



UNIVERSIDAD DE MÁLAGA  
FACULTAD DE FILOSOFÍA Y LETRAS

**El lenguaje económico en los tiempos de  
la crisis global:  
un estudio longitudinal de análisis de  
sentimiento**

**TESIS DOCTORAL**

presentada por  
Javier Fernández Cruz  
para optar al título de Doctor por la Universidad de Málaga


Director:  
Dr. Antonio Jesús Moreno Ortiz  
Departamento de Filología Inglesa, Francesa y Alemana.

Programa de doctorado en Lingüística, Literatura y Traducción  
Málaga, 2019



UNIVERSIDAD  
DE MÁLAGA

AUTOR: Javier Fernandez Cruz

 <http://orcid.org/0000-0002-9007-9976>

EDITA: Publicaciones y Divulgación Científica. Universidad de Málaga



Esta obra está bajo una licencia de Creative Commons Reconocimiento-NoComercial-SinObraDerivada 4.0 Internacional:

<http://creativecommons.org/licenses/by-nc-nd/4.0/legalcode>

Cualquier parte de esta obra se puede reproducir sin autorización pero con el reconocimiento y atribución de los autores.

No se puede hacer uso comercial de la obra y no se puede alterar, transformar o hacer obras derivadas.

Esta Tesis Doctoral está depositada en el Repositorio Institucional de la Universidad de Málaga (RIUMA): [riuma.uma.es](http://riuma.uma.es)





UNIVERSIDAD  
DE MÁLAGA

AUTOR: Javier Fernández Cruz  
ORCID: 0000-0002-9007-9976



Esta obra está bajo una licencia de Creative Commons ReconocimientoNoComercial-SinObraDerivada 4.0 Internacional:

<http://creativecommons.org/licenses/by-nc-nd/4.0/legalcode>

Cualquier parte de esta obra se puede reproducir sin autorización, pero con el reconocimiento y atribución del autor.

No se puede hacer uso comercial de la obra y no se puede alterar, transformar o hacer obras derivadas.



Málaga a 16 de julio de 2019

Antonio Jesús Moreno Ortiz, profesor de la Facultad de Filosofía y Letras (Universidad de Málaga),

### HACE CONSTAR

Que JAVIER FERNÁNDEZ CRUZ, con DN/NIE/pasaporte 74.942.806-C es estudiante de doctorado del Programa de Doctorado "Lingüística, Literatura y Traducción", con matrícula activa, y que ha realizado bajo mi dirección, la Tesis Doctoral titulada

"EL LENGUAJE ECONÓMICO EN LOS TIEMPOS DE LA CRISIS GLOBAL: UN ESTUDIO  
LONGITUDINAL DE ANÁLISIS DE SENTIMIENTO"

Revisado el presente trabajo estimo que reúne los requisitos establecidos según la normativa vigente. Por lo tanto, **AUTORIZO** la admisión a trámite y defensa pública de esta Tesis Doctoral para optar al grado de Doctor en la Universidad de Málaga.

Y para que así conste, lo firmo en Málaga a 16 de julio de 2019,



Fdo. Antonio Jesús Moreno Ortiz



Campus de Teatinos s/n. 29071 Málaga  
Tel.: 952 13 16 83/1684/1685/1687/3432/3435 - Fax: 952 13 18 23





*A mi madre Carmina Cruz del Águila y a mi padre Rafael Fernández Arce.*

*Todo lo que somos es gracias a sus corazones y manos.*



# Índice

<b>ÍNDICE.....</b>	<b>VII</b>
<b>ÍNDICE DE TABLAS .....</b>	<b>X</b>
<b>ÍNDICE DE FIGURAS.....</b>	<b>XII</b>
<b>AGRADECIMIENTOS.....</b>	<b>XV</b>
<b>1. INTRODUCCIÓN .....</b>	<b>1</b>
1.1 PLANTEAMIENTO DEL PROBLEMA .....	1
1.2 JUSTIFICACIÓN Y MOTIVACIÓN.....	6
1.3 PREGUNTAS FORMALES DE INVESTIGACIÓN .....	9
1.4 DEFINICIÓN DE OBJETIVOS .....	9
1.4.1 Objetivo general .....	9
1.4.2 Objetivos específicos.....	10
1.5 ORGANIZACIÓN DE LA TESIS .....	10
<b>2. CONTEXTO DE LA INVESTIGACIÓN: EL MUNDO Y EL LENGUAJE ANTE LA CRISIS GLOBAL</b>	<b>13</b>
2.1 CONTEXTUALIZACIÓN HISTÓRICA, POLÍTICA Y SOCIAL DE LA CRISIS (2007-2018).....	13
2.1.1 La crisis de las hipotecas <i>subprime</i> y la crisis del crédito.....	13
2.1.2 La crisis de la deuda .....	17
2.1.3 Política y ciudadanía ante la <i>Gran Recesión</i> .....	24
2.2 LA INFLUENCIA DE LA CRISIS Y LA PRENSA EN EL LENGUAJE ECONÓMICO.....	29
2.3 CONTENIDOS EN RED Y OPINIÓN.....	36
2.4 CONCLUSIONES .....	38
<b>3. EL LENGUAJE EVALUATIVO .....</b>	<b>41</b>
3.1 INTRODUCCIÓN.....	41
3.2 TÉRMINOS UTILIZADOS PARA ESTUDIAR EL LENGUAJE EVALUATIVO .....	44
3.3 DEFINICIÓN DE EVALUACIÓN .....	48
3.4 FUNCIONES DE LA EVALUACIÓN.....	50
3.5 IDENTIFICACIÓN DE SEGMENTOS EVALUATIVOS .....	53
3.5.1 Marcadores evaluativos .....	53
3.5.2 Evaluación y textualidad.....	56
3.5.3 Identificación de patrones en el género periodístico .....	57
3.6 GRAMÁTICA LOCAL: FORMA Y FUNCIÓN DE LA EVALUACIÓN .....	60
3.7 LA TEORÍA DE LA VALORACIÓN.....	65
3.7.1 La teoría de la valoración situada en el modelo sistémico funcional .....	67
3.7.2 Sistema general de la teoría de la valoración .....	69
3.7.3 Actitud .....	70

3.7.4 Compromiso.....	79
3.7.5 Gradación.....	87
3.7.6 Procesamiento de la valoración y análisis de sentimiento .....	91
3.8 EL ESTUDIO DE LA OPINIÓN DESDE LA PERSPECTIVA DEL ANÁLISIS DE SENTIMIENTO .....	92
3.8.1 Aproximación a la expresión de la opinión.....	93
3.8.2 Opinión y análisis de sentimiento .....	95
3.8.3 El <i>Nuevo</i> modelo dinámico del lenguaje evaluativo.....	99
3.9 CONCLUSIONES .....	102
<b>4. LENGUAJES ESPECIALIZADOS Y TERMINOLOGÍA .....</b>	<b>105</b>
4.1 INTRODUCCIÓN.....	105
4.2 LOS LENGUAJES ESPECIALIZADOS .....	112
4.3. TEORÍAS DE LA TERMINOLOGÍA .....	121
4.3.1 La teoría general de la terminología de Wüster .....	121
4.3.2 Las dimensiones lingüística y comunicativa de la terminología de Juan C. Sager .....	126
4.3.3 La teoría comunicativa de la terminología de Teresa Cabré.....	131
4.3.4 La teoría sociocognitiva de la terminología de Rita Temmerman.....	135
4.4 LOS TÉRMINOS: DENOMINACIONES Y FORMAS LINGÜÍSTICAS .....	140
4.5 LA GESTIÓN TERMINOLÓGICA .....	143
4.5.1 Enfoque textual de la terminología.....	143
4.5.2 Las bases de conocimiento terminológicas.....	147
4.5.3 Los tesauros .....	148
4.5.4 Las ontologías .....	150
4.5.5 Extracción terminológica.....	152
4.6 CARACTERÍSTICAS DEL LENGUAJE PERIODÍSTICO EN LENGUA INGLESA .....	157
4.7 CARACTERÍSTICAS DEL LENGUAJE ECONÓMICO EN LENGUA INGLESA.....	164
4.8 CONCLUSIONES .....	169
<b>5. ANÁLISIS DE SENTIMIENTO .....</b>	<b>171</b>
5.1 INTRODUCCIÓN Y DEFINICIÓN .....	171
5.2 PROPIEDADES DEL ANÁLISIS DE SENTIMIENTO .....	174
5.3 DESCRIPCIÓN DE LOS PRINCIPALES NIVELES DE CLASIFICACIÓN .....	176
5.3.1 Clasificación de sentimiento a nivel de documento .....	176
5.3.2 Clasificación de sentimiento a nivel oración .....	177
5.3.3 Clasificación de sentimiento a nivel aspecto .....	180
5.4 ENFOQUES DE ANÁLISIS DE SENTIMIENTO .....	185
5.4.1 Aprendizaje automático ( <i>Machine learning</i> ).....	185
5.4.2 Enfoque basado en léxico .....	189
5.4.3 Enfoques híbridos.....	192

5.4.3 Ventajas, desventajas y desafíos de los distintos enfoques.....	193
5.5 MODIFICADORES CONTEXTUALES.....	194
5.6 COMPILACIÓN DE LEXICONES DE SENTIMIENTO.....	198
5.6.1 Enfoque basado en diccionario.....	199
5.6.2 Enfoque basado en corpus.....	200
5.6.3 Lexicones de sentimiento en la literatura.....	201
5.7 CLASIFICACIÓN DE SENTIMIENTO EN DOMINIOS ESPECÍFICOS .....	203
5.7.1 Lexicones de dominio económico-financiero .....	206
5.7.2 Características del dominio económico-financiero en relación con el análisis de sentimiento .....	209
5.7.3 Sistemas de análisis de sentimiento económico-financiero .....	211
5.8 CAMBIOS EN LA ORIENTACIÓN SEMÁNTICA Y ANÁLISIS DE SENTIMIENTO .....	214
5.9 APLICACIONES COMERCIALES DE ANÁLISIS DE SENTIMIENTO .....	218
5.10 CONCLUSIONES.....	220
<b>6. MÉTODOS DE INVESTIGACIÓN .....</b>	<b>223</b>
6.1 PARADIGMAS Y ENFOQUES METODOLÓGICOS.....	223
6.2 DISEÑO DE LA INVESTIGACIÓN.....	228
6.3 MUESTRA: EL ESMEALDAS CORPUS OF BUSINESS ENGLISH.....	233
6.3.1 Recopilación del corpus.....	239
6.3.2 Descripción del corpus .....	240
6.4 HERRAMIENTAS.....	242
6.4.1 <i>Lingmotif</i> , un software de análisis de sentimiento basado en léxico .....	242
6.4.1.1 Descripción del software.....	242
6.4.1.2 El lexicón nuclear de sentimiento.....	245
6.4.1.3 Reglas contextuales .....	247
6.4.1.4 Lexicones <i>plug-in</i> .....	249
6.4.2 <i>The SketchEngine</i> , una suite de gestión de corpus y análisis textual.....	250
6.5 LA CONSTRUCCIÓN DE <i>LEXIECON</i> , UN LEXICÓN DE DOMINIO ECONÓMICO FINANCIERO PARA <i>LINGMOTIF</i> .....	251
6.5.1 Extracción terminológica.....	251
6.5.2 Justificación y fundamentos de <i>LexiEcon</i> .....	253
6.5.3 Etiquetado de entradas.....	256
6.5.4 Evaluación de <i>LexiEcon</i> .....	259
6.6 CONCLUSIONES .....	262
<b>7. ANÁLISIS DE DATOS.....</b>	<b>265</b>
7.1 TÉRMINO 1: CREDIT.....	267
7.2 TÉRMINO 2: <i>DEBT</i> .....	282

7.3 TÉRMINO 3: <i>INFLATION</i> .....	301
7.4 TÉRMINO 4: <i>HOUSING</i> .....	317
7.5 TÉRMINO 5: <i>MARKETS</i> .....	330
7.6 TÉRMINO 6: <i>INVESTOR</i> .....	350
7.7 TÉRMINO 7: <i>PUBLIC</i> .....	363
7.8 TÉRMINO 8: <i>REFORM</i> .....	382
7.9 TÉRMINO 9: <i>SPENDING</i> .....	399
7.10 TÉRMINO 10: <i>COMPANY</i> .....	414
<b>8. RESULTADOS Y DISCUSIÓN.....</b>	<b>425</b>
<b>9. CONCLUSIONES FINALES.....</b>	<b>435</b>
9.1 CONSECUCIÓN DE OBJETIVOS.....	435
9.2 LIMITACIONES Y PERSPECTIVAS DE FUTURO.....	443
<b>REFERENCIAS .....</b>	<b>445</b>
<b>ANEXO 1. RESUMEN DE LA TESIS.....</b>	<b>I</b>
<b>ANEXO 2. THESIS SUMMARY.....</b>	<b>X</b>
<b>ANEXO 3. CONCLUSIONS.....</b>	<b>XIX</b>

## Índice de tablas

<b>TABLA 1</b> RESUMEN DE CATEGORÍAS DE AFECTO Y EVIDENCIALIDAD. BIBER Y FINNEGAN (1989) .....	56
<b>TABLA 2</b> TIPOS DE AFECTO SEGÚN MARTIN Y WHITE (2005).....	72
<b>TABLA 3</b> EJEMPLOS DE AFECTO IRREALIS (ADAPTADO DE MARTIN Y WHITE, 2005) .....	73
<b>TABLA 4</b> EJEMPLOS DE IN/FELICIDAD (ADAPTADO DE MARTIN Y WHITE, 2005). .....	74
<b>TABLA 5</b> RESUMEN DEL SISTEMA DE APRECIACIÓN (MARTIN Y WHITE, 2005), TRADUCCIÓN. ....	77
<b>TABLA 6</b> CLASIFICACIÓN DE MEDIDAS ESTADÍSTICAS EN LAS DIMENSIONES ESTADÍSTICA Y LINGÜÍSTICA (PAZIENZA, PENNANCCHIOTTI, & ZANZOTTO, 2005). ....	155
<b>TABLA 7</b> RESUMEN DE LOS PRINCIPALES LEXICONES DE SENTIMIENTO .....	203
<b>TABLA 8</b> RESUMEN DE LOS PRINCIPALES ESTUDIOS DE ANÁLISIS DE SENTIMIENTO RELACIONADO CON EL DOMINIO ECONÓMICO-FINANCIERO (KIRSHNAMOORTHY, 2017) .....	214
<b>TABLA 9</b> LISTA DE TÉRMINOS DE ANÁLISIS Y NÚMERO DE ORACIONES ANALIZADAS. ....	232
<b>TABLA 10</b> DATOS DEL ECBE, DESGLOSE POR AÑO DE LOS DATOS SOBRE EL NÚMERO TOTAL DE PALABRAS Y DE ARTÍCULOS RECOPIADOS .....	240
<b>TABLA 11</b> DESCRIPCIÓN DEL CORPUS ECBE SEGÚN FUENTE .....	241
<b>TABLA 12</b> EJEMPLO DE LAS DOS PRIMERAS ENTRADAS DEL DATASET DEL TÉRMINO DEBT.....	243
<b>TABLA 13</b> EJEMPLOS DE DIRECCIONALIDAD EN EL INDICADOR “EMPLOYMENT” .....	255
<b>TABLA 14</b> RESUMEN ESTADÍSTICO DE <i>LEXIÉCON</i> .....	259
<b>TABLA 15</b> MATRICES DE CONFUSIÓN (FILA = <i>REFERENCIA</i> ; COLUMNA = <i>TEST</i> ) .....	260

<b>TABLA 16</b> RESUMEN DE VERDADEROS POSITIVOS, FALSOS POSITIVOS, VERDADEROS NEGATIVOS Y FALSOS NEGATIVOS .....	260
<b>TABLA 17</b> CIFRAS DE RENDIMIENTO .....	260
<b>TABLA 18</b> RESUMEN ESTADÍSTICO DEL ANÁLISIS DE LAS ORACIONES QUE INCLUYEN <i>CREDIT</i> .....	267
<b>TABLA 19</b> PRINCIPALES COLOCACIONES DE “ <i>CREDIT</i> ” DURANTE EL PRIMER INTERVALO .....	270
<b>TABLA 20</b> PRINCIPALES COLOCACIONES DE “ <i>CREDIT</i> ” DURANTE EL SEGUNDO INTERVALO .....	274
<b>TABLA 21</b> PRINCIPALES COLOCACIONES DE “ <i>CREDIT</i> ” DURANTE EL TERCER INTERVALO .....	279
<b>TABLA 22</b> RESUMEN ESTADÍSTICO DEL ANÁLISIS DE LAS ORACIONES QUE INCLUYEN “ <i>DEBT</i> ” .....	285
<b>TABLA 23</b> PRINCIPALES COLOCACIONES DE “ <i>DEBT</i> ” DURANTE EL PRIMER INTERVALO .....	286
<b>TABLA 24</b> PRINCIPALES COLOCACIONES DE “ <i>DEBT</i> ” DURANTE EL SEGUNDO INTERVALO .....	290
<b>TABLA 25</b> PRINCIPALES COLOCACIONES DE “ <i>DEBT</i> ” DURANTE EL TERCER INTERVALO .....	298
<b>TABLA 26</b> RESUMEN ESTADÍSTICO DEL ANÁLISIS DE LAS ORACIONES QUE INCLUYEN “ <i>INFLATION</i> ” .....	304
<b>TABLA 27</b> PRINCIPALES COLOCACIONES DE “ <i>INFLATION</i> ” DURANTE EL PRIMER INTERVALO .....	305
<b>TABLA 28</b> PRINCIPALES COLOCACIONES DE “ <i>INFLATION</i> ” DURANTE EL SEGUNDO INTERVALO .....	309
<b>TABLA 29</b> PRINCIPALES COLOCACIONES DE “ <i>INFLATION</i> ” DURANTE EL TERCER INTERVALO .....	314
<b>TABLA 30</b> RESUMEN ESTADÍSTICO DEL ANÁLISIS DE LAS ORACIONES QUE INCLUYEN “ <i>HOUSING</i> ” .....	319
<b>TABLA 31</b> PRINCIPALES COLOCACIONES DE “ <i>HOUSING</i> ” DURANTE EL PRIMER INTERVALO .....	321
<b>TABLA 32</b> PRINCIPALES COLOCACIONES DE “ <i>HOUSING</i> ” DURANTE EL SEGUNDO INTERVALO .....	327
<b>TABLA 33</b> RESUMEN ESTADÍSTICO DEL ANÁLISIS DE LAS ORACIONES QUE INCLUYEN “ <i>MARKETS</i> ” .....	333
<b>TABLA 34</b> PRINCIPALES COLOCACIONES DE “ <i>MARKETS</i> ” DURANTE EL PRIMER INTERVALO .....	334
<b>TABLA 35</b> PRINCIPALES COLOCACIONES DE “ <i>MARKETS</i> ” DURANTE EL SEGUNDO INTERVALO .....	340
<b>TABLA 36</b> PRINCIPALES COLOCACIONES DE “ <i>MARKETS</i> ” DURANTE EL TERCER INTERVALO .....	345
<b>TABLA 37</b> PRINCIPALES COLOCACIONES DE “ <i>MARKETS</i> ” DURANTE EL CUARTO INTERVALO .....	348
<b>TABLA 38</b> RESUMEN ESTADÍSTICO DEL ANÁLISIS DE LAS ORACIONES QUE INCLUYEN “ <i>INVESTOR</i> ” .....	352
<b>TABLA 39</b> PRINCIPALES COLOCACIONES DE “ <i>INVESTOR</i> ” DURANTE EL PRIMER INTERVALO .....	353
<b>TABLA 40</b> PRINCIPALES COLOCACIONES DE “ <i>INVESTOR</i> ” DURANTE EL SEGUNDO INTERVALO .....	356
<b>TABLA 41</b> PRINCIPALES COLOCACIONES DE “ <i>INVESTOR</i> ” DURANTE EL TERCER INTERVALO .....	361
<b>TABLA 42</b> RESUMEN ESTADÍSTICO DEL ANÁLISIS DE LAS ORACIONES QUE INCLUYEN EL ADJETIVO “ <i>PUBLIC</i> ” .....	365
<b>TABLA 43</b> PRINCIPALES COLOCACIONES DE “ <i>PUBLIC</i> ” DURANTE EL PRIMER INTERVALO .....	366
<b>TABLA 44</b> PRINCIPALES COLOCACIONES DE “ <i>PUBLIC</i> ” DURANTE EL SEGUNDO INTERVALO .....	370
<b>TABLA 45</b> PRINCIPALES COLOCACIONES DE “ <i>PUBLIC</i> ” DURANTE EL TERCER INTERVALO .....	375
<b>TABLA 46</b> PRINCIPALES COLOCACIONES DE “ <i>PUBLIC</i> ” DURANTE EL CUARTO INTERVALO .....	379
<b>TABLA 47</b> RESUMEN ESTADÍSTICO DEL ANÁLISIS DE LAS ORACIONES QUE INCLUYEN “ <i>REFORM</i> ” .....	384
<b>TABLA 48</b> PRINCIPALES COLOCACIONES DE “ <i>REFORM</i> ” DURANTE EL PRIMER INTERVALO .....	385
<b>TABLA 49</b> PRINCIPALES COLOCACIONES DE “ <i>REFORM</i> ” DURANTE EL SEGUNDO INTERVALO DE ESTUDIO .....	388
<b>TABLA 50</b> PRINCIPALES COLOCACIONES DE “ <i>REFORM</i> ” DURANTE EL TERCER INTERVALO DE ESTUDIO .....	392
<b>TABLA 51</b> PRINCIPALES COLOCACIONES DE “ <i>REFORM</i> ” DURANTE EL CUARTO INTERVALO .....	396
<b>TABLA 52</b> RESUMEN ESTADÍSTICO DEL ANÁLISIS DE LAS ORACIONES QUE INCLUYEN “ <i>SPENDING</i> ” .....	401
<b>TABLA 53</b> PRINCIPALES COLOCACIONES DE “ <i>SPENDING</i> ” DURANTE EL PRIMER INTERVALO DE ESTUDIO .....	402





<b>TABLA 54</b> PRINCIPALES COLOCACIONES DE “SPENDING” DURANTE EL SEGUNDO INTERVALO DE ESTUDIO .....	407
<b>TABLA 55</b> PRINCIPALES COLOCACIONES DE “SPENDING” DURANTE EL TERCER INTERVALO DE ESTUDIO ..	411
<b>TABLA 56</b> RESUMEN ESTADÍSTICO DEL ANÁLISIS DE LAS ORACIONES QUE INCLUYEN “COMPANY” .....	415
<b>TABLA 57</b> PRINCIPALES COLOCACIONES DE “COMPANY” DURANTE EL PRIMER INTERVALO DE ESTUDIO ..	417
<b>TABLA 58</b> PRINCIPALES COLOCACIONES DE “COMPANY” DURANTE EL SEGUNDO INTERVALO DE ESTUDIO .....	418
<b>TABLA 59</b> PRINCIPALES COLOCACIONES DE “COMPANY” DURANTE EL TERCER INTERVALO DE ESTUDIO ..	423

## Índice de figuras

<b>FIGURA 1</b> SISTEMA GENERAL DE LA TEORÍA DE LA VALORACIÓN.....	70
<b>FIGURA 2</b> NUEVO MODELO DINÁMICO DEL LENGUAJE EVALUATIVO (BENAMARA ET AL., 2017).....	100
<b>FIGURA 3</b> MODELO DE ACTUALIZACIÓN DEL DISCURSO EVALUATIVO (BENAMARA ET AL., 2017) .....	101
<b>FIGURA 4</b> MODELO DE INSTANCIACIÓN DE LA EVALUACIÓN SEGÚN LOS DISTINTOS NIVELES LINGÜÍSTICOS (BENAMARA ET AL., 2017).....	101
<b>FIGURA 5</b> MODELOS DE RELACIÓN ENTRE LENGUA GENERAL Y LENGUAJES ESPECIALIZADOS (SANTAMARÍA PÉREZ, 2009) .....	114
<b>FIGURA 6</b> PUNTOS DE CONVERGENCIA LÉXICA ENTRE EL DOMINIO FINANCIERO Y LOS TEXTOS DIVULGATIVOS SOBRE LAS FINANZAS. ....	120
<b>FIGURA 7</b> ESQUEMA DE LA TEORÍA DE LAS PUERTAS (CABRÉ, 2002).....	133
<b>FIGURA 8</b> RESUMEN DE LAS PRINCIPALES CARACTERÍSTICAS DEL ANÁLISIS DE SENTIMIENTO .....	176
<b>FIGURA 9</b> FÓRMULAS PARA CALCULAR LA <i>PRECISIÓN</i> , LA <i>EXHAUSTIVIDAD</i> Y EL <i>VALOR-F</i> .....	189
<b>FIGURA 10</b> FÓRMULA DE CÁLCULO DEL SENTIMIENTO DE LINGMOTIF CUANDO LAS UNIDADES POSITIVAS SUPERAN A LAS NEGATIVAS .....	244
<b>FIGURA 11</b> FÓRMULA DE CÁLCULO DEL SENTIMIENTO DE LINGMOTIF CUANDO LAS UNIDADES NEGATIVAS SUPERAN A LAS POSITIVAS .....	244
<b>FIGURA 12</b> UNA ENTRADA TÍPICA SEGÚN LAS ESPECIFICACIONES DE <i>LINGMOTIF</i> .....	246
<b>FIGURA 13</b> EJEMPLO DE UNA ENTRADA DEL LÉXICO <i>PLUG-IN</i> .....	250
<b>FIGURA 14</b> LAS DIEZ PRIMERAS ENTRADAS DEL CONJUNTO DE DATOS DE “DEBT” TRAS SER PROCESADAS POR <i>LINGMOTIF</i> .....	266
<b>FIGURA 15</b> GRÁFICO LONGITUDINAL DE LAS MEDIAS ANUALES DE LAS PUNTUACIONES DE SENTIMIENTO PARA “CREDIT” Y LA FRECUENCIA RELATIVA DEL TÉRMINO POR CADA 1000 TOKENS .....	269
<b>FIGURA 16</b> EJEMPLO DE ENTRADAS RELACIONADAS CON “DEBT” EN <i>LEXIECON</i> .....	282
<b>FIGURA 17</b> GRÁFICO LONGITUDINAL DE LAS MEDIAS ANUALES DE LAS PUNTUACIONES DE SENTIMIENTO PARA “DEBT” Y LA FRECUENCIA RELATIVA DEL TÉRMINO POR CADA 1000 TOKENS.....	284
<b>FIGURA 18</b> EJEMPLO DE ENTRADAS RELACIONADAS CON “INFLATION” EN <i>LEXIECON</i> .....	302
<b>FIGURA 19</b> GRÁFICO LONGITUDINAL DE LAS MEDIAS ANUALES DE LAS PUNTUACIONES DE SENTIMIENTO PARA “INFLATION” Y LA FRECUENCIA RELATIVA DEL TÉRMINO POR CADA 1000 TOKENS .....	304

---

<b>FIGURA 20</b> GRÁFICO LONGITUDINAL DE LAS MEDIAS ANUALES DE LAS PUNTUACIONES DE SENTIMIENTO PARA “HOUSING” Y LA FRECUENCIA RELATIVA DEL TÉRMINO POR CADA 1000 TOKENS. ....	318
<b>FIGURA 21</b> GRÁFICO LONGITUDINAL DE LAS MEDIAS ANUALES DE LAS PUNTUACIONES DE SENTIMIENTO PARA “MARKETS” Y LA FRECUENCIA RELATIVA DEL TÉRMINO POR CADA 1000 TOKENS .....	332
<b>FIGURA 22</b> GRÁFICO LONGITUDINAL DE LAS MEDIAS ANUALES DE LAS PUNTUACIONES DE SENTIMIENTO PARA “INVESTOR” Y LA FRECUENCIA RELATIVA DEL TÉRMINO POR CADA 1000 TOKENS.....	351
<b>FIGURA 23</b> GRÁFICO LONGITUDINAL DE LAS MEDIAS ANUALES DE LAS PUNTUACIONES DE SENTIMIENTO PARA EL ADJETIVO “PUBLIC” Y LA FRECUENCIA RELATIVA DEL TÉRMINO POR CADA 1000 TOKENS .	364
<b>FIGURA 24</b> GRÁFICO LONGITUDINAL DE LAS MEDIAS ANUALES DE LAS PUNTUACIONES DE SENTIMIENTO PARA “REFORM” Y LA FRECUENCIA RELATIVA DEL TÉRMINO POR CADA 1000 TOKENS .....	383
<b>FIGURA 25</b> NÚMERO DE BÚSQUEDA DE NOTICIAS DEL TÉRMINO “REFORM” EN <i>GOOGLE</i> DURANTE EL INTERVALO 2008-2015 ( <i>GOOGLE TRENDS</i> ).....	384
<b>FIGURA 26</b> GRÁFICO LONGITUDINAL DE LAS MEDIAS ANUALES DE LAS PUNTUACIONES DE SENTIMIENTO PARA “SPENDING” Y LA FRECUENCIA RELATIVA DEL TÉRMINO POR CADA 1000 TOKENS. ....	400
<b>FIGURA 27</b> GRÁFICO LONGITUDINAL DE LAS MEDIAS ANUALES DE LAS PUNTUACIONES DE SENTIMIENTO PARA “COMPANY” Y LA FRECUENCIA RELATIVA DEL TÉRMINO POR CADA 1000 TOKENS.....	415
<b>FIGURA 28</b> EVOLUCIÓN DE LA FRECUENCIA RELATIVA DE LOS DISTINTOS TÉRMINOS DE ESTUDIO .....	428
<b>FIGURA 29</b> EVOLUCIÓN DEL SENTIMIENTO DE LOS DISTINTOS TÉRMINOS DE ESTUDIO.....	428



## Agradecimientos

Escribir esta tesis se ha asemejado por momentos a un largo viaje por mares procelosos. Por ello creo que es injusto atribuirme el mérito exclusivo de descubrir las artes y destrezas que implica toda navegación, desde la botadura al arribo a puerto. Es necesario superar el tópico de la soledad del doctorando, pues este viaje de cinco años, que veces se antojan como cinco vidas, lo he realizado bien acompañado: hay un poco de muchas personas a lo largo de este texto.

Comenzaré reconociendo la excelente labor de Antonio Moreno-Ortiz, mi director de tesis, quien me animó a pergeñar el presente trabajo y me reveló los desafíos de esta forma tan especial de aproximarnos al lenguaje. Más allá de acercarme a las distintas teorías, metodologías, enfoques y técnicas, gracias a él he aprendido que al final se trata de tener pasión por lo que se emprende y que no hay límites en el aprendizaje, sino un ejercicio de constancia, de horas y horas de dedicación. En la Universidad de Málaga he contado con el ejemplo también de la profesora Chantal Pérez y de la excelente disposición de Rosario Arias, coordinadora del programa de doctorado. Agradezco igualmente la labor del resto del equipo de Tecnolengua, especialmente de Jorge Galván, por la ayuda técnica a la hora de procesar los datos. Debo mencionar explícitamente a los evaluadores externos de esta tesis, Olivier Méric y Haydée Ramírez, por sus necesarias críticas, comentarios y aportaciones.

Esta travesía ha sido la historia de un eterno retorno que durante cinco años me ha arrastrado y devuelto del Mediterráneo a las verdes aguas del Pacífico. Doy las gracias a la República del Ecuador, que me abrió las puertas de par en par y me permitió recorrer sus playas, ríos, manglares, volcanes y selvas. Me sigue pareciendo extraordinaria la vocación de progreso del país, que supo encender un lucerito que se divisaba desde tanta distancia y en una época en la que académica y laboralmente atravesaba una larga noche en España. Y así arribé a Esmeraldas, tierra verde y patria chica del pueblo afroecuatoriano, donde en la colina de Santa Cruz se encuentra la Pontificia Universidad Católica del Ecuador Sede Esmeraldas. Este trabajo se ha podido llevar a cabo gracias al apoyo de su pro rector Aitor Urbina, la directora académica Liris Bernabé e Ingrid Grijalva, la directora de la Escuela de Lingüística Aplicada. Junto a mis compañeros en la sede (Silvia Cabrera, Marjorie Perlaza, Rebeca Naranjo, Tony Rodríguez, Cristina Marmolejo, Solange España, José Luis Romero, Cristóbal Baixauli, Gabriela Cisneros y

tanta gente a la que injustamente omito), han hecho sencillo lo imposible: me han despejado el camino ante los múltiples y remotamente previsibles obstáculos que me he encontrado durante este viaje. No puedo olvidar los paseos por el Malecón y la tregua que pocas veces dio la brisa de la tarde durante este verano de dos años y medio en los que veíamos pasar la vida frente a nuestra terraza asomada a la playa de Las Palmas.

Del Ecuador es también el desarrollo de LexiEcon. Este se pudo llevar a cabo con los fondos obtenidos a partir del Concurso de Investigadores PUCESE 2017, que sirvieron para ampliar los fondos de nuestra biblioteca y por añadidura pusieron en mi camino la quasi omnisapientia de Ander Mujica Uranga. Además de ser un experto en el dominio de las finanzas, me guio siempre de buena gana a través de las lúgubres tardes equinocciales en nuestro espacio de trabajo, mientras demostraba que siempre se podía hilar más sutil y, así, ampliar y afinar los pequeños matices de un dominio a veces tan obtuso como es lenguaje de la economía y las finanzas.

El capítulo que versa sobre lenguajes especializados sería totalmente distinto sin la inspiración y las aportaciones de Laurent Gautier, mi responsable científico en mi primaveral estancia dijonesa en la Maison des Sciences de l'Homme de la Universidad de Borgoña. Gracias a su buena disposición y a las aportaciones de quienes allí conocí pude ampliar enormemente las perspectivas y aplicar en otros dominios las técnicas aprendidas durante este trabajo. El equipo dijónés me ayudó a profundizar en la importancia de la observación detallada de las fuentes en el discurso y las enormes posibilidades que ofrece la terminología sensorial. Con una importancia equivalente, aprendí a disfrutar si cabe aún más de la excelencia de los vinos de la Côte d'Or y la gastronomía borgoñona. Ya que el vino se disfruta mejor en compañía, he de mencionar (o brindar con ellas) a todas las personas que tan bien me ayudaron y acogieron tanto académica como socialmente en la Universidad de Borgoña y la Universidad del Franco Condado: Mariele Mancebo, Maxime Tissier-Sevestre, Hédi Maazaoui, Rémi Badian, Matthieu Bach, Cyril Nguyen Van, Virginie Lethier, Marion Landre, Fabien Labondance, Candice Lemaire y Judit Rodrigues-Balbuena.

A mis amigos (Dani, Curro, Rafa, Nacho, Ale, José, Ángel, Gonzalo, Kike, Marta, Juan, Santi, Ramón, Germán, Yendéh, Carmen, Ana, Rocío, Rosa, Toret, Pitu, Adrián, Mamen, Manu, Viviana, Geovanna, Cristian, Gaby, Iván, Fernanda, espero que me puedan perdonar las omisiones dolorosas) les doy las gracias por sus impresiones, visiones y sobre todo por las experiencias compartidas. Con otros, además de ser amigos,

he compartido hogar a lo largo de mis sempiternas mudanzas estos años (Falete, Jasmanny, Francis, Nuria, Afo, Pedro y Jorge), y me han aguantado con buen humor mis paseos por la casa a deshora, además de obligarme a hacer cosas tan esenciales como tomarnos una cerveza a la fresca o comer una pizza juntos. Y, cómo no, mi agradecimiento a Ana Luna, por acompañarme así de bien y hacerme tan ligeros estos últimos meses de interminables revisiones.

Finalmente doy las gracias a mis padres, Carmina y Rafael, y a mi hermana Alba, por enseñarme a mirar la vida con las gafas de la humildad y por haber aceptado de grado mis largas ausencias durante los buenos y los malos momentos, siempre aconsejándome y poniendo todo su corazón en su apoyo. Más allá de los clichés de la igualdad de oportunidades, no es tan sencillo que el nieto de unos jornaleros llegue a defender una tesis doctoral. Sin unos padres que se despertasen antes que el sol durante toda una vida para ir a la fábrica, a limpiar otros hogares, al campo o al camión y sin una universidad pública universal y asequible, aunque tan solo una vez me concedió beca, no estaría escribiendo estas líneas. Nunca habrá palabras suficientes para describir el peso de los valores que he aprendido de mis mayores.

*Con un pie en la Cala Varques y otro en el varadero de la Playa de San Andrés, en julio de 2019 (uno de los veranos más calurosos que se recuerdan).*

Ítaca te brindó el espléndido viaje.  
Sin ella no te habrías puesto en camino.  
No puede ya ofrecerte nada más.

Y si pobre la encuentras, Ítaca no te engañó.  
Con la sabiduría que has alcanzado, con tu experiencia,  
ya habrás comprendido qué significan las Ítacas.

*Ítaca* (1911), Cavafis (1863-1933)



# 1. Introducción

## 1.1 Planteamiento del problema

¿Han influido las consecuencias de la crisis económica mundial en la forma en que usamos el lenguaje? Los cambios en las condiciones económicas y las consecuencias sociales que conllevan pueden dar lugar a cambios que en última instancia se reflejan en las expresiones lingüísticas; particularmente, en la expresión de la polaridad. Si a comienzos del verano de 2019 alguien introdujera en un buscador la palabra “rescate”, este encontraría principalmente noticias sobre las ayudas a los productores estadounidenses en plena guerra comercial con China. Un año antes, esta búsqueda hablaría de la ONG *Open Arms* y el rescate de seres humanos en las costas de Libia durante la dramática crisis migratoria en el Mediterráneo. Si esta acción se hubiese realizado en 2008, muy probablemente se estaría hablando sobre el hundimiento del sistema financiero mundial.

Más allá de ser una negra analogía, este ejemplo nos sirve para ilustrar la profunda carga subjetiva que tienen los textos que cada mañana encontramos al abrir el periódico. Los sentimientos que en cada caso nos evoca la palabra “rescate” son bien distintos. Por ello cabe preguntarse: ¿es la comunicación y, en consecuencia, la terminología que en ella se emplea, un instrumento para proyectar el sentimiento de la política económica de manera subliminal, el cual hoy en día aún no alcanzamos a dominar, pero del que comenzamos a vislumbrar su potencial? El análisis de sentimiento o análisis de opinión depende profundamente del contexto y, además, del dominio específico en cuestión. Las lenguas de especialidad tienen una terminología, expresiones y construcciones propias, que distan de ser inalterables, y que además reflejan el espíritu de los tiempos. Todos estos factores salen a la luz en la manera en la que se expresan las evaluaciones y opiniones en los textos.

Continuando en 2008, el día 9 de agosto el Banco Central Europeo y la Reserva Federal estadounidense debieron intervenir en un sistema bancario congelado al que inyectaron 100.000 millones de euros. Fue el día en que “cambió el mundo”, según Adam



Appelgarth, entonces cabeza visible del banco británico *Northern Rock* y una de las veinticinco caras del desastre económico publicadas por el diario británico *The Guardian* (Finch, 2009). A partir de ese verano, la actividad del mercado mayorista comenzó a caer en picado y meses después tuvo lugar el evento icónico de la presente crisis mundial: el 15 de septiembre de 2008 quiebra el banco de inversión estadounidense *Lehman Brothers* y con esto se inicia un episodio global que una década después sigue repercutiendo en nuestra vida cotidiana. Aunque para *Lehman* no hubo rescate, fue la primera vez que el público empezó a oír la palabra “rescate” en un contexto económico, y que connotaba, más que un umbral de esperanza, un halo de preocupación.

Este otoño fue clave en términos políticos, económicos y, como no, lingüísticos. Las tres semanas que siguieron al colapso de *Lehman Brothers*, los mercados financieros entraron en pánico y fue la primera vez en la que el público oyó términos hasta entonces utilizados solo en el argot técnico, por ejemplo, el paradigmático *credit crunch*: un término que hacía referencia inicialmente a que los bancos privados se negaban a realizar préstamos entre ellos y que acabó extendiéndose como metonimia de esa etapa de la crisis. Desde entonces, palabras como “crédito” o “deuda” pasaron a formar parte de todos los foros y tertulias.

Los gobiernos prometieron salvar al sistema mediante refuerzos tanto al sector bancario como a la economía pública y privada. Ante la gravedad de la situación, aún no del todo palpable para la población, se produjo un escenario de acción discursiva, más performativo que filosófico, y los actores principales a nivel global, como Nicolas Sarkozy, se pronunciaron sobre un sistema capitalista desbocado que había devenido en un ente caduco y que debía ser refundado en las bases del trabajo y la moralidad (Sarkozy, 2010)<sup>1</sup>. En España, el gabinete de Rodríguez Zapatero decidió utilizar eufemismos como “desaceleración acelerada”, “dificultad que nos viene de fuera”, o “ajuste duro” (Garrea, 2008), de manera que se evitase el uso de cualquier referencia directa a la explosión de la crisis. Estos ejemplos en distintas lenguas europeas simbolizan el importante revulsivo lingüístico que la Gran Recesión supuso en el discurso político, el cual queda plasmado

---

<sup>1</sup> “...nous sauverons le capitalisme et l'économie de marché, en le refundant, oserai-je le mot, en le moralisant.” (“... salvaremos el capitalismo y la economía de mercado, lo refundaremos, me atrevo a decir la palabra, moralizándolo”.)

inmediatamente en la prensa y, consecuentemente, provoca reacciones perlocutivas en los distintos actores sociales.

Como respuesta al cambio de paradigma social que supuso la crisis, especialmente alrededor de 2011, florecieron nuevos movimientos ciudadanos a escala global, los cuales ganaron una relevancia social extraordinaria e inusitada hasta entonces; hasta el punto en que los manifestantes fueron reconocidos como figura del año en la revista *Time*. El giro discursivo que provocaron actores sociales ciudadanos como el *Movimiento 15-M* en España, el desborde del sindicalismo tradicional en Grecia, o la emergencia de los campamentos protesta de *Occupy Wall Street* inicialmente en los Estados Unidos (y su posterior expansión al resto del mundo) abrieron una brecha para nuevos debates impulsados por la ciudadanía que ponían en cuestión el discurso oficialista del experto en el contexto de las políticas de austeridad económica.

La intervención ciudadana en el discurso político supuso una aceleración exponencial en el intercambio de ideas de carácter transformador, ya que el debate caló profundamente en todos los estratos de la sociedad y, como consecuencia, tuvo repercusiones políticas profundas. Desde una traducción eminentemente divulgativa<sup>2</sup> de las decisiones políticas y económicas, se puso rostro humano a las consecuencias de las políticas de austeridad: como el drama de los desahucios o las consecuencias de los recortes en servicios públicos. En el contexto institucional, y de manera muy esquemática, la repercusión de estos movimientos se ha traducido en el fin del sistema bipartidista en España tras la emergencia de nuevos actores políticos a ambos lados del eje político izquierda-derecha. En el resto del mundo, se han propulsado representaciones políticas otrora minoritarias, como el ascenso de Alexis Tsipras al gobierno de Grecia o la puesta en escena de nuevas figuras en el mundo anglosajón, como el liderazgo de Jeremy Corbyn en el Partido Laborista británico o la meteórica candidatura de Bernie Sanders a las primarias del Partido Demócrata, ambas con un respaldo eminentemente popular fuera del impulso de los lobbies y otros grupos de interés tradicionales.

La popularización de una serie de elementos discursivos especializados viene dada, entre otros factores, por la necesidad de poner negro sobre blanco y describir el desarrollo de los eventos de la crisis. Paralelamente, emergían la prensa online y las redes

---

<sup>2</sup> Por ejemplo, *El capital del siglo XXI* de Thomas Piketty (2013).

sociales, en detrimento de los medios tradicionales, a cuyo desarrollo discursivo acompañó un giro brusco en los factores económicos, sociales y políticos. Durante esta década, el acceso a Internet se ha universalizado: los sitios web dinámicos y las redes sociales son hoy el núcleo de un mundo estrechamente hiperconectado que ha llegado más lejos que nunca y tiene efectos de mayor calado en la vida de las personas, especialmente si lo comparamos con la proyección tradicional de los medios de comunicación masivos. En la actualidad, Internet es un espacio multidireccional de socialización e intercambio de opinión, un ágora mundial de intercambio de ideas. Al calor de la difusión de noticias en los muros de las redes sociales surgen debates entre personas de todo el mundo. Como consecuencia, la información, y con ella los nuevos usos del lenguaje, se difunden ahora no solo desde la prensa y los medios tradicionales (que en gran medida siguen marcando las pautas de la opinión pública, puesto que se concentran en manos de muy pocos conglomerados privados<sup>3</sup>), sino también desde la voz del público en las redes sociales. En una situación de crisis económica, la difusión de noticias dedicadas a explicar la crisis al gran público pone de manifiesto la necesidad de divulgar el funcionamiento básico de un dominio tan especializado como la economía, con el fin de aportar una interpretación, nunca aséptica, de la situación político-económica, y servir de atalaya de observación de la opinión en los debates a pie de calle. La prensa generalista online es, dado su papel mediador, uno de los mejores reflejos del discurso de una época, y es un instrumento próximo para obtener información veraz de la lengua en uso (Gouws, Metzler, Cai, & Hovy, 2011; Toret, 2015).

En este contexto de crisis es muy pertinente, además, poner sobre la mesa una serie de preguntas centrales en este trabajo: ¿qué palabras marcan el tono del discurso de un texto? ¿Existen palabras específicas de una especialidad con una orientación semántica positiva o negativa, o son estas un ente totalmente aséptico e independiente del contexto? Dado el caso: ¿es este sentimiento permanente o evoluciona con los tiempos? Si esto es así, ¿podemos determinar de alguna manera los factores externos que hacen que este sentimiento varíe? Se parte de la hipótesis de que el lenguaje es un dispositivo altamente dinámico que es reflejo de la evolución histórica y permanente de las connotaciones subjetivas, actitudes, evaluaciones y postura de las palabras que se usan a diario. Las

---

<sup>3</sup> En relación con esto existe una amplia literatura, por ejemplo, en España, *Traficantes de información*, de Serrano (2010).

palabras y expresiones, pertenecientes tanto a los lenguajes generales como a los especializados, pueden transmitir sentimientos, opiniones y creencias, que son el núcleo de la subjetividad en el lenguaje humano (Liu, 2012). Moesner (2003) considera que las palabras cambian a medida que se asocian con conceptos más o menos favorables que antes, en relación con la máxima de Firth (1957) “You shall know a word by the company it keeps” (p. 179), el cual considera que parte del significado de una palabra se deriva de las palabras con las que coocurre:

Meaning by collocation is an abstraction at the syntagmatic level and is not directly concerned with the conceptual or idea approach to the meaning of words. One of the meanings of night is its collocability with *dark*, and of *dark*, of course, collocation with *night*. (p. 196).

En términos psicológicos, Aitchison (1997) considera que los humanos determinan el significado de las palabras desconocidas a través de las palabras con las que coocurren, de manera que el significado es altamente dependiente del contexto producido por los eventos. Un *evento* sería un término general para referirse a situaciones que ocurren en un momento determinado. Los eventos pueden ser puntuales o durar un período de tiempo. También se consideran eventos aquellos predicados que describen estados o circunstancias en las que algo se obtiene o se mantiene (Pustejovsky *et al.*, 2004). En esta línea, estamos de acuerdo con la teorización de las *palabras evento*<sup>4</sup> como las entiende Moirand (2007, 2016): determinadas unidades se utilizan inicialmente vinculadas a las lenguas de especialidad, es decir, parten de un punto de vista superficialmente aséptico y científico. Posteriormente, estas unidades léxicas adquieren significados más politizados tras un evento o situación en concreto. Estas se relacionarán con otros elementos con un sentimiento más o menos positivo y, a través de estas asignaciones, se convertirán en medios de transmisión definatorios de los hechos comunicativos inicialmente especializados (p. ej., la *deuda soberana* o una *recapitalización*) para el consumo y reutilización del gran público. Finalmente, estos terminan adquiriendo nuevas dimensiones léxicas de la mano de los eventos, las distintas opiniones e interpretaciones producidas tanto por las autoridades como por el público y, finalmente, las repercusiones que tienen estas en la memoria de una comunidad de hablantes.

---

<sup>4</sup> Traducción directa de *mots-événement* en el original en francés.

Este estudio rastrea la evolución de la orientación semántica de una serie de términos en lengua inglesa durante un evento de profundo calado, como fue la Crisis Global durante un intervalo de tiempo relativamente corto (ocho años). Especialmente, se centrará en el estudio de dos tipos de evolución en la orientación semántica, como lo discutieron Bloomfield (1933) o, más recientemente, Traugott y Dasher (2002): *mejora* y *peyoración*. Según Moesner (2003), el contenido semántico de las palabras se modifica a medida que se asocian con conceptos más o menos favorables de lo que solían hacerlo. En este proceso, la orientación semántica de una determinada palabra puede evolucionar al intensificar su positividad o negatividad, o la palabra podría incluso invertir su polaridad por completo a lo largo de un intervalo de tiempo determinado.

A pesar de que la variación en la orientación semántica es indudablemente difícil de rastrear exclusivamente a través de la introspección, tal observación se ve ahora facilitada por la disponibilidad actual de grandes cantidades de texto y el desarrollo de soluciones de procesamiento del lenguaje natural. Se aprovecharán las oportunidades que ofrecen las cabeceras editoriales digitales ya que ofrecen una gran cantidad de textos que son fácilmente accesibles en Internet y que permiten su rápido procesamiento automático. En este caso se utilizaron los textos de la sección *Business* del intervalo 2007-2015 de The Guardian y The New York Times, cabeceras referentes a nivel mundial. La observación de cambios en la orientación semántica es un terreno poco estudiado en profundidad y parece ser una dirección prometedora para profundizar en el conocimiento de variaciones en este concepto a muy corto plazo. De esta manera, se puede observar con inmediatez si se producen nuevos sentidos en el uso del léxico. Asimismo, se ofrece la posibilidad de establecer relaciones con las situaciones históricas que motivan estos cambios (Crowley y Bower, 2010; Cook y Stevenson, 2010; Cavallin, 2012).

## 1.2 Justificación y motivación

La velocidad a la que se lanzan las noticias ha llegado a su punto máximo durante esta década, donde las tecnologías de la información y la sociedad están influyendo en la comprensión del mundo que nos rodea a golpe de clic. La combinación del desarrollo tecnológico, por un lado, y una crisis social, cultural y económica por el otro, ha producido nuevos desafíos para el estudio de la influencia de la sociedad en el lenguaje y viceversa. El estudio del lenguaje evaluativo y la evolución de la orientación semántica es relevante en sí mismo desde una perspectiva puramente lingüística. Esto también lo es

desde la perspectiva más técnica del procesamiento del lenguaje natural ya que el estudio de este aspecto del lenguaje se ha vuelto muy relevante en los últimos años.

Las organizaciones siempre han buscado conocer el sentimiento de su público objetivo (clientes, consumidores, usuarios) hacia los productos y servicios que ofrecen. Tradicionalmente, esto se realiza a partir de expertos en consumo y medios de comunicación de manera manual a través de encuestas de satisfacción y otros documentos similares. Con la llegada de las redes sociales y la capacidad espontánea que tienen sus usuarios para generar contenido en la Web, la cantidad de documentos relevantes es tan grande que simplemente no es posible analizarla sin la ayuda de herramientas computacionales.

La *Web 2.0* actual se centra en el contenido del usuario. Para 2013, el 70% de los usuarios de Internet ya estaban registrados en una o más plataformas sociales. Tras la emergencia de las redes sociales, en detrimento de los medios tradicionales, Internet ha ofrecido a la ciudadanía el poder de convertirse en creadores de información de cualquier tipo, pasando a ser un elemento activo para las instituciones o empresas, que se organiza en la red de manera espontánea y distribuida. El contenido generado por el usuario (*user generated content*, UGC) se define como la información (por ejemplo, reseñas de productos, comentarios, tutoriales, etc.) que los usuarios ponen a disposición de sus pares en Internet. Este contenido se comparte fuera de las prácticas profesionales sin ninguna expectativa de beneficio, sino simplemente como medio de expresión personal y expectativas de retroalimentación. Por lo tanto, los usuarios pueden obtener una idea de los datos reales, no analizados por los medios de comunicación regulares. Con respecto a esto, los sitios web de las cabeceras periodísticas tradicionales como The Guardian o The New York Times se han adaptado a los tiempos de manera progresiva. Más allá de cambios superficiales, como la aceptación de esta multidireccionalidad a través de una participación de los usuarios (p. ej., en forma de texto, fotografías o vídeos) o la aparición de nuevos formatos en las cabeceras (*blogs* o *microblogs*), en gran medida, el discurso está modulado de manera directa o indirecta por la mediación de sus usuarios a través de votos y comentarios en las redes sociales. De hecho, el reto de la adaptación de la cabecera del New York Times al periodismo Web 2.0 pasó por la consolidación discursiva que ofreciera una combinación de “periodismo analítico” y “lenguaje simple” (Nafra, 2017 p. 236). La cantidad de contribuciones a estas plataformas de UGC ha crecido hasta tal punto que los investigadores de muchos campos las consideran fuentes de información inestimables para obtener una visión más profunda y una mejor comprensión de las

tendencias y evolución de intereses de las sociedades (Balasubramaniam, 2009, Krumm, Davies, y Narayanaswami, 2008).

Desde la década de 2000, algunas de las principales redes sociales, como *Reddit*, *Twitter* o los comentarios de los usuarios en periódicos basados en la web, han incluido arquitecturas abiertas que permiten la extracción de contenido por los desarrolladores de aplicaciones, las llamadas interfaces de programación de aplicaciones (API). Gracias a las API, los investigadores pueden extraer grandes cantidades de contenido generado por el usuario que incluye, entre otros, información lingüística de casi cualquier tipo de dominio. Dichas herramientas han sido objeto de una intensa investigación y desarrollo durante las últimas dos décadas, dando lugar a lo que se conoce como minería de textos (*Text Mining*) o análisis textual (*Text Analytics*). El análisis de sentimiento, o minería de opinión (*Opinion Mining*), es el subcampo principal de esta disciplina y se ha convertido en una temática cada vez más presente en la literatura especializada, ya sea como un desafío científico y técnico, principalmente la lingüística y la informática, como en sus aplicaciones prácticas experimentales en ciencias políticas, mercadotecnia o periodismo.

El análisis de sentimiento se ocupa del tratamiento computacional de la opinión y la subjetividad en los textos. A través de distintos enfoques y algoritmos es posible calcular con gran precisión la *polaridad* u *orientación semántica* de elementos léxicos, oraciones y textos. Básicamente, una herramienta de análisis de sentimiento clasifica un texto como positivo o negativo, ya sea como una simple clasificación binaria (enfoque *thumbs up / thumbs down*) o asignarle una puntuación dentro de una escala (enfoque *seeing stars*) (Turney, 2002; Pang, et al., 2002, 2005).

Este estudio pretende analizar la variación en la orientación semántica de determinados términos financieros aparecidos en la prensa durante la crisis global a través del análisis de sentimiento. Para ello se utilizará *Lingmotif*, un sistema de análisis de sentimiento desarrollado por el grupo Tecnolengua en la Universidad de Málaga (Moreno-Ortiz et al., 2017, 2018). Esta herramienta se basa en el léxico y utiliza un método de cálculo que pondera la frecuencia de segmentos positivos, negativos y neutros para determinar la orientación semántica. Así, se pretende completar y profundizar en la problemática planteada en las conclusiones de mi trabajo de fin de máster y contribuir a los hallazgos de investigaciones anteriores.

Los esfuerzos previos se centraron en la extracción de un conjunto inicial de términos especializados que sirviese para evaluar las posibilidades de diseño de un módulo léxico de análisis de sentimiento en el dominio económico-financiero (Moreno



Ortiz y Fernández-Cruz, 2015). A partir de estos trabajos, se evidenció que la orientación semántica es en gran medida dependiente del dominio. Para abordar esta problemática, en este proyecto se propuso una metodología sencilla y eficiente que permitió integrar el análisis de sentimiento de dominio específico en un clasificador diseñado inicialmente para textos de lenguaje general. En este proyecto, se ampliará la metodología de forma que, a partir de fuentes textuales especializadas, se extraiga un amplio lexicón de sentimiento del dominio económico-financiero con un alto índice de rendimiento.

Finalmente, el interés principal de este trabajo se justifica también en la escasez de estudios de la evolución de la orientación semántica basado en un enfoque de análisis de sentimiento. Los resultados futuros pueden abrir la posibilidad de desarrollar metodologías aplicadas con el fin de detectar patrones de cambio en la orientación semántica. Las herramientas de análisis de sentimiento ofrecen una amplia gama de posibilidades que permiten investigar el impacto de las nuevas condiciones socioeconómicas pueden haber producido en el uso del lenguaje evaluativo y, por lo tanto, obtener un mejor conocimiento lingüístico de la evolución actual del lenguaje.

### 1.3 Preguntas formales de investigación

El presente trabajo toma como punto de partida las siguientes preguntas:

1. ¿Varían los términos a medida que se asocian con conceptos más o menos favorables que antes?
2. ¿Impulsan los eventos históricos de mayor orden, como es el caso de las crisis, cambios rápidos en la orientación semántica de los términos?

### 1.4 Definición de objetivos

#### 1.4.1 Objetivo general

**Objetivo general.** Analizar las fluctuaciones en la orientación semántica de determinados términos económicos dentro del período 2007-2015 a través de la caracterización del impacto de los eventos de mayor orden en las variaciones semánticas de las unidades léxicas.



### 1.4.2 Objetivos específicos

Dados el objetivo general antes mencionado, los siguientes objetivos específicos están en orden:

**OE1.** Determinar un lexicón de sentimiento de dominio específico en lengua inglesa para el dominio de las noticias económicas de alta precisión y exhaustividad a partir de un corpus de texto especializado.

**OE2.** Definir un subconjunto de datos longitudinal en forma de oraciones que contienen los términos de estudio y que serán el *input* del análisis de sentimiento.

**OE3.** Para los términos económicos-financieros objetos de estudio, identificar los factores (sociales, políticos, económicos, etc.) que han acompañado a cambios en el uso, desde la perspectiva de sentimiento, durante el período 2007-2015.

**OE4.** Analizar cualitativamente las posibles variaciones en la prosodia semántica de los términos de estudio a través de los años.

## 1.5 Organización de la tesis

La tesis consta de nueve capítulos:

El Capítulo 1 es la presente Introducción, que incluye el planteamiento inicial del problema, la justificación, los objetivos y las preguntas formales de la investigación.

El Capítulo 2 realiza una revisión del contexto del estudio: una revisión histórica de la crisis y una reflexión sobre la influencia de la prensa en el discurso económico de acuerdo con los factores de esta década.

Los Capítulos 3, 4 y 5 componen el marco teórico. El Capítulo 3 realiza una revisión bibliográfica sobre el estudio del lenguaje evaluativo y sus principales teorías, el Capítulo 4 trata sobre los lenguajes de especialidad, las principales teorías sobre la terminología y los principales desarrollos terminográficos. Por último, el Capítulo 5 revisa el estado del arte y los desarrollos técnicos y teóricos del análisis de sentimiento.

El Capítulo 6 recoge la metodología. Primero revisa los paradigmas y enfoques principales de investigación y posteriormente desgrana el diseño de la investigación. Se justifica aquí el diseño del *Esmeraldas Corpus of Business English*, fuente principal de datos para este trabajo. Además, se exponen las dos herramientas utilizadas en este trabajo: la suite de análisis de corpus *The Sketch Engine* y el clasificador de sentimiento *Lingmotif*. Finalmente justifica el diseño y posterior desarrollo de *LexiEcon*, un *plug-in* de dominio económico-financiero para *Lingmotif*.

En el Capítulo 7 se analizan cualitativa y cuantitativamente todas las oraciones que incluyen los diez términos de estudio de este trabajo, recopilados en sendos conjuntos de datos extraídos del corpus.

El Capítulo 8 interpreta y discute de manera conjunta los resultados obtenidos en el capítulo previo, en torno al marco teórico.

Finalmente, el Capítulo 9 expone las conclusiones definitivas junto con las limitaciones de este estudio y las perspectivas futuras de investigación.



## 2. Contexto de la investigación: el mundo y el lenguaje ante la crisis global

« From that standpoint we are not changing our language. We have the same language – we have only one language – which we have repeated very strongly, not only since the start of the crisis, but since the setting-up of the euro. And at times we have had to fight for maintaining, for preserving this soundness of our governance, particularly as regards fiscal consolidation, but also in other domains. »

Jean-Claude Trichet (1942), presidente del Banco Central Europeo (2003-2011)  
Rueda de prensa mensual del BCE, 3 de febrero de 2011

« The language is ineffective. The language is not going to change things. Policy has to be adapted. What we need is real action. »

Dominique Strauss-Kahn (1949), director del Fondo Monetario Internacional (2007-2011)  
The Guardian. 10 de octubre de 2010

### 2.1 Contextualización histórica, política y social de la crisis (2007-2018)

#### 2.1.1 La crisis de las hipotecas *subprime* y la crisis del crédito

El 15 de septiembre de 2008 se convierte en la fecha icónica del estallido de la Gran Recesión tras la quiebra del banco de inversiones Lehman Brothers. Sin embargo, la crisis global se gestaba desde principios de siglo e iba tomando forma a través de la especulación inmobiliaria y el aumento del endeudamiento de las familias debido a los bajos salarios. Las entidades reguladoras no supieron prever a tiempo esta crisis, hecho que se combinó con errores contables y éticos. Tras el 11-S y la crisis de las puntocom, la Reserva Federal estadounidense bajó los tipos de interés y fomentó la inversión inmobiliaria. A partir del mito infundado de que “el precio de la vivienda nunca baja”,

estas inversiones aparentemente de carácter conservador estaban siendo financiadas a través productos crediticios cada vez más exóticos.

En febrero de 2007 estalló la burbuja debido a la caída en los precios de la vivienda y a la subida de los tipos de interés por parte de la Reserva Federal de los Estados Unidos. Comienza entonces la crisis de las hipotecas *subprime*, y con ella quiebran 50 empresas estadounidenses especializadas en estos productos de alto riesgo que estaban clasificados como inversión segura por las agencias de calificación crediticia.

Tras una época con dificultades económicas, millones de personas se quedan sin condiciones de cumplir sus pagos mensuales y se produjo una reacción en cadena. Las *subprime* se ofrecían a familias con finanzas frágiles y sin garantía alguna de la aprobación de la deuda. Mientras esto ocurría, muchos analistas y prensa especializada minimizaron el impacto de esta crisis y ofrecían perspectivas optimistas considerando que este evento tendría un impacto mínimo en la economía.

No obstante, en el mes de junio el problema de la deuda hipotecaria *subprime* empezó a contaminar los mercados financieros internacionales: los *hedge funds* ‘fondos de cobertura’ del banco de inversiones Bear Stearns revelaron que sus fondos *subprime* tenían dificultades importantes. Los sistemas financieros nunca habían contemplado la posibilidad de que el sistema se viniera abajo, por tanto, lo que a principios de año tenía un valor de varios millones de dólares, poco después tenía un precio indeterminado (Kensil y Margraf, 2015; Observatorio Metropolitano, 2011).

El problema aquí fue la gran distribución que tenían estos productos: entre los compradores de estos activos tóxicos se encontraban bancos de todo el planeta. Se produjo una escasez de liquidez al congelarse la realización de préstamos interbancarios, de manera que los bancos centrales deben intervenir e inyectar miles de millones de liquidez. Esta perturbación se incrementó durante el verano y como consecuencia se desataba la fase activa de la crisis que se extendió a Europa. El 9 de agosto de 2007 el BNP Paribas francés rescindió sus propias inversiones en préstamos de alto riesgo, que se manifestó como una crisis de liquidez. A comienzos de 2008, el gobierno británico debe nacionalizar al Northern Rock, el quinto mayor banco del país. En marzo, JP Morgan-Chase adquirió el Bear Stearns por un precio simbólico mientras trataba de evitar el incipiente colapso (Mueller, 2018; Wiggins, Piontek y Metrick, 2014).

El banco de inversión Lehman Brothers declaró la quiebra el 15 de septiembre a pesar de los intentos de rescate del gobierno de Estados Unidos. Mientras los principales bancos del mundo capeaban la crisis, Lehman Brothers había cerrado 2007 publicando

un récord de resultados anuales con un beneficio neto de 4.200 millones de dólares. El veterano banco de inversión de Wall Street no solo no hizo ninguna provisión para subsanar esta crisis sino que, no mucho antes, su carismático portavoz Dick Fuld había alardeado de superponerse a los ciclos del mercado gracias a la diversificación de capitales. Una semana después de esas declaraciones, Lehman Brothers anunciaría por primera vez 2.800 millones de dólares de pérdidas trimestrales, aunque ya el 2 de junio, la agencia de calificación de riesgo Standard and Poor's le había bajado la clasificación de A+ a A. El 10 de septiembre se publicaron resultados catastróficos. La institución busca cierta manera de aumentar su liquidez, principalmente en la búsqueda de compradores, potencialmente Bank of America y Barclays, sin éxito al considerarlo un banco *tóxico*. El sábado 13 de septiembre el presidente del Banco de la Reserva Federal de Nueva York, Timothy F. Geithner y Hank Paulson, secretario del Tesoro estadounidense, plantearon en una reunión de emergencia los escenarios de salida, planteándose una liquidación de emergencia de sus activos. Ninguna entidad bancaria llegó a materializar una adquisición de emergencia y transcurrida la medianoche del lunes 15 de septiembre formalizaron la declaración en bancarota. Su colapso sacudió los mercados a ambos lados del Atlántico: el índice Dow Jones cayó 500 puntos en un solo día, una marca no superada desde el 11-S, paralizando el flujo mundial de crédito. La onda expansiva llegó inmediatamente a Europa: Alistair Darling, el ministro británico de Hacienda, se dirigió a su homólogo americano para comunicarle que iba a hacer todo lo posible para evitar importar el *cáncer* de Estados Unidos.

A pesar de lo icónico del evento, la quiebra de un Lehman Brothers “too big to fail” era una consecuencia del espíritu de los tiempos, pues el sentimiento alrededor de los mercados estaba muy deteriorado. Como se ha visto aquí, el mundo financiero contaba con un rescate como salida de emergencia. Sin embargo, la *caída del gigante* se había convertido en un revulsivo para los inversionistas: a partir de entonces todo era posible, se abría paso a la incertidumbre y esto provocó una aceleración en las ventas. El Dow Jones había caído un 44% seis meses después, lo que se consideró el desplome más atronador de Wall Street desde la Segunda Guerra Mundial (Mueller, 2018; Wiggins, Piontek y Metrick, 2014). Sobre el proceso post colapso, BBC Mundo (2009) relataba la naturaleza no cuantificable de la liquidación a través del relato de los registradores que gestionaron las montañas de documentos repletos de jerga técnica, cuya gesta compararon con “la autopsia de un animal gigantesco y desconocido”.

Más allá de los datos macroeconómicos, las páginas económicas se cargaron no solo de cifras e interpretación sino también de lenguaje relacionado con el ámbito de las emociones. Las historias tras las finanzas recogidas en el texto relatan algunos detalles sobre las emociones y la batalla por la interpretación de textos especializados especialmente oscuros:

Han tenido que lidiar con lágrimas, ira, un vocabulario a veces incomprensible y cifras que no cuadran. Todo para llegar a un descubrimiento que no deja de preocupar: no se sabe a ciencia cierta quién le debe a Lehman o a quién le debe Lehman (BBC Mundo, 2009).

Después de Lehman se produjo una onda expansiva de quiebras, rescates y adquisiciones bancarias. El rescate más paradigmático fue el de AIG por 180 mil millones de dólares y la toma del control por parte de este, ya que su quiebra hubiera supuesto la caída del llamado Bulge Bracket, los bancos más poderosos del mundo: Goldman Sachs, Morgan Stanley, Bank of America y Merrill Lynch junto con decenas de bancos europeos. Las economías se desaceleraron durante este período en todo el mundo, la falta de confianza aumentó, mientras que la facilidad de crédito disminuía, cayó el comercio internacional y los valores sufrieron grandes pérdidas durante 2008 y principios de 2009. La falta de liquidez provocó el colapso de empresas clave, una significativa pérdida de empleos y como consecuencia el declive del poder adquisitivo. Más allá de la descripción de las cifras, esto supuso consecuencias dramáticas en millones de hogares a ambos lados del Atlántico que repentinamente vieron como se deprimían sus estándares de vida (Ocampo, 2008; Stiglitz, 2009).

Los gobiernos y los bancos centrales respondieron con paquetes de estímulo fiscal sin precedentes, generalizándose los rescates institucionales. En 2009, los Estados Unidos aprobaron una Ley de Recuperación y Reinversión para compensar este descenso en el consumo y la capacidad de préstamo. La Reserva Federal de los Estados Unidos desembolsó inicialmente casi 14 mil millones de dólares, convirtiéndose de facto en el único prestamista con solvencia para un importante sector de la economía. Esta ley permitió al gobierno de los Estados Unidos inyectar un grueso total de 350.000 millones de dólares para la compra de activos. Esta decisión frenó en gran medida la velocidad del colapso y como resultado, los Estados Unidos salieron de los niveles de contracción económica a mediados de 2009. España, sin embargo, solo logró salir de la recesión en 2010, volviendo a hundirse en 2011 como consecuencia de la inacción ante la crisis. El presidente del gobierno de España, José Luis Rodríguez Zapatero, decidió utilizar complejas estrategias retóricas para negar la existencia de una crisis, considerándola un

fenómeno estadounidense. Inicialmente, los bancos debieron superar los problemas intentando minimizar los costos de la crisis a corto plazo. La inacción generó incertidumbre, los activos inmobiliarios mayormente especulativos comenzaron a desplomarse y el flujo de crédito se estancó, inflando la generación de la deuda en una economía escasamente endeudada (Jorrín, 2018).

Esta paralización de la actividad económica condujo a la Gran Recesión de 2008-2012 y contribuyó a la formación de la siguiente crisis: la crisis de la deuda soberana europea. La debacle de Lehman no fue en ningún caso un fenómeno americano, sino que englobó por completo a las economías del Atlántico Norte. Los bancos europeos estuvieron contagiados desde un principio, pues en 2006 un 30% de las hipotecas de alto riesgo estaban en manos de bancos europeos. La antigua firma belga Dexia, con fuertes inversiones en los Estados Unidos, fue la primera corporación europea en ser rescatada conjuntamente por los gobiernos de Bélgica, los Países Bajos y Francia tan solo quince días después de la caída de Lehman. Las consecuencias de la explosión de la burbuja inmobiliaria fueron especialmente dolorosas en los países del sur de Europa e Irlanda. De hecho, el antiguamente denominado Tigre Celta debió ser rescatado, pues su economía se colapsó a causa de su especial dependencia del capital estadounidense, ya que actuaba como un paraíso fiscal y puerta de entrada de entrada de la inversión estadounidense en la Zona Euro (Fratianni y Marchionne, 2013; Pérez, 2018; Teather, 2008).

### **2.1.2 La crisis de la deuda**

En 2009 estalla en Europa una nueva crisis a partir de un entorno de déficit estructural excesivamente alto y la aceleración de los niveles de deuda. La banca privada había alcanzado niveles alarmantes de deuda y su caída amenazaba con arrastrar la estabilidad financiera de la economía de muchos estados en Europa. Por esto, debieron rescatar a los bancos más afectados mediante préstamos de recapitalización, lo que provocó un fuerte deterioro de la ratio deuda / PIB. Cuatro estados de la Zona Euro (Irlanda, Portugal, Chipre y Grecia) debieron ser salvados formalmente por programas de rescate soberanos



proporcionados conjuntamente por la *troika*<sup>5</sup>, mientras que España se acogió a un programa de recapitalización bancaria equiparable en muchas medidas. Cuatro fueron también las principales crisis que se desataron en Europa: (a) crisis de crecimiento: un crecimiento real y potencial débil, (b) la debilidad competitiva, (c) crisis de la deuda soberana: altos ratios de deuda / PIB; (d) crisis bancaria en importantes *stocks* de responsabilidad civil (sector público, privado y no privado). A principios de 2010, los prestamistas exigían altas tasas de interés a los países con mayores niveles de deuda y déficit. Esto dificultó la financiación de los déficits presupuestarios y el reembolso de la deuda pública existente principalmente a varios gobiernos de la Zona Euro con bajas tasas de crecimiento económico y altos niveles de deuda extranjera: Grecia, Irlanda, Italia, España y Portugal. Si estos estados no alcanzasen a saldar sus deudas, el sistema bancario en su conjunto amenazaba con desplomarse y llevar a los estados a la bancarrota. Estas medidas sirvieron como pretexto para realizar imposiciones de austeridad presupuestaria con el objetivo inicial de aliviar los problemas causados por el estrés soberano que estancaba el crecimiento. Tras estas políticas de austeridad, se inició una marcada agenda de flexibilización y precarización de las condiciones de trabajo, una ola de privatizaciones en el sector público y recortes presupuestarios especialmente en los derechos sociales, salud y educación (Shambaugh, Reis y Rey, 2012; Noeth et al., 2012).

Ya en octubre de 2008, el gobierno de Alemania se había cerrado en firme a una solución europea conjunta. Durante la presidencia de Jean-Claude Trichet, el Banco Central Europeo seguía unas líneas de actuación muy similares a las del *Bundesbank*, aunque finalmente se puso en marcha una respuesta conjunta del Eurogrupo y el BCE que pudo responder al problema de falta de liquidez mediante una política expansiva. Los estados europeos respondieron en una triple dirección: (1) aprobación de programas de expansión del gasto público y del consumo para tratar de mantener la actividad, (2) inyectar una enorme cantidad de liquidez a las entidades financieras para reactivar el crédito y (3) comprometer parte de los gastos en paliar los efectos inmediatos de la crisis a través de, principalmente, subsidios de desempleo. Mientras se pretendía evitar que la crisis bancaria llevara a una quiebra financiera, el gasto público disparó el déficit y rápidamente la crisis crediticia se convirtió en una crisis de deuda soberana. Según el

---

<sup>5</sup> Las tres organizaciones financieras internacionales que representan a los acreedores de rescate: el Fondo Monetario Internacional, la Comisión Europea, y el Banco Central Europeo.

Observatorio Metropolitano de Madrid (2011), la burbuja de la deuda europea fue combatida principalmente mediante la dupla rescate bancario y austeridad en los sistemas sociales de protección. Esta estaba fundamentada en torno a cuatro aspectos: (1) control del déficit público y castigo a la expansión al gasto público, (2) una alta tasa de desindustrialización en los países periféricos, (3) anteposición del rescate a las entidades financieras, (4) una amenaza constante de que el impago de la deuda de uno o varios países del sur de Europa arrastrase consigo a la economía de la Zona Euro y a la moneda única.

En 2007, España tenía un superávit presupuestario de alrededor del 2% del PIB y la deuda soberana era solo del 40% del PIB. Cinco años después, en 2012, la deuda pública española se había duplicado para alcanzar aproximadamente el 90% del PIB. Esto era similar en otros países de Europa, como Irlanda. En 2009, la Zona Euro no pudo responder conjuntamente ante el creciente problema de la deuda soberana de Grecia. Con la meta de evitar que un estado quiebre, como el gigante Lehman, a partir de 2010 aplicaron políticas de austeridad más próximas a un castigo que a una solución estructural, pues empobrecían el motor económico, lo que generaba incertidumbre e inflaba aún más la deuda. La aplicación de estas políticas fue especialmente dolorosa en España, la región más azotada por la crisis inmobiliaria (Jordà et al., 2016).

Dado que el BCE no puede emitir bonos de deuda, ni inicialmente comprar los emitidos por los estados, Trichet mezcló una política monetaria obsesionada en combatir una inflación a efectos prácticos insignificante con dos subidas de los tipos de interés en 2011, lo que provocó una segunda recesión. Se dispararon los tipos de interés en los países del sur, y la prima de riesgo española llegó a subir hasta los 640 puntos, hundiendo el valor de los activos. El ministro español Luis de Guindos firmó dos decretos que exigían provisiones milmillonarias a los bancos. Estos cargaron las provisiones contra resultados, generando pérdidas millonarias. Estas decisiones fueron políticas que no solo no generaron confianza, sino que provocó una estampida de inversores que hizo imposible la supervivencia del sector. El estado español solicitó ser rescatado en junio de 2012. La gestión había sido tan lenta como funesta y, lo que es peor, calificada como inevitable, pues se sigue alegando que no había forma alguna de recuperar el dinero inyectado. El propio Banco Central Europeo puso de manifiesto en 2015 que las fórmulas empleadas entonces por España hacían que se produjese un enorme desembolso público. Estas estaban diseñadas de una forma que hacía muy difícil la recuperación de la inversión de capital público (Jorrín, 2018).

Para los rescates, España utilizó destacablemente el esquema de protección de activos (EPA). Este consiste en respaldar con dinero público operaciones de venta de un banco nacionalizado a la banca privada mientras se protege al inversionista privado de pérdidas. En otros casos, las entidades financieras nacionalizadas fueron malvendidas por un euro tras ser rescatadas con dinero procedente de impuestos. Además de acciones en Bankia, el único tipo de activos que mantiene el Estado español es la Sareb, el *banco malo* creado ad-hoc para absorber los activos inmobiliarios tóxicos, el cual siguió generando pérdidas en 2018. En otras palabras, el estado favoreció a entidades privadas de manera totalmente altruista, regalando cantidades ingentes de dinero recaudado con impuestos que de otra manera hubiesen sido invertidos en el bienestar público. En contraste con esto y simplificando, los Estados Unidos recapitalizaron los bancos mediante la compra de acciones. Así, el estado solo debía vigilar que el negocio de los bancos mejorase y vender esos títulos para recuperar el dinero del rescate (Jorrín, 2018).

Un caso de especial relevancia es el de Grecia. Hasta mediados de la década de 2000, la economía helénica fue una de las que más creció en Europa. Durante la crisis financiera, las principales industrias griegas, el transporte y el turismo, tornaron muy sensibles. A principios de 2010 se descubre que el déficit helénico era mucho mayor que el presentado a Europa, pues ocultaban deudas importantes fuera de los balances y ocultando partidas de gastos en los presupuestos. La reacción no se hizo esperar: el mismo oligopolio de calificación que había calificado a Lehman Brothers como excelente unas semanas antes de su desplome (Standard & Poor's, Moody's o Fitch), calificaba ahora a la deuda griega como bono basura. Esto hizo que la prima de riesgo de los bonos de deuda se disparase hasta 1000 puntos por encima del bono alemán, el cual se utiliza como referente. Aunque Grecia corría un serio riesgo de quiebra, la Unión Europea tardó en reaccionar y, tras interminables jornadas de negociación, la *troika* decidió inyectar el 40% del PIB griego con el fin de insuflar vida a la economía griega. A cambio, se antepuso el pago de intereses a cualquier tipo de gasto social, lo que se traducía en una reducción drástica del sector público: disminución de sueldos, recortes de las ayudas públicas y privatizaciones masivas. Este fue el primero de los tres paquetes de rescate. Los países centrales de la UE tomaron el ejemplo griego y se autoimpusieron austeridad ante la amenaza de presión de los mercados (Observatorio Metropolitano, 2011).

En 2011, el primer ministro Papandreu consultó a la población sobre las medidas a través de un referéndum sobre un nuevo plan de rescate. Fue el primer intento de intervención democrática ante la crisis dentro de la UE, aunque tuvo que dar marcha atrás

dada la presión del bloque central, que amenazaba con paralizar las ayudas. Esta serie de eventos provocó su dimisión y acordó designar al tecnócrata Lucas Papademos para implementar las medidas de austeridad previas al segundo rescate. En un ambiente de amenazas de expulsión de la zona euro, el temido *Grexist*, en marzo de 2012, Grecia dejó de pagar la deuda. Esto se convirtió en el mayor incumplimiento por parte de un gobierno en la historia. La reestructuración se materializó a partir de una nueva ley aprobada por el gobierno para que los titulares privados de deuda aceptaran el cambio de condiciones de 206 mil millones de euros en este tipo de bonos (Bakers y Sassard, 2012).

Tras varios años de leve mejoría en los datos macroeconómicos, Grecia entra de nuevo en nueva recesión en 2014. En las elecciones de ese año, se produce la victoria de Syriza, la coalición de izquierda radical, tras un ascenso meteórico apoyado por los sectores populares deprimidos y en constante movilización desde el inicio de la crisis. La plataforma liderada por Alexis Tsipras se negaba a respetar el rescate, los términos de austeridad e intervención in situ. Aludiendo a la incertidumbre política, la *troika* intervino unilateralmente ante cualquier ejercicio de control democrático sobre las finanzas, retirando toda la ayuda programada hasta que aceptasen las medidas. Esto fue acompañado de una nueva crisis de liquidez y un aumento de las tasas de interés privado, lo que impedía este medio alternativo de financiación (RTÉ News, 2014).

El nuevo gobierno helénico debió negociar una prórroga del préstamo que vencía el 28 de febrero de 2015 y el flamante ministro de economía Yanis Varoufakis lideró esta negociación en el Eurogrupo y con el Fondo Monetario Internacional. Sin embargo, la *troika* se negó a que el nuevo gobierno revisara los términos previos, y a aceptar cualquier reestructuración de la deuda. Estas fueron entonces las únicas semanas donde la confrontación sobre el modelo económico fue real en el Eurogrupo, destacando el antagonismo entre el ministro de finanzas alemán Schäuble y su homólogo griego. En estas negociaciones, el gobierno griego hizo frecuentes concesiones a la *troika*, frecuentemente con el propio ministro en contra, a pesar de que finalmente no llegaron a un acuerdo final. Ante el cierre de las negociaciones, Tsipras llamó a un referéndum sobre la propuesta del Eurogrupo para el 5 de julio. La histórica campaña y posterior victoria del OXI (“NO”) fue amplia y esto provocó que los índices bursátiles de todo el mundo se desplomasen. A pesar de un respaldo popular rotundo, el gobierno de Tsipras había decidido aceptar las condiciones de Europa, lo que provocó la dimisión inmediata de Varoufakis y de gran parte de la bancada de su partido, debiendo convocar nuevas elecciones. El 13 de julio, los líderes del Eurogrupo acordaron un tercer rescate,

esencialmente el propuesto el mes anterior, con condiciones estrictas para la población como reducción de las pensiones y mayores recortes en la seguridad social, sanidad y educación y de ejercer grandes presiones fiscales sobre la clase media y trabajadora. Tras estos enormes sacrificios, entre 2016 y 2017, Grecia obtuvo un superávit de un 4%, por encima de las exigencias de los acreedores (Valero, 2018).

Volviendo al BCE, el 22 de enero de 2015, el presidente Mario Draghi anuncia el llamado Plan Draghi, un agresivo programa ampliado de compra de activos, conocido como expansión cuantitativa<sup>6</sup> para combatir la creciente amenaza de deflación. El poder del discurso se puso de manifiesto el 26 julio de 2012 en Londres cuando el presidente Draghi lanzó espontáneamente la enunciación que marcó el cambio de paradigma: “Within our mandate, the ECB is ready to do whatever it takes to preserve the euro. And believe me, it will be enough”<sup>7</sup>. Peregrinamente, el discurso que dio el pistoletazo de salida a las políticas de alivio comenzó con Draghi dibujando una torpe analogía que comparaba el euro con un abejorro: “it shouldn’t fly but instead it does”<sup>8</sup>. (Krugman, 2012; Marsh, 2016)

Según Pérez (2012), a Draghi “[I]e bastó arquear una ceja, y un instante después todos corrían como conejos”. Los operadores reaccionaron inmediatamente ante la resolución directa del director del BCE, con una respuesta de calma de parte de los inversores y una bajada en los bonos. Aunque las cifras eran sintomáticas, en España la prima de riesgo cayó 50 puntos de golpe (Nelson, 2017). Finalmente, se lanzó el programa por el que se comprarían mensualmente 60.000 millones de euros de bonos de todo tipo procedentes de gobiernos, agencias e instituciones europeas con una duración inicial de 19 meses a partir de marzo de 2015. Por ejemplo, esto cubriría hasta un 29% de los bonos alemanes y más del 20% de los bonos centrales neerlandeses y españoles (Valiante, 2017). La medida fue bien recibida y supuso la inyección de liquidez en el mercado de manera que depreciaba controladamente a la moneda haciendo baratas las exportaciones europeas y, sobre el papel, colaboraba en la recuperación.

En agosto de 2018, tras acordar una prórroga en los pagos de intereses y la amortización de los préstamos, el gobierno griego terminó de cumplir con las

---

<sup>6</sup> Referenciada frecuentemente en inglés, *Quantitative Easing* o, abreviada, QE.

<sup>7</sup> “El BCE hará todo lo necesario para sostener el euro. Y, créanme, eso será suficiente”.

<sup>8</sup> “No debería volar y sin embargo lo hace”.

obligaciones de un tercer rescate, implementando por completo las medidas exigidas por los acreedores, aunque lejos de superar la austeridad. “Es el fin de un ciclo para Grecia pero también el fin de la crisis de la Zona Euro” fueron las palabras del comisario europeo Moscovici (Ríos, 2018). Tras años de dolorosas medidas de austeridad, privatizaciones y empobrecimiento, el otrora insumiso primer ministro Tsipras dio un discurso simbólico en Ítaca mientras declaraba el fin de la *odisea* moderna del pueblo griego y daba por concluida la etapa de austeridad y desertificación social. Grecia había recuperado, al menos nominalmente, su autonomía financiera (Psara y Lázaro, 2018).

Ya en agosto de 2016, un informe de auditoría interna del FMI aseguraba que esta entidad sufrió injerencias políticas procedentes del bloque central de la UE en los rescates a Grecia. Además, reconocieron errores de cálculo de magnitud extraordinaria que hundieron deliberadamente la economía durante la intervención, por ejemplo, a través del uso de proyecciones optimistas que no tenían en cuenta el impacto de los ajustes fiscales en el crecimiento y en la dinámica de la deuda. Asimismo, se reconoció que se cargó al estado griego con los gastos de la prevención del contagio que deberían haber sido compartidos, en todo caso, entre toda comunidad internacional. Así, el *sacrificio* de la economía griega que salvó al euro se había ejecutado para, principalmente, proteger a la banca francesa y alemana, acreedores y a la par responsables parciales de la deuda griega (Godin, 2016).

La crisis del euro fue deliberadamente gestionada de manera antidemocrática y sin tener en cuenta los intereses y necesidades de la ciudadanía. También en esa fecha, la directora del FMI Christine Lagarde declaró que la intervención griega había sido un éxito al haber logrado salvar el euro, a pesar de que los informes señalaron la ausencia absoluta de transparencia y de estructuras democráticas a la hora de llegar a acuerdos, como afirmaba el exministro Varoufakis años atrás. El informe señala la alta irregularidad en la toma de decisiones, ya que la mayoría de estas se tomaron extraoficialmente y sin dejar constancia de los actores que intervinieron en las decisiones. Mientras el rescate firmado por Tsipras continúa con la imposición de objetivos fiscales irrealizables, la *troika* y el Eurogrupo siguen señalando al gobierno de Atenas como culpable del fracaso de los rescates. El pago de la deuda y el coste humano de las políticas de austeridad no fue asumido por los bancos franceses y alemanes, los auténticos rescatados en esta narrativa, sino por la ciudadanía de los países de la Zona Euro con enormes secuelas sociales y políticas (Contexto y Acción, 2016; Godin, 2016).

### 2.1.3 Política y ciudadanía ante la *Gran Recesión*

La crisis global ha servido como injusto redistribuidor de la riqueza y poder de la ciudadanía, de manera que el capital, así como buena parte de los derechos, han sido reacopiados por los grandes poderes financieros. El agregado total de los motivos que se han resumido en los dos apartados anteriores indica que esta *Gran Recesión* no fue una mera bajada cíclica del capitalismo especialmente grave, sino que además esta época trascendió hacia un cambio global que ha transformado las relaciones política-ciudadanía (crisis de representación), los movimientos migratorios (crisis migratoria y crisis de refugiados) y las relaciones interestatales<sup>9</sup>. Si se compara la Gran Recesión con la Gran Depresión del siglo XX, Estefanía (2018) considera que los cambios fueron menos profundos, pero sí más amplios y más complejos. Además, destacados expertos (p. ej., Roubini, 2018) consideran el presente periodo de crecimiento es tan solo una recuperación temporal que preconiza una nueva crisis de mayor calado acuciada por los famélicos resultados de consumo, la precarización del trabajo, una nueva burbuja inmobiliaria y el aumento de las cifras de desigualdad poscrisis (Gordon, 2019)<sup>10</sup>.

La disminución de la riqueza y la disminución del consumo tienen una relación directa con el motor económico: la inversión privada y la inversión pública. Mientras que el sector empresarial estaba debilitado por el desplome en el consumo, el sector público estaba comprometido por los desvíos de fondos hacia el salvamento de la deuda de la banca privada. A pesar del rescate a la banca, los gobiernos no ejecutaron políticas de alivio a los sectores de la población deprimidos por la crisis. Desde que el sistema financiero mundial implosionó, la crisis ha sido un motor cultural generacional. La crisis afectó a los sectores más vulnerables, ya que, paradójicamente, fueron estos los que firmaron un número extremado de hipotecas de alto riesgo a pesar de que las entidades bancarias eran conscientes de que no estaban capacitados para conocer la dimensión del riesgo, traspasándose en muchas ocasiones los límites de la ética y la legalidad. La pérdida prolongada de puestos de trabajo, y la consiguiente sustitución por empleos precarios,

---

<sup>9</sup> Hablamos aquí del *Brexit* en el año 2016, las guerras comerciales de Estados Unidos contra la UE, Canadá y China en 2018, la consolidación de China como superpotencia o el colapso de la ola progresista en América Latina (desde 2016 en adelante).

<sup>10</sup> El número de superricos en España aumentó un 150% durante la crisis (EFE, 2018), mientras que el número de pobres en España pasó de un 23,8% en 2008 a un 27,9% en 2016.



junto con la consecuente caída del poder adquisitivo de los prestatarios provocó tasas de incumplimiento significativamente mayores que los deudores con hipotecas tradicionales de características similares. Como consecuencia, se produjo un nivel desproporcionado de desahucios en todo el mundo y se gestó una crisis de la vivienda en EE. UU. y Europa que continúa una década después. Entre 2006 y 2008, los ahorros y fondos de inversión de los ciudadanos estadounidenses perdieron 8,3 billones de dólares (Alman, 2009) y los niveles de endeudamiento hipotecario de los Estados Unidos en relación con el PIB aumentaron de un 46% en los años 90 a un 73% durante 2008 (Barr, 2009). Tras el estallido de la burbuja española en 2007-2008 la tasa del desempleo se disparó del 8% al 26%<sup>11</sup>, la mitad de esta cifra se produjo entre 2008 y 2009, en los albores de la crisis del euro.

Durante la crisis de la deuda, el PIB griego se desplomó un 6,9% en 2011<sup>12</sup> y 111.000 empresas fueron a la bancarrota. Como resultado, los griegos perdieron un 40% de poder adquisitivo, fomentado por unas cifras asfixiantes de desempleo que alcanzaron un récord histórico de 27,9% en 2013, con un desempleo juvenil por encima del 60% (Castiella, 2013). Para las mismas fechas, España había perdido un 10% de poder adquisitivo y las cifras de desempleo ascendían a 27,16% y un 57,2% de jóvenes estaban sin trabajo (López Manjón, 2013). Las cotas de malestar eran muy altas; sin embargo, los actores sociales institucionalizados como las grandes centrales sindicales españolas estuvieron especialmente ausentes. La reacción ciudadana europea en muchos casos tomó las calles de manera inicialmente espontánea, y posteriormente de manera auto organizada.

Tras de las distintas crisis (inmobiliaria, financiera y de deuda), subyacía una crisis de valores democráticos. Como se ha visto anteriormente, aparte del boicot de los distintos referéndums griegos por parte de la *troika* y/o el Eurogrupo, la opinión pública no fue motivo de consulta, mucho menos un sujeto de participación. La ausencia de un plan de rescate ciudadano que hiciera frente a las nefastas consecuencias de la crisis supuso un vacío de representatividad en Europa que ensanchó aún más la brecha entre la ciudadanía y los gobernantes. Asimismo, el desembolso de cifras milmillonarias trasladadas de los servicios sociales y dirigidas a rescatar la economía financiera se llevó

---

<sup>11</sup> Datos del Instituto Nacional de Estadística: <http://www.ine.es/jaxiT3/Tabla.htm?t=4247&L=0>

<sup>12</sup> Datos de la OECD : <https://data.oecd.org/greece.htm>



a cabo sin ningún tipo de consulta. Uno de los ejemplos más característicos de esta crisis de representatividad fue la reforma de la otrora “intocable” Constitución Española que, sin debate, modificó su artículo 135 tras la petición de Merkel y Sarkozy. Esta se llevó a cabo por la puerta de atrás en pleno verano de 2011. En el texto se introducía el concepto de “estabilidad presupuestaria”, primando la devolución de la deuda pública frente a cualquier gasto presupuestario del estado. A pesar de las movilizaciones en todo el país y las voces de autoridad en contra, la actuación monolítica de los dos grandes partidos llevó a cabo la reforma por la vía rápida sin necesidad de un referéndum (Verdú, 2011).

En esta década se produjo un cambio de paradigma social. A partir del cambio de década, Europa vivió el mayor ciclo de movilizaciones desde *mayo del 68*. Antes del Movimiento 15-M en España, se contaba con los antecedentes de decenas de huelgas generales antiausteridad en Grecia y la ola prodemocrática de la *Primavera Árabe*. El primer antecedente de protestas de nuevo cuño fue la llamada *Revolución de las Cacerolas* en 2009, tras el colapso de la entonces próspera economía financiera de Islandia. Tras varias jornadas multitudinarias de protesta frente al Parlamento en Reikiavik, el gobierno dimitió en bloque y se inició un nuevo proceso electoral que culminó con el rescate de la población, que impulsó un proceso constituyente que integró las nuevas tecnologías en los procesos de participación y culminó con un referéndum de profunda reforma constitucional en 2012 y varias consultas que decidieron no hacerse cargo de la deuda del fondo *Icesave*, de capital angloholandés, junto con el auge de nuevas políticas de profundización democrática y de limitación del poder del mercado.

Por otra parte, a principios de 2011, se inició un ciclo histórico de protestas en todo el mundo árabe. Tras el suicidio del joven vendedor ambulante tunecino Mohamed Bouazizi comenzó una oleada de ocupación masiva de plazas llevada a cabo por la ciudadanía organizada espontáneamente a través de las redes sociales. Entre las demandas estaban las mejoras en las condiciones de vida y la implantación de sistemas democráticos y de derecho. Este ciclo de empoderamiento ciudadano se extendió desde el Magreb a Oriente Medio, pasando por el cuerno de África. Tras meses de movilizaciones cayeron longevos regímenes, como el tunecino y egipcio, y se iniciaron otros nuevos: en el primer caso un régimen con algunas mejoras democráticas, mientras que el segundo devino en régimen militar.

Por otro lado, las consecuencias de las revueltas en países como Libia o Siria sustituyeron las demandas prodemocráticas de la población por el asedio de grupos terroristas de corte islamista que acabaron en guerra civil. El dictador de Libia, Muamar

el Gadafi, fue ejecutado a manos de la turba en 2011 y el estado libio, con uno de los mayores niveles de desarrollo humano en África hasta entonces, se desmembró a partir de dos guerras provocadas por injerencias internacionales por la explotación de sus vastas reservas de hidrocarburos, el regreso de las disputas tribales y la insurgencia de organizaciones terroristas internacionales.

Por otro lado, las facciones revolucionarias de Siria fueron derrotadas por insurgencias radicales islamistas que provocaron una guerra civil de escala internacional desde marzo de 2011. En amplias zonas del oriente del país, se erigió en el poder el llamado *Estado Islámico de Irak y Siria*, que dejaban a su paso un reguero de crímenes contra la humanidad en las ciudades que ocupaban y que detonaron la mayor ola de refugiados desde la Segunda Guerra Mundial. El estado de la República Árabe Siria no solo debió enfrentar a los terroristas, sino a las injerencias internacionales, principalmente el apoyo tácito de Turquía al Estado Islámico y servir como campo de batalla de los intereses imperialistas de Estados Unidos y Rusia. Entre los puntos de claridad de este contexto, hay que destacar la insurgencia de las Fuerzas Democráticas Sirias, una alianza multiétnica con predominancia kurda que, además de ser un actor no alineado fundamental para la liberación de territorios ocupados por los terroristas, llevó a cabo las demandas populares de autonomía. Este proceso ha generado un proceso constituyente de abajo hacia arriba con amplia participación de todos los sectores populares, que tiene como ejes ideológicos transversales el feminismo y el ecologismo.

Por lo que respecta a España, la ola de protestas que dieron lugar a los nuevos movimientos sociales en España fue, sin duda, uno de los fenómenos populares más influyentes en este tiempo y resultaron realmente innovadoras en cuando a praxis organizativa y en amplitud de objetivos. El 15 de mayo de 2011 se organizaron más de cincuenta manifestaciones en ciudades de todo el estado bajo un lema que resumía de manera clara las demandas de mayor participación ciudadana y respuesta ante las medidas de austeridad imperantes: “Democracia Real Ya: No somos mercancías en manos de políticos y banqueros” (Verdú, 2011). Un grupo de marchantes decidieron emular las acampadas de la *Primavera Árabe* y tomar la Puerta del Sol en Madrid. Tras ser desalojados por la fuerza la mañana siguiente, centenares de personas convocaron una asamblea popular en el mismo lugar los cuales habían sido citados descentralizadamente mediante distintas plataformas de redes sociales y consolidaron la ocupación de la plaza hasta finales del mes de junio. Dos días después, más de un centenar de ciudades habían

secundado el llamamiento de ocupar las plazas de las principales ciudades españolas y europeas, extendiendo las demandas a amplias capas de la población.

Sobre el estudio del Movimiento 15-M a través de sus datos, es necesario destacar el trabajo coordinado por Toret Medina (2015) sobre tecnología, multitudes conectadas y efectividad comunicativa. Entre otros factores, en el texto se recoge que el movimiento fue pedagógico, especialmente en educación económica, y rompió con el consenso discursivo imperante. Sus estrategias de comunicación no solo pusieron sobre la mesa asuntos relacionados con la austeridad económica (p. ej., la popularización de términos como *deuda ilegítima*), sino como sistema de transmisión de emociones y simbología. A este fenómeno comunicativo le llamaron “contagio tecnológicamente estructurado”. En resumen, se puede hablar del Movimiento 15-M como un *movimiento-red*, que hizo evidente el vacío de representación imperante y que propuso nuevos métodos de difusión del discurso especializado imperante, y divulgó, mediante una combinación virtuosa de canales tradicionales y digitales, conceptos y términos especializados a amplias capas de la población que no estaban relacionadas con el discurso económico y político.

Esta vulgarización del discurso económico fomentó la comprensión de los detalles más obtusos y escandalosos del sistema político y económico y abrió una grieta en el régimen bipartidista imperante. Además de otros factores, supuso la entrada de nuevos actores sociales y políticos en el país como una praxis experimental que dio lugar a fenómenos de naturaleza muy diversa. Esta diversidad se ha materializado políticamente, i.e. la red de candidaturas municipalistas que lograron formar gobierno en los ayuntamientos de Madrid, Barcelona y Zaragoza o la entrada de Podemos en el arco parlamentario en diciembre de 2015. Fuera de la política institucional, la plataforma *15MpaRato* financiada a través de técnicas novedosas de financiación popular (*crowdfunding*) que fue alentada mediante una innovadora campaña de difusión y de divulgación de las actividades ilícitas detectadas a través de filtraciones y ciberactivismo. Esta campaña consiguió la condena a prisión al exministro español de economía y ex jefe del FMI Rodrigo Rato así como a más de medio centenar de consejeros de *Bankia*, una banca nacionalizada, tras destapar varios casos de corrupción como las conocidas como *tarjetas black* (Serrano, 2014).

El auge de los movimientos sociales surgidos a partir de la crisis supuso un revulsivo político progresista a nivel mundial. No obstante, muchos países del mundo occidental han sufrido reveses populistas significativos. En 2016, el éxito electoral de Donald Trump en los Estados Unidos es el mayor indicador de una serie de movimientos

populistas que vienen fortaleciéndose en las clases populares de Europa y América. En junio de ese año, la palabra *brexit*<sup>13</sup> inundaba los titulares de medio mundo alentados por las alas extremistas del Partido Conservador y el euroescéptico *UKIP*.<sup>14</sup> Lo que en apariencia había sido un término y un movimiento de carácter marginal, obtuvo un 51,9% de los apoyos en el referéndum celebrado ese día. En el continente europeo, el Frente Nacional francés fue la primera fuerza de voto tras recibir el 27% en las elecciones regionales de 2015, mientras que en Italia unas elecciones sin mayorías impulsaban a la vicepresidencia a Matteo Salvini, abiertamente caracterizado por un discurso y una praxis política euroescéptica y racista. El mundo germánico vio la llegada del primer ministro Sebastian Kurz en Austria, mientras que Alemania vio el auge electoral de Alternativa para Alemania (*Alternative für Deutschland*), ambos tras movimientos ultraderechistas alentados por la austeridad y la propaganda política antagónica tras la mayor oleada de refugiados desde la Segunda Guerra Mundial, procedente de la crisis en Oriente Medio.

El discurso económico y su calado en la población ha sido un campo de batalla durante la crisis. Si bien hay causas específicas para cada sociedad, el principal motivo es una crisis de representación y de modelo económico tras el abandono de los pactos sociales y el salvataje a toda costa de la banca en detrimento de las clases populares. Como ya hemos evidenciado, todo esto surge en gran parte como reacción de los sectores mayoritarios de la población a las arengas carismáticas y discursos anti-élites simplistas de los nuevos líderes populistas, que proponen respuestas rápidas ante situaciones nuevas, como la inmigración o los desafíos culturales de la posmodernidad.

## 2.2 La influencia de la crisis y la prensa en el lenguaje económico

Este trabajo, está motivado a partir del estudio del lenguaje basado en la especialidad y el análisis de la evolución del lenguaje de opinión. En él se han utilizado extensivamente muestras de lenguaje periodístico especializado, por lo que es pertinente justificar por qué estudiar el discurso económico a partir de textos periodísticos. Bednarek y Caple (2012) justifican esta motivación por tres razones principales: (a) la abundancia de textos de este

---

<sup>13</sup> Este término *portmanteau* se había popularizado en la prensa como analogía al *grexit* que estuvo en boga en la prensa durante las negociaciones entre el gobierno de Tsipras y el Eurogrupo en el que se amenazaba con la salida de Grecia de la UE.

<sup>14</sup> *United Kingdom Independence Party*.

tipo, (b) la facilidad para obtenerlos y, finalmente, (c) la gran exposición que el público tiene a ellos, a través de los distintos medios, por lo que este discurso genera una influencia destacada en el consumo lingüístico, y posterior producción, de los hablantes de una lengua.

En el contexto de esta década, dichas autoras consideran que las noticias se comparten y difunden más que nunca en la historia, hecho en el que las redes sociales han jugado un rol primordial. De esta manera, los medios tienen más poder que nunca en la historia para influenciar la opinión y moldear las opiniones, de manera que el uso del lenguaje modifica la opinión, ideas y creencias, así como el significado y valoración de las palabras utilizadas. El uso exponencial de las redes sociales ha servido como gran trampolín de difusión de nuevos usos y sentidos de palabras y términos a través de los artículos que los creadores de opinión han ido generando a lo largo de estos últimos años, contribuyendo estos, entonces, a la construcción de un sistema de valores, según Thompson y Hunston (2000), construido por la comunidad de hablantes a través de las evaluaciones. Este sistema trasciende como componente de la ideología que rezuma a través de las combinaciones y construcciones lingüísticas de cada uno de los textos.

La crisis ha sido considerada un campo de pruebas para experimentar con todo tipo de recursos expresivos para describir la naturaleza de una situación económica extremadamente complicada. El estallido de la burbuja inmobiliaria, además, se tradujo en consecuencias nefastas para las condiciones de vida de la población: desempleo prolongado, recortes en los servicios básicos y la asistencia social, además de una crisis de vivienda que supuso cientos de miles de ejecuciones hipotecarias y desalojos. A partir de estos eventos se inauguró una nueva época discursiva en la que primó el uso del lenguaje económico en el discurso. El caso discursivo inaugural, y quizá el que mejor refleja el cambio de paradigma de esta época, fue el llamamiento de Nicolas Sarkozy para refundar el capitalismo expuesto anteriormente (Luna-Alonso, 2016).

El empeoramiento de las condiciones de vida en grandes sectores de la población ha producido cambios sociopolíticos drásticos que han puesto en el foco a las noticias económicas y, por lo tanto, al lenguaje económico. Lo relevante para este trabajo es que tales reacciones políticas y sociales a las consecuencias de las medidas de austeridad han influido en el discurso público y, consecuentemente, el uso del léxico. Por ello, los términos financieros especializados que solían ser utilizados exclusivamente por expertos se han convertido de la noche a la mañana en un fenómeno común en las noticias generales y han sido adoptados progresivamente por el público (Holborow, 2015).

Al examinar el uso del vocabulario de la crisis en medios de comunicación influyentes como The Guardian o The New York Times, nos enfrentamos a un lenguaje que se refiere a eventos mediante el uso de términos cuidadosamente seleccionados que parten del dominio especializado de la economía, mezclado con elementos que referencian la política contemporánea y la vida pública. Como si de una correa de transmisión lingüística se tratase, estos términos del metalenguaje son asimilados rápidamente por el público y reutilizados de manera natural por otros medios de comunicación. Como consecuencia, el lenguaje se ha adaptado a esta nueva realidad y evolucionado a la necesidad expresiva de su tiempo. Los términos financieros y económicos se pusieron en primer plano, y las actitudes hacia ellos son susceptibles de haber cambiado. Por ejemplo, expresiones como “toxic assets” o “green shoots” han sido utilizadas especialmente de manera metafórica y en gran parte eufemística.

Para responder a la demanda del público, se puso en marcha a la maquinaria editorial para explicar la terminología y a partir de 2008 una amplia gama de recursos lexicográficos salió a la luz: desde *glosarios de la crisis* a diccionarios terminológicos completos (p. ej., Pinedo y Martínez, 2012) para dar luz a la jerga que inundó las portadas y los informativos televisados. Por otro lado, en muchos casos los periodistas ejercieron de agentes divulgadores de manera explícita para restar opacidad a los términos y ofrecer una comprensión asequible a sus lectores<sup>15</sup>. Otras estrategias de divulgación económica han tenido éxito a la hora de observar las distintas estrategias seguidas para reformular la terminología: desde el parafraseado a conjuntos pregunta-respuesta o el empleo de metáforas. Detrás de estas técnicas se revela la necesidad de hablar del contexto y el mundo que rodea a los términos, de forma que estas son la muestra de la reelaboración de un discurso a partir de otro discurso previo (Rojo López y Orts Llopiz, 2010; Janot, 2012).

Sobre este proceso de asimilación lingüística a partir de la cadena *poder financiero-prensa-público*, Haase (2010) expresa que, especialmente en la descripción de crisis en la economía, los datos no hablan por sí mismos y destaca la enorme importancia

---

<sup>15</sup> Existen una gran cantidad de ejercicios de recopilación de glosarios en muchos idiomas destacando aquí algunos. En español: “Nuevo diccionario de la crisis” (*BBC Mundo*, 16/08/2011), “El diccionario de la Gran Recesión” (*El Periódico*, 08/08/2017). En inglés encontramos, entre muchos: “Financial Crisis Glossary” (Web del Parlamento Europeo, 01/06/2010) o “Eurozone crisis glossary” (*The Guardian*, 01/06/2012). En francés y en italiano, Janot (2012) destacan “Le mots de la crise” (*L’Expansion* 12/12/2008) o “Il nuovo ABC dell’economia” (*Il Sole 24 Ore*, 10/2009).

de la retórica utilizada por los medios masivos a la hora de explicar el calado de la crisis económica a sus lectores. El léxico utilizado en la difusión de las noticias contribuye así al efecto global de la crisis que se presenta en los medios de comunicación, ya que la presentación carece de información diferenciada de trasfondo. Este relato no es monodireccional y, como contraparte, la jefa de creación de contenido en *Oxford Dictionaries* recogió algunos apuntes sobre el léxico económico de la célebre acampada protesta *Occupy Wall Street* y realizó un pequeño estudio longitudinal del mismo. Martin (2012) considera que, si bien estas palabras alcanzaron gran difusión en la prensa y se extendieron a otras ciudades movilizadas, estas no tuvieron un impacto significativo en la forma en que la ciudadanía usaba el inglés. Sin embargo, algunos términos alcanzaron relevancia puntual, por ejemplo, “the 99%” (en relación con la ciudadanía afectada por la crisis) o “1-percenter” ‘los poderes financieros’.

Existe un uso intencional del lenguaje especializado a la hora de transmitir no solo información técnica, sino ideología. En relación con esto, Moirand (2007) afirma que el consumo del lenguaje de los medios se expresa con un costo financiero y social colectivo. En muchas ocasiones los medios y las instituciones hacen un uso innecesariamente opaco de la terminología, que a su vez genera una ráfaga semántica que incluye connotaciones de sentimiento. La ideología imperante del neoliberalismo empapa el discurso con sus valores, y autoras como Massey (2013) opinan que el vigente sistema neoliberal ha copado con una rapidez virtuosa el lenguaje de lo cotidiano, de forma que moldea la concepción de nosotros mismos y nuestra forma de aproximarnos del mundo. Así, en determinados entornos han empezado a denominar con naturalidad como *customers* a los pasajeros del tren o a los pacientes de un hospital. Sobre la velocidad de estos cambios, algunos autores, p. ej., Vincent (2011), hablan de la existencia de rápidos cambios en el sentimiento de las palabras que no pueden ser reflejados en el estatismo de los diccionarios de cabecera. El autor califica este fenómeno como “guerra semántica”, donde incluso la mera mención a *Wall Street* rezuma emoción.

En estos términos de orientación semántica, desde el inicio de la crisis en 2008 el uso polarizado del lenguaje financiero fue cuestionado por especialistas y críticos. En muchos casos se utilizaron términos con una orientación semántica positiva para blanquear efectos beneficiosos principalmente para las grandes finanzas, de manera que puedan ser mejor aceptados por el público. Mientras la *troika* hablaba de “solidarity”, este término pasó de hablar sobre la cooperación y apoyo mutuo a una *solidaridad* de contratos firmes. El bono de 86.000 millones de euros de solidaridad a Grecia en 2015 supuso,



primero, férreas condiciones para la devolución de ese mismo préstamo a los prestatarios (50.000 millones de euros), 25.000 para recapitalizar a la banca privada y solo 11.000 a reactivar la economía del país, con la condición de profundizar en las medidas de austeridad, que a la vez la hacían más dependientes de inyecciones de capital y a una austeridad que agrava el problema de la economía del país receptor de tal solidaridad.

Otro ejemplo es el uso de “structural reforms”, el verbo “reformar” ha tenido de manera preponderante una orientación semántica positiva: algo que es reformado es algo que va a modificarse progresivamente para funcionar mejor. La colocación con la palabra “structural”, por tanto, debe hacer referencia a una mejora de las estructuras de un país para que las falencias mejoren. Sin embargo, las reformas estructurales (a cambio de su “solidaridad”) impuestas por el FMI a Grecia, cinco años después de la primera imposición, supusieron un crecimiento de la deuda del 100 al 187% en relación con su PIB sin una mejora cualitativa de la economía, más bien todo lo contrario: despido de funcionarios, la subida de impuestos a las rentas bajas, y la privatización de los recursos del país en beneficio de fondos de inversión extranjeros, esto es, los “mercados” (Jiménez Gómez, 2015). Finalmente, la fascinación de los sistemas capitalistas con el crecimiento, sin tener en cuenta otros factores, hace medir cualquier situación económica en relación con este concepto. Es el caso de la popularización de “negative growth”, que ha sido utilizado ofuscadamente, pues generalmente ha reemplazado un uso técnico de “recession”, que tiene connotaciones abiertamente negativas. De esta manera, la evaluación negativa queda embebida en una estructura difícil de procesar tanto por hablantes no especializados como por algoritmos analizadores del sentimiento. Por ejemplo, en el discurso de Jean-Claude Trichet, encontramos el uso de este término en oraciones en las que el término negativo queda difuminado dentro de un sintagma relativo en la oración:

Taking all these measures and their effects into account, as well as the pronouncedly NEGATIVE GROWTH in the first quarter of this year and the most recently published forecasts by public institutions, the risks to this outlook remain broadly balanced.<sup>16</sup>

Alim (2011), refleja en una columna para el New York Times la influencia lingüística de los movimientos contestatarios surgidos como reacción para defender a la

---

<sup>16</sup> Rueda de prensa del BCE del 07/05/2009, consultada en el corpus Dijon ECBPress-Conf (Fernández-Cruz, *et al.*, en prensa).



población de los embates de la crisis. Afirma que, por ejemplo, el fenómeno *Occupy Wall Street* de 2011 inició un giro en la narrativa del debate de tal forma que movió algunos de sus términos del foco central, tales como “debt ceiling” y “budget balance”, y colocaron en el candelero a otros como “inequality” y “greed” como términos de debate socioeconómico. Considera que este cambio discursivo produjo variaciones en la argumentación discursiva del “establishment” sobre las razones del colapso financiero y sus consecuencias. En la misma época, el director del Knight Centre for Digital Media Entrepreneurship de la Arizona State University, Gillmor (2011), reflexiona en *The Guardian* sobre la evolución de los usos del lenguaje económico, y como sus connotaciones en la prensa suelen reflejar los intereses de los grandes poderes. Pone de ejemplo que uno de los significados del verbo “earn” es “to profit financially”, que va siempre combinado en la prensa con la segunda acepción asignada al merecido de la ganancia. Cuestiona entonces sobre si, dada la situación del momento, las enormes ganancias y activos acumulados durante la crisis del 1%<sup>17</sup> merecen el uso de la palabra *earn* en toda su acepción, cuando pone sobre la mesa las acciones y la impunidad de los ejecutivos de Wall Street durante los eventos que afectaron gravemente a la economía mundial. Asimismo, cuestiona el uso partidista del lenguaje que los lobbies proponen a la prensa, un uso del lenguaje especializado cargado de subjetividad, como es el caso del reemplazo del aséptico “inheritance tax” favorecido por el uso del subjetivo “death tax” con vistas a influir en la opinión política. Pero el lenguaje emocionalmente sesgado, especialmente cuando ha sido moldeado a conciencia por los poderes mediáticos, ha servido de herramienta esencial para los poderes financieros y sus aliados a obtener ganancias tan grandes como polémicas.

Trascendiendo el dominio de las finanzas, recoge Massey (2013) que, en los últimos años, el término “work”, ha evolucionado en su significado, de tal forma que solo se tienen en cuenta las transacciones de dinero como pertenecientes a la economía, invisibilizando los trabajos no remunerados relacionados con los cuidados o la creación, en gran parte feminizados, de manera que son solo apreciados y disfrutados en la intimidad de las familias y las comunidades. La autora reflexiona que, especialmente en los últimos años, la palabra *trabajo* solo se visibiliza con connotaciones desagradables, y

---

<sup>17</sup> “The 1%” es un término popularizado a partir del Movimiento Occupy Wall Street, y hace referencia a los super ricos, en contraposición al “99%”.

que tan solo ofrece como recompensa el disfrute de bienes de consumo. Pone otros ejemplos de empleo de términos económicos en el vocabulario general, como es el caso de la referencia a los términos agrupados en torno a la inversión y el gasto, y las connotaciones morales que van implícitas junto en su uso. Mientras que una “inversión” (“investment”) implica actividad, incluso sacrificio, con vistas a un futuro mejor resultando en un elemento de sentimiento positivo, la selección de palabras con significado similar, como es el caso de “gasto” (“expenditure”), aparece simplemente un expendio o una carga económica. De esta manera, cuando se habla de la construcción de la infraestructura de una escuela hablaremos de investment, pero el pago a los docentes, al cuerpo administrativo, o al personal de comedor implicará hablar de expenditure.

Finalmente es necesario destacar la relevancia de la prensa como factor principal de transmisión de opinión. En ella, la evidencialidad se expone a través del uso del lenguaje especializado, pues será el que asigne mayor o menor relevancia a los eventos narrados (Haase, 2010). Este uso de la lengua de especialidad tiene un estilo uniforme altamente retórico, metafórico y evaluativo. Arrese y Vara-Miguel (2015) revisan la bibliografía entre las distintas lenguas europeas. Muchas investigaciones durante esta década han observado este asunto desde múltiples perspectivas a través del uso de los informes económicos para la prensa en distintas lenguas europeas con conclusiones similares: a pesar la frecuencia de uso y las variables idiomáticas, la naturaleza de las distintas metáforas económicas resulta bastante uniforme en distintas lenguas, países y tipos de cabecera de periódico. Además, Rojo López y Orts Llopiz (2010), en un estudio comparativo entre el inglés y el español observaron que los artículos periodísticos previos a la crisis tienen un mayor número de metáforas con sentimiento negativo en artículos periodísticos en ambos idiomas. De hecho, cerca de cumplir el rol de cliché, los términos se presentan como unidades semánticas en los escritos de los medios como concepto semiótico de las rupturas que ocurren en el lenguaje de dominio económico. En el caso del primer estudio, el lexicón económico de la crisis que se usó incluye algunos de los términos que se van a revisar en este trabajo, como “credit”, “inflation” o “spending”. Estos términos se encontraban prominentemente en las secciones económicas de los periódicos o en la jerga económica y, tras la crisis, saltaron al dominio general. Ahora parecen encontrarse más fácilmente en el uso cotidiano de la lengua y suelen utilizarse en segmentos evaluativos.

Como muestran estos estudios, los diarios son uno de los formatos textuales más interesantes para el análisis del lenguaje especializado económico y financiero, pues

presentan de manera asequible al público el discurso entre las élites mientras que, a su vez, influyen al resto de las secciones, y recopilar la información de estos es muy accesible. El modelo genérico de medio de comunicación de masas ha sido complementado por análisis en los procesos de comunicación entre las élites, que hacen uso de publicaciones especializadas, y especialmente el papel de la prensa como mediador entre el público y los actores que toman las decisiones económicas. La información de los artículos estaría dentro de la agenda de los mercados y tendería a excluir los intereses del público, de manera que “financial and business news coverage reproduces the prevalent ideas, norms and values of those who work in these sectors”<sup>18</sup> (Davis, 2011, p.2, citado en Arrese y Vara-Miguel, 2015). La cadena de opiniones de los profesionales de la economía, los miembros de las instituciones y, en el último eslabón, la ciudadanía se sustenta en gran medida la construcción formada desde las que otros tienen sobre esos mismos eventos a partir de lo publicado en las cabeceras editoriales y sus distintos avatares divulgadores: periódicos, blogs y redes sociales.

Las noticias financieras pertenecen a géneros de comunicación considerados “recursos estratégicos” para el sentido colectivo y la persuasión (Fairclough, 2015). El análisis de la voz y opinión, representados en las noticias de índole económica o financieras, permite una aproximación a las relaciones de poder entre los grandes grupos editoriales y los grupos económicos, ya que estos últimos poseen la influencia económica o simbólica para poder ampliar sus puntos de vista con el objetivo de llegar al gran público. Se observa entonces que corporaciones y gobiernos copan la mayoría de los titulares y, por ende, sus evaluaciones se convierten en las perspectivas dominantes, especialmente si estas son asimiladas por los periodistas y finalmente no encuentran cuestionamiento por la ciudadanía. Como consecuencia, los medios se convierten en una forma de entrada del lenguaje al gran público y uno de los principales transmisores de opinión sobre las visiones del mundo (Rocci, Palmieri y Gautier, 2015).

### 2.3 Contenidos en red y opinión

El clima de ansiedad y descontento que ha impregnado uniformemente el discurso de la crisis ha avivado la necesidad de buscar nuevos enfoques para poder estudiar con mayor

---

<sup>18</sup> La cobertura de la prensa financiera y de negocios reproduce las ideas preponderantes, las normas y los valores de quienes trabajan en esos sectores.

profundidad la influencia del sentimiento en la comunicación y en los mercados, buscando nuevas formas de establecer formas de medirlo de manera confiable y precisa. El análisis de sentimiento surge a principios del siglo XXI para abordar el problema del análisis de la opinión a través de un enfoque computacional y es en esta época cuando gana mayor prominencia.

Tradicionalmente se han venido realizando costosas mediciones a través de encuestas de opinión o, dado el caso, análisis textuales para realizar estudios sobre la opinión y conocer sus tendencias. El acceso a esta valiosa información, dada la necesidad de despliegue, es muy costosa y exclusiva, ya que solo las organizaciones pudientes pueden realizar estos estudios. La ciudadanía de a pie y pequeñas organizaciones suelen no poder acceder de primera mano a estas informaciones, sino a través de los informes que estas envían por obvio interés propio a los medios de comunicación. Este sería el caso de los sondeos de opinión internos que, por ejemplo, los partidos políticos ponen al servicio de la prensa en el caso de que les sean favorables. Por esto, el auge de la Web 2.0 revolucionó el mundo de la información y las comunicaciones al democratizarlas y dotarlas de mayor horizontalidad a partir de la implementación de redes distribuidas y multidireccionales. Por otro lado, la Web proporciona mayor difusión y acceso a medios de comunicación de toda índole y procedencia. En ella caben desde las nuevas plataformas digitales de los imperios mediáticos (hablamos aquí de los oligopolios informativos *transmedia* como News Corp o PRISA en el mundo hispanohablante), a nuevos métodos de base, tales como radios y televisiones comunitarias y medios online ya históricos como IndyMedia o Democracy Now y medios principalmente sostenidos por sus suscriptores como eldiario.es o La Marea, y el universo de *podcasts* y *blogs*. Además de las líneas editoriales, no se puede obviar en este apartado lo que se ha llamado *user-generated content* o UGC. Su aplicación en el territorio de la opinión, el *electronic word-of-mouth* (eWOM) supone un reemplazo 2.0 tradicional del boca a boca pues aporta información contrastable sobre cualquier entidad procedente de pares en lugar de la publicidad ofertada por el propio proveedor del servicio o producto (Cheung y Thadani, 2012).

Los UGC se entienden como materiales digitales creados por los usuarios de manera no amateur para difundir ideas o contenidos por el mero hecho de compartir, o para influenciar a otros. Estos contenidos pueden ser de distinta naturaleza: entradas de blogs, sección de comentarios en noticias, páginas de reseñas de restaurantes, podcasts, etc. No sorprendentemente, dada la portentosa plasticidad del capitalismo, en el último

lustro han surgido centenares de llamados *influencers*, usuarios que promocionan de manera semiprofesional productos a partir de reseñas y opiniones remuneradas o, directamente, las compañías generan contenidos que simulan su estética con fines puramente comerciales y que, además, evidencia la potencia de estos formatos. Tanto los UGC como los materiales editoriales de esta generación Web facilitan la extracción de información mediante técnicas computacionales a partir del procesamiento de muestras de lenguaje real rico en opiniones (Balasubramaniam, 2009; MacKinnon, 2012)

A través de distintos enfoques y técnicas informáticas englobadas en lo que se ha llamado procesamiento del lenguaje natural (PLN o, en inglés, NLP), se pueden estudiar estas opiniones, destacando los estudios de análisis de sentimiento. Los sistemas de análisis de sentimiento suponen una gran ventaja ante los métodos seguidos tradicionalmente como el análisis de informes financieros o de opinión pública, o en índices publicados periódicamente como el *Indicador de Sentimiento Económico de la Zona Euro*, entre otros. En la actualidad, podemos disponer al minuto de información del sentimiento directamente sobre los distintos sectores económicos desde las valoraciones inscritas en los textos de las fuentes deseadas: prensa, instituciones o compañías (Malo et al., 2014).

## 2.4 Conclusiones

Este capítulo ha planteado un marco contextualizador de la tesis que permite mejorar la comprensión de los factores sociales que influyeron en el planteamiento del estudio. Se ha dividido en tres partes: una contextualización histórica, una contextualización de la influencia de la crisis y la prensa en el uso del lenguaje económico y, por último, una breve exposición sobre la influencia y posibilidades de los contenidos en red y la opinión en la actualidad.

La primera parte revisó de manera cronológica los eventos principales de la crisis global: primero el estallido de la burbuja inmobiliaria que dio lugar primero a la crisis de las hipotecas subprime y posteriormente a la crisis crediticia y, después, la crisis de la deuda, centrada principalmente en la Zona Euro. Posteriormente trató sobre el impacto de las consecuencias de la *Gran Recesión* en la vida cotidiana de la ciudadanía.

La segunda parte observó la influencia de la prensa como mediadora del lenguaje económico durante la crisis. Se revisa por qué, repentinamente, los términos financieros

---

y económicos se pusieron en primer plano y se cuestiona sobre los usos ideológicos y los posibles cambios en la actitud del público hacia el discurso económico.

Finalmente, la tercera parte repasó la relación asimétrica del discurso en Internet y la emergencia del análisis de sentimiento. En esta esfera colisionan las grandes cabeceras y la creación de contenidos por parte de los usuarios. Si bien las oportunidades son mayores hoy en día, existen nuevos desafíos a la hora de conocer la opinión sobre la actualidad económica.



### 3. El lenguaje evaluativo

« Las opiniones no se borran con fuego. »

José Mejía (1775-1813), diputado por Quito en las Cortes de Cádiz. Cita inscrita en el mural *Imagen de la patria* (1988) de Oswaldo Guayasamín (1919-1999) que preside desde entonces la Asamblea Nacional del Ecuador

#### 3.1 Introducción

Tanto las emociones como las opiniones condicionan la manera en la que los humanos se comunican entre sí y motivan sus acciones. El lenguaje evaluativo, según Stubbs (1986), se usa para expresar creencias personales y adoptar posiciones, para expresar acuerdo o desacuerdo con otras personas, para establecer alianzas, contratos y compromisos o, de manera alternativa, para reflejar un sistema de valores, disociar al hablante de determinados puntos de vista y vaguedad conscientemente o no comprometerse con lo expresado (Du Bois, 2007; Hunston, 2011; Moreno-Ortiz y Pérez-Hernández, 2014).

En el uso cotidiano del lenguaje existen determinadas palabras que tienen un potencial ilimitado para expresar opinión. Palabras como “excellent” o expresiones como “it’s a shame that...” tienen una aplicación casi universal para expresar opinión en cualquier dominio. Mientras, otras expresiones tienen un contenido más referencial, pero en su significado se incluye un elemento de opinión que incluye más o menos carga emocional y más o menos inherente a la expresión. Por ejemplo, *jam* puede entenderse como “mucho tráfico” + “malo”. Aunque no es siempre una cuestión cristalina, la opinión es una característica del lenguaje esencial para poder describir completamente el significado de un texto (Thompson y Hunston, 2000).

El estudio del lenguaje evaluativo ha despertado el interés de múltiples disciplinas: desde los distintos campos de las humanidades como la filosofía, la sociología o la ciencia política, como para la ingeniería a través la informática afectiva y la inteligencia artificial. Esto último ha acompañado a la emergencia de Internet 2.0, especialmente a través de las redes sociales que permiten a sus usuarios la difusión de emociones, evaluaciones y opiniones a una escala a la vez horizontal y mundial (Benamara, Taboada y Mathieu, 2017).



El estudio humanístico de los valores se remonta a las disertaciones clásicas sobre ética y retórica, por ejemplo, en el estudio de la universalidad de la “idea del bien” en su sentido platónico. Con la emergencia del idealismo alemán, Kant (1785) considera el valor moral como un valor absoluto, único e identificable universalmente. Esta universalidad de los valores fue puesta en cuestión por el pragmatismo norteamericano (p. ej., la teoría de la valoración<sup>19</sup> de Dewey, 1939) que negaba la existencia del valor intrínseco como una propiedad inherente a las cosas, relativizando su valoración y, por lo tanto, necesitaría de un estudio en profundidad del discurso y la naturaleza intersubjetiva de los valores. En la actualidad, Habermas y Apel, filósofos de la segunda generación de la Escuela de Frankfurt, han reformulado la ética deontológica kantiana para estudiar el análisis de las estructuras comunicativas de una manera intersubjetiva: se discute la validez de una corrección normativa que depende enteramente de la intersubjetividad, es decir, las normas dependen del mutuo acuerdo entre todos los actores que forman la comunidad discursiva (Alba-Juez y Thompson, 2014).

El estudio de los sistemas de valores humanos es una tarea muy compleja que puede afrontarse en buena medida a partir del análisis del estudio del lenguaje evaluativo en uso. Desde un punto de vista lingüístico, se han propuesto distintos modelos de observación para establecer un modelo de relación entre el contenido factual o referencial y el contenido afectivo o expresivo. Sin embargo, la tarea de identificar el lenguaje evaluativo no es una tarea simple, y desde el siglo XX se han propuesto algunos modelos lingüísticos para resolver esta tarea, entre otros: representativo/expresivo (Buhler, 1934), descriptivo/expresivo (Lyons, 1977), ideacional/interpersonal (Halliday y Matthiesen, 2014), expresivo/relacional (Fairclough, 2015). Aunque la mayoría de los autores se han centrado en distintos tipos de texto, y características del lenguaje evaluativo, algunas teorías han ofrecido recuentos más concisos de la naturaleza del lenguaje actitudinal, como *modalidad y modulación* (Halliday y Matthiesen, 2014); *orientaciones evaluativas* (Lemke, 1998); *posición o postura* (Biber et al., 1999; Conrad y Biber 2000; Hunston, 2011, *inter alia*); *evaluación* (Thompson y Hunston, 2000; Bednarek, 2006; 2008) o *valoración* (Martin, 2000; Martin y White, 2005).

---

<sup>19</sup> *Theory of valuation*, en su redacción original, no confundir con la teoría de la valoración de Martin y White (2005), llamada en inglés *Appraisal theory*.

En este capítulo se resumen algunos conceptos y términos asociados que se han venido utilizando por los distintos autores para el estudio del lenguaje evaluativo y, con especial empeño, se revisarán los aspectos principales de la teoría de la valoración (Martin y White, 2005) y los enfoques de evaluación de Hunston (2011), fundamentales para entender los principios generales de la lengua que rigen la expresión de la evaluación. Estos sirven, en gran medida, como base para el estudio de la evaluación desde un punto de vista computacional que se estudiará con más profundidad en el capítulo siguiente.

Los enfoques computacionales del estudio de la evaluación están representados principalmente por la disciplina del análisis de sentimiento. Esta trata los estudios automatizados de la expresión y la extracción automática de la opinión en tres niveles principales de estudio (Liu, 2012, 2015): el documento, la oración y el aspecto. El primer nivel tiene como objetivo clasificar los documentos globalmente como positivos o negativos, mientras que el segundo trata de determinar si la orientación es subjetiva, para posteriormente catalogar la opinión (positiva o negativa) de cada una de las oraciones; el tercero extrae cada uno de los distintos aspectos o entidades tratados en un texto, para posteriormente asignarles una puntuación positiva o negativa.

Exceptuando la teoría de la valoración de Martin (Martin y White, 2005), ninguna teoría había propuesto observar holísticamente los mecanismos lingüísticos utilizados en la expresión de la evaluación. Antes de entrar con detenimiento a dar cuenta de esta teoría, se realizará una panorámica de las principales aproximaciones al lenguaje evaluativo desde la lingüística. En este campo, se utiliza una amplia gama de términos para hacer referencia a zonas similares de uso del lenguaje: *valoración*, *postura*, *sentimiento*, *lenguaje evaluativo*, *lenguaje actitudinal*, o *lenguaje afectivo*. Aunque estos comparten un vasto espacio de trabajo y en usos generalistas pueden considerarse *grosso modo* equivalentes, es posible discutir en amplitud las diferencias entre dichos términos.

Tras las distinciones entre ellos, hay una variación en el tipo de fenómeno de evaluación que se lleva a cabo. De esta manera, se destacan las siguientes características comunes:

- La evaluación es una función realizada por el texto al completo o por una parte de un texto. Se trata el texto como una entidad independiente de las personas que interactúen en él.
- En su vertiente más sencilla, la evaluación se observa inicialmente como un conjunto de palabras y sintagmas que expresan significado evaluativo (Conrad y Biber, 2000).

- La evaluación es una acción, es decir, algo que alguien desempeña de manera consciente y activa. Dicha acción puede ser realizada y estudiada a partir del uso verbal abierto o de manera encubierta a través de una expresión aparentemente no evaluativa.
- La evaluación es un conjunto de significados que pueden ser expresados en cualquier texto a través del uso de una amplia variedad de recursos lingüísticos. A partir del desarrollo sistémico-funcional de la teoría de valoración por Martin y White (2005) se hace referencia a un sistema de significados por el cual un hablante usa la lengua como herramienta para aprobar y desaprobar, entusiasmarse y aborrecer, e instar a sus oyentes a hacer lo mismo. Desde una perspectiva dialógica heredera de Bajtin, observan las acciones dentro del discurso, en particular las acciones tripartitas de expresar actitud, posicionamiento y alineamiento con respecto a una audiencia determinada o un interlocutor.

Teniendo en cuenta lo anterior, es necesario hacer referencia a la fase cognitiva previa a la expresión textual de la evaluación basada en la experiencia y el contexto de los hablantes. Aunque su estudio traspasa los límites de este trabajo, eminentemente lingüístico, en esta fase el hablante realiza internamente una triple elección: si llevará a cabo el acto evaluativo, qué postura tomará y qué elementos lingüísticos utilizará para reflejar su evaluación interna (Alba-Juez y Thompson, 2014).

### 3.2 Términos utilizados para estudiar el lenguaje evaluativo

Aunque similares en su naturaleza, los sistemas terminológicos utilizados para describir el lenguaje evaluativo han sido de naturaleza diversa y dependiente del enfoque. Entre los principales, Martin y White hablan de *valoración (appraisal)*, Conrad y Biber utilizan el término *postura (stance)*, mientras que Thompson y Hunston, además del uso de *postura*, hablan de *evaluación*. En esta sección se definirán de manera breve los términos principales utilizados a lo largo de la literatura de las últimas tres décadas y se abordará la cuestión de la definición del término *evaluación* a lo largo de la literatura.

Uno de los estudios más veteranos fue el realizado por Chafe (1986), que parte de la escasez de trabajos sobre la evaluación en anglística si se compara con la bibliografía existente en las lenguas orientales. Según este trabajo, esto se debe a que se asumía inicialmente que la evidencialidad quedaba relegada a elementos (para)lingüísticos implícitos como el tono, la duración o el énfasis. Aquí se sentaron las bases para la

discusión de los primeros términos usados para reflejar conceptos relacionados con la expresión de la evaluación. Definió la *evidencialidad* (*evidentiality*) como las actitudes expresadas por el hablante hacia el conocimiento: su *fiabilidad*, el modo de conocer, y lo adecuado que considera el hablante sus enunciados. De esta manera, se alinean la enunciación y la realidad física de su contexto. El autor distingue tres aspectos marcadores de *evidencialidad*:

- fiabilidad intrínseca del conocimiento, eminentemente adverbial: usa partículas como “maybe”, “probably”, “surely”.
- modo de conocimiento (creencia, inducción y rumor): la primera con construcciones como “I think” o “I guess”, la segunda con “must” o “seem” y la tercera con modales como “should”, “could” o “presumably”.
- fuente de conocimiento (creencia, evidencia, informe verbal e hipótesis): expresada principalmente a partir de verbos relacionados con los sentidos como “see”, “hear” o “feel”.

Según esto, algunos autores (Vincze et al., 2008; Sauri y Pustejovsky, 2009) consideran que a pesar de ser un factor importante a la hora de, p. ej., detectar la fuente de opinión y también su fiabilidad, la evidencialidad no ha recibido mucha atención en el análisis de sentimiento, a pesar de aportar un conjunto de recursos muy útiles para determinar la fiabilidad de una opinión. Algunos aspectos de esto, sin embargo, se capturan en el sistema de *compromiso* dentro de la teoría de la valoración, que se revisará más adelante.

Dos años antes, Labov (1984) usó el término *intensidad* (*intensity*) como la expresión emocional de la orientación social hacia la proposición lingüística, en otras palabras, el compromiso del sujeto hacia dicha proposición. Se distingue así entre *amplificadores* y *enfáticos*: los primeros serían los elementos verbales que reflejan la fuerza y fiabilidad de una proposición en términos positivos (“absolutely”, “completely”, “extremely”); los segundos solo marcarían la presencia de certeza en una proposición (“for sure”, “really”...).

La *postura* (*stance*) es un término amplio que se usa de dos maneras distintas en la literatura. Biber y colegas publicaron una serie de artículos (Biber y Finegan 1988, 1989; Conrad y Biber 2000), que enfrentaban el lenguaje evaluativo a partir de un entonces novedoso enfoque computacional: la lingüística de corpus. En ellos describen la postura como la expresión de las actitudes, los sentimientos y el juicio del hablante, a la

vez que reflejan los dispositivos de *evidencialidad* que se utilizan para expresar el compromiso con el mensaje, así como el *afecto*, que además de actitudes personales, como emociones, sentimientos o humor, reflejan la evaluación positiva o negativa del hablante (Ochs y Schiefflin, 1989). Mediante el estudio del corpus London Lund, Biber y Finnegan (1989) crearon una clasificación de postura en seis categorías, acompañada de sus características lexicogramaticales y géneros más frecuentes: expresión enfática de afecto, postura sin rostro, evidencialidad interaccional, expresión expositora de duda, persuasión predictiva, persuasión oral controvertida. A partir de aquí, determinaron que las funciones pragmáticas de la postura se clasificaban en (a) *afecto* y (b) *evidencialidad* y a su vez, distinguían dos ejes semánticos: positivo / negativo (para el afecto) y certeza / duda (para la evidencialidad).

Una década después, Conrad y Biber (1999) simplificaron la clasificación y distinguieron tres tipos de postura: la *postura epistémica*, la *postura actitudinal* y la *postura estilística*:

- La *postura epistémica* (*epistemic stance*) indica el grado de certeza del hablante y se manifiesta de cinco maneras: (a) grado de certeza o duda sobre la proposición (“perhaps”, “certainly”, etc.), (b) comentar la realidad de la proposición (“actually”, “in fact”, etc.), (c) indicando vaguedad o imprecisión (“sort of”, “nearly”, etc.), (d) identificando la fuente original de información (“according to”) y (e) identificando la perspectiva del hablante (“apparently”, “from our perspective”).
- La *postura actitudinal* (*attitudinal stance*) expresa juicios y opiniones sobre lo que se está expresando (“fortunately”, “most surprising of all”, etc.).
- La *postura estilística* (*style stance*) se refiere al comentario del hablante sobre la presentación de la cláusula (“honestly”, “more simply put”, etc.).

Un cambio de paradigma fue el estudio de la *toma de postura* (*stance-taking*), el cual además de tener en cuenta la expresión del hablante, revisa la actividad comunicativa humana que hay detrás (Du Bois, 2007). El *triángulo de postura* fue diseñado para representar cómo se lleva a cabo la comunicación durante el diálogo en un acto que se compone de evaluación, posicionamiento y alineación. Es decir, alguien que toma una postura evalúa un objeto, posiciona un sujeto (generalmente a sí mismo) y se alinea con otros sujetos. Partiendo de este triángulo, es necesario tener en cuenta la relación de la

postura subjetiva / ideológica (“dónde tú y yo nos situamos”) con la postura intersubjetiva / interaccional (“cómo tú y yo nos relacionamos”).

El estudio de la *modalidad* (*modality*), que hace referencia no exclusiva al uso de los verbos modales, se ha utilizado tradicionalmente para determinar la opinión del hablante en relación con la probabilidad de que ocurran determinados eventos mencionados en el discurso. Halliday y Matthiessen (2003) o Perkins (1982) hacen una revisión exhaustiva del andamiaje lexicogramatical del sistema de modalidad. En cuanto a los términos utilizados por los enfoques sistémico-funcionales, Halliday et al. (2014) reflexionan sobre la necesidad de separar la modalidad del significado actitudinal y la exploran más profundamente estableciendo dos subcategorías: *modalización* (relacionada con la probabilidad y la usualidad) y *modulación* (relacionada con la obligación e inclinación). Coloca a ambos significados dentro de la categoría metafuncional interpersonal de su sistema, aunque esta categoría también incluya otras áreas como el *humor* (*mood*) del hablante y en su sistematización gramatical no establezca una subdivisión intermedia que englobe exclusiva y específicamente a estos dos conceptos.

La *actitud* es un hiperónimo usado por una red distribuida de estados mentales y emocionales, afecto, creencias, certeza, compromiso, disposiciones, emociones, ideología, postura, estado mental o cualquier otra condición interior relacionada, ya sea pasajera o permanente (Jullian, 2009). Como concepto, la actitud es clave en el estudio de la evaluación que hoy en día tiene mayor sostén teórico: la teoría de la valoración (Appraisal Theory) de Martin y White (2005), que encaja como una expansión del significado actitudinal de la metafunción interpersonal de la lingüística sistémica funcional establecida por Halliday. El eje de actitud se divide en tres subcategorías que se detallarán más adelante: *afecto* (*affect*), *juicio* (*judgement*) y *apreciación* (*appreciation*). La valoración no solo aporta una descripción muy rica de diferentes aspectos de la evaluación y el posicionamiento subjetivo, sino que es quizá una de las que mayor aceptación ha tenido, con una amplia literatura que la utiliza como marco de análisis, además de ser un referente ineludible para las implementaciones de análisis de sentimiento (Benamara et al., 2017).

Tras revisar los términos clave para aproximarse al lenguaje evaluativo (evidencialidad, modalidad, postura, actitud), es importante relacionarlo con el *lenguaje de las emociones*. Si bien tiene claras diferencias, este lenguaje está estrechamente relacionado con el lenguaje evaluativo ya que comparten temas de estudio (p. ej., a través de la relación del afecto con las emociones). Sobre esto, Bednarek (2008) diferencia entre

*habla de emoción (emotion talk)* y *habla emocional (emotional talk)*. El primero se constituye a partir de expresiones que denotan afecto/emoción, p. ej., “love”, “hate”, “joy”, “envy”, “sadness”, “disgust”, etc. e incluye también expresiones más complejas como “He had a broken heart”. El habla emocional hace referencia a elementos lingüísticos y extralingüísticos que sirven para expresar afecto o emoción (p. ej., interjecciones, intensificaciones). Relacionándolo con Martin y White (2005), mientras que el *afecto* (en particular, sus procesos mentales y estados) encaja mejor en el habla de emoción, los fenómenos de apreciación y juicio pueden tratarse potencialmente en el paradigma de habla emocional. Sobre esto, Bednarek opina que es necesario establecer límites claros entre afecto y emoción ya que, principalmente, la autora considera que la evaluación no refleja necesariamente las emociones o los sentimientos reales de los hablantes.

### 3.3 Definición de evaluación

Una vez definidos todos los conceptos y términos fundamentales, es necesario agruparlos deliberadamente. Para ello, Thompson y Hunston (2000) propusieron un término jerárquicamente superior: *evaluation*, de manera que este fuese lo suficientemente amplio como para abarcar globalmente la expresión de la actitud o la sensación del hablante hacia las entidades o proposiciones de las que se está hablando. Por ello, propusieron una definición ya clásica en los estudios del lenguaje evaluativo: “the broad cover term for the expression of the speaker or writer's attitude or stance towards, viewpoint on, or feelings about the entities or propositions that he or she is talking about” (p. 5).

Alba-Juez y Thompson (2014) actualizaron esta definición y la ampliaron teniendo en cuenta las innovaciones teóricas desarrolladas a lo largo de esta década. La definición revisada se enmarca más allá del texto y evidencia su participación en procesos dialógicos complejos:

a dynamical subsystem of language, permeating all linguistic levels and involving the expression of the speaker's or writer's attitude or stance towards, viewpoint on, or feelings about the entities or propositions that s/he is talking about, which entails relational work including the (possible and prototypically expected and subsequent) response of the hearer or (potential) audience. This relational work is generally related to the speaker's and/or the hearer's personal, group or cultural set of values (p. 13).

Ante la diversidad de teorías y aproximaciones al lenguaje evaluativo, Hunston (2011) marca seis puntos de consenso entre los distintos enfoques al lenguaje evaluativo. Son los siguientes:



1. La evaluación es subjetiva e intersubjetiva. En cuanto a lo primero: las oraciones evaluativas comparten opiniones personales y las expresiones evaluativas son respaldadas solo por su emisor. Sobre la intersubjetividad, podemos relacionarlo con el triángulo de Du Bois (2005), o con el alineamiento entre hablante y oyente de Martin y White (2005). Sobre esto, Hunston (2011) opina que las funciones de la evaluación sirven para construir y mantener relaciones humanas.
2. La evaluación infiere una ideología compartida entre hablante y oyente. Aunque en el primer punto hemos visto que la evaluación es personal y subjetiva, esta tiene lugar dentro de un mínimo marco ideológico y de valores compartido entre ambos actores y se lleva a cabo aunque los hablantes estén en desacuerdo cuando expresen sus opiniones.
3. El significado evaluativo está cubierto por un amplio rango de léxico y otros indicadores de significado evaluativo. Aunque las distintas escuelas difieran en qué se incluye en el término *evaluación* (p. ej., si la cuestión de la modalidad encaja o no dentro de este marco), existe consenso en que estudiar lo deseable e indeseable se aproxima a lo expresado por las distintas corrientes. Otro punto en común es que la evaluación depende profundamente de que los elementos estudiados estén contextualizados. También es un punto de consenso el que los distintos elementos que constituyen el lenguaje evaluativo son tan amplios que el mero hecho de listarlos supone una ardua tarea. Algunos elementos léxicos que contextualizados pueden indicar evaluación: sustantivos (p. ej., “success”), verbos, (p. ej., “fail”), adjetivos (p. ej., “excellent”), adverbios (p. ej., “sadly”) o expresiones multipalabra como “number one”, mientras que “number” y “one” no tienen implicaciones evaluativas). Asimismo, remarcar que existe consenso en que esto también pueda realizarse de manera implícita a través de distintos mecanismos.
4. La evaluación es contextual y acumulativa. Lo primero quiere decir que no es posible determinar si una palabra es evaluativa sin conocer su contexto. Por lo tanto, si una palabra es potencialmente positiva de manera aislada, no se puede determinar su polaridad con certeza sin otros elementos textuales (p. ej., negaciones, ironía, etc.). Por otro lado, los significados evaluativos tienden a unirse, de forma que podemos reconocer un segmento como evaluativo al detectar que distintos elementos evaluativos actúan en conjunto.



5. Toda evaluación implica un objetivo, un objeto y una fuente. Estos conceptos son centrales al mencionado triángulo de postura (*stance triangle*) de Du Bois (2007) y están incluidos en la quintupla de opinión de Liu (2015) al ser elementos clave para el análisis de sentimiento.
6. La tarea de identificar la evaluación se complica al discernir los factores evaluativos. La subjetividad y las “gafas ideológicas” que nos ponemos cuando nos sentamos a analizar un texto suelen ser muy influyentes a la hora de determinar si un texto es de carácter evaluativo o factual.

Gozdz-Rosztowski y Hunston (2016) revisan estos puntos al considerar mayor la complejidad de la evaluación al ser considerada un concepto problemático (por las dificultades de establecer cánones para identificarla) a la vez que importante (al construir una conexión entre ideología y discurso). Por ello proponen que la evaluación:

7. Representa el lenguaje que puede ser subjetivo o actitudinal y que puede ser distinguido del objetivo o factual.
8. Es multifuncional, ejecuta roles para interpretar relaciones entre los participantes en una interacción y para estructurar el discurso. Puede darse el caso de que el objetivo de la evaluación sea, por ejemplo, cambiar de tema, o expresar opinión.
9. La evaluación puede ser inscrita / explícita o evocada / implícita y para detectar en cuál de los ejes encaja deben explorarse las gramáticas locales y los registros.
10. Aunque la base de la evaluación sea la polaridad “bueno / malo” existen otros patrones de distinta naturaleza.
11. Tiene en consideración los conceptos de multiglosia e intertextualidad al poder realizarse mediante expresiones atribuidas a otros hablantes.
12. El lenguaje evaluativo es un campo clave para referenciar e interpretar las distintas ideologías (hegemónicas y de resistencia) ya que la evaluación inscrita o implícita presupone un conjunto de valores compartido.

### 3.4 Funciones de la evaluación

Ante tal cantidad de sistemas, visiones y enfoques existentes sobre el lenguaje evaluativo, cabe plantearse el motivo de la importancia de la evaluación como tema de estudio. Hunston (2011), señala que en cualesquiera de los actos evaluativos enunciados por el hablante recaen, al menos, una de estas tres funciones: (1) la expresión de la opinión del hablante de manera que refleje su sistema de valores y el de su comunidad, (2) la

construcción y el mantenimiento de relaciones hablante-oyente y (3) la organización del discurso.

Obviamente, la expresión de la opinión es la función más frecuente, ya que así se puede expresar el alineamiento o sentimiento sobre un tema en específico. La identificación de lo que piensa el hablante permea algo más que sus ideas: cada acto de evaluación tiene como telón de fondo un sistema de valores común y cada uno de los actos de evaluación contribuye a la construcción colectiva del mismo. Este sistema de valores es un engranaje más de un sistema ideológico indirectamente expuesto tras cada texto. Por esto, el análisis de los textos nos permite conocer más sobre la ideología y la comunidad de los hablantes. Aunque estas no son explícitas la observación de los textos evaluativos, nos permiten conocer (según la voz del texto) qué códigos o enunciados son esperados y deseables, qué es bueno, qué es malo, qué debe y qué no debe ocurrir.

La segunda función de la evaluación es la construcción y mantenimiento de las relaciones entre el hablante y el oyente. Este constructo social ha sido estudiado en relación con tres ejes de observación: *manipulación*, *cobertura* y *cortesía*. En cada una de estas áreas, el hablante explota los recursos de evaluación para construir un tipo particular de relación con el lector. La evaluación suele usarse con el objetivo de persuadir al oyente de manera que este comience a ver las cosas de una manera determinada. Por ejemplo, si el hablante considera determinado asunto como “un problema”, este forzosamente debe hacer uso del lenguaje para etiquetar lo enunciado como conflictivo. Es difícil que el lector no acepte esta percepción del problema como tal, en el caso de que dialógicamente se exprese un solo punto de vista. A partir de aquí, generalmente, el lector debería aumentar sus niveles de atención para extraer el punto de vista o ideología subyacente a la problemática. Ya que mucha información relevante a la evaluación se da por sentada y omitida, el hablante debe mostrar si su evaluación es el punto principal de un discurso a través de distintas estrategias de organización textual. Por otro lado, el lector no suele posicionarse y decidir si está de acuerdo con lo evaluado y, normalmente, asume la posición acríticamente (Hoey, 2000).

Francis (1986) indica que determinadas etiquetas del discurso tienen un rol muy importante cuando se lleva a cabo la labor de resumir, distribuir el texto y mantener los niveles de alineamiento entre lo que el oyente y el hablante consideran aceptable. Estas etiquetas de discurso pueden, de una manera simple, resumir (“these words”, “this question”, “to sum up”) o evaluar (“this claim”, “this nonsense”). Hay funciones de organización del discurso que juegan un factor clave en la influencia del oyente, p. ej.,

cuando se ofrece una información considerada hegemónica de manera factual, el oyente baja la guardia y no cuestiona su validez. Por ello, mientras menos obstruya la evaluación en una determinada cláusula, mayor éxito se suele tener cuando se pretende manipular al lector. Cuando se propone al oyente que lo que se expone es un punto en común para ambos interlocutores se está indicando qué es aceptable y qué no lo es de una manera indirecta. Además, la evaluación ejerce como garante de la veracidad en las declaraciones. Uno de los fenómenos interpersonales más frecuentes aquí es el *hedging*, mediante el que se relaciona el grado de certeza de un enunciado con un grado de conocimiento previo común. Por ejemplo, en la escritura académica, podemos hablar tanto de elementos discursivos (palabras como “insignificant”, “somewhat” o expresiones como “I’m not an expert but...”) como de terminología que aumenten tanto el nivel de precisión como la apariencia de veracidad de lo expuesto. El uso de estos elementos garantiza una versión final del enunciado negociada de una forma más precisa a través de modificaciones, que permite ajustar la alineación entre interlocutores. Por otro lado, existen algunos factores textuales que dificultan la alineación entre interlocutores. Un ejemplo de esto sería la adición de cláusulas adversativas al discurso que prevean puntos conflictivos que puedan añadir dificultad a la comprensión de una enunciación (Hunston, 2011).

Finalmente, además de alinear a los interlocutores, los segmentos evaluativos también se usan para organizar el discurso de forma que se pueda decir “esta parte es el principio del texto, esta parte indica cómo cuadra el argumento y esta sección es el final de nuestra interacción”. Hunston (2011), menciona que Sinclair ya discutía sobre la valuación y su papel organizador en el discurso. Desde el “Yes, that’s right” de un profesor después de una contribución en clase, por ejemplo, o la manera en la que un ponente destaca los puntos clave de su discurso como “the interesting part of this is that...”. Además de organizar la narración de un texto, cita a autores como Labov, que consideran que la evaluación es la cláusula principal de un enunciado, pues considera que un relato no puede existir sin algún tipo de evaluación. El narrador debe destacar el por qué estos hechos deben ser contados y por qué se les debe prestar atención. Aunque la evaluación ocurre durante toda la narración de una manera más o menos abierta, esta toma relevancia y es preeminente en distintos sectores de la narración: al principio, en el nudo, en el desenlace y justo al final antes del cierre de forma que suspenda la acción. Indica, para concluir, que la evaluación permite la realización de la narrativa, pues la convierte en un texto interactivo y hace que cumpla con su función comunicativa.

## 3.5 Identificación de segmentos evaluativos

### 3.5.1 Marcadores evaluativos

La evaluación se considera conceptualmente comparativa, subjetiva y cargada de valor, por lo que identificarla implica detectar señales de comparación, subjetividad y valor social. A lo largo de este capítulo se han recogido múltiples maneras de identificación de patrones evaluativos y, por ejemplo, a través de técnicas de corpus es posible identificar una gran cantidad de segmentos valorativos con relativa sencillez. No obstante, el lenguaje evaluativo presenta dificultades en su análisis porque no hay un conjunto claro de formas lingüísticas (ya sean gramaticales, léxicas o discursivas) que cubran el rango total de expresiones de evaluación. Esto no quiere decir que p. ej., una gran parte de adjetivos y un amplio conjunto de adverbios expresen significado evaluativo, o que sus patrones estén siempre asociados con las expresiones de evaluación (Hunston, 2011).

En el caso de la naturaleza comparativa de la evaluación, es necesario realizar una labor de contraste con las normas establecidas en una comunidad determinada. Así, el uso de una cláusula negativa se suele comparar con lo esperado o aceptable (o no) en una comunidad de hablantes.

Banks no longer have the power to approve loans.

Segundo, la naturaleza subjetiva de la evaluación es de carácter cognitiva y se realiza a partir de la *evaluación interna* de un sujeto en el que se muestra la reacción introspectiva del hablante ante un evento o situación.

Angela Merkel regrets the outcome of last night's vote in the Cypriot parliament...

Con respecto a la carga de valor, se sugiere que lo “bueno” o lo “malo” tiene relación con el cumplimiento de metas prefijadas. Una entidad con un valor “bueno” ayudaría a completar objetivos. Por ejemplo, un cartel en el que se diga “only available during desk hours” no permite conseguir la meta de sacar un libro de una biblioteca, por lo que se evalúa de manera negativa si nuestro objetivo es llevarlo a casa. Un cartel que indique “Honor bar, please help yourselves” permite a los usuarios de un transporte público obtener refrescos de un minibar dejando una moneda sin que nadie esté vigilando si lo que queremos es simplemente tomar un refresco, esto sería negativo si lo que queríamos es tener una agradable charla con un camarero mientras tomamos un refresco

(Hunston, 2011). Entre los elementos léxicos que tienen la evaluación como su función principal de significado se encuentran (1) adjetivos: “marvelous”, “terrible”, “amazing”, “obvious”... (2) adverbios: “happily”, “unfortunately”, “possibly”, “necessarily”... (3) sustantivos: “success”, “failure”, “tragedy”, “triumph”, “likelihood” y (4) verbos: “succeed”, “fail”, “win”, “lose”, entre otros.

Aunque existe un grado elevado de consenso a la hora de determinar el significado evaluativo de las palabras a partir de la introspección y la experiencia vivida dentro de una comunidad de hablantes, establecer criterios para distinguir los elementos evaluativos del resto es una tarea sumamente compleja. La autora propone comparar dos cláusulas por ejemplo “Jane’s a genius” y “Jane is a student”, a través de ella podemos establecer con facilidad que la primera es una oración de naturaleza evaluativa ya que “genius” es un sustantivo que se puede comparar por encima de la norma y ser un genio es algo valorado socialmente. La segunda palabra, “student”, se puede determinar como una palabra objetiva y descriptiva de la ocupación de una persona, mientras que en algunos contextos específicos puede tener una connotación positiva (alguien que forma parte de una élite intelectual) o negativa (desordenado, irresponsable...). Gracias a la lingüística de corpus podemos observar el discurso de un dominio específico o una comunidad determinada de forma que podamos ver el valor evaluativo real de una palabra a partir de su observación dentro de un contexto, ya que la intuición puede no bastar para conocer el aspecto evaluativo de la misma. Por otro lado, un elemento léxico proporciona más información que la evaluativa y como resultado su estatus de palabra evaluativa puede ser más que discutible. Este es el caso de los verbos y los sustantivos. En el caso del significado del ejemplo propuesto por la autora “being involved in something or taking part in an activity” teniendo a verbos inherentemente positivos como es el caso de “assist” o “help” desde el punto de vista de la otra persona implicada en la acción. Aun sea “assist in a robbery”, el ladrón, agente principal de ese robo, ve esa colaboración como positiva. En cuanto a ejemplos negativos, tendríamos a “interfere” o “meddle”, que son evaluativamente lo opuesto a “assist” y “help”.

Más allá del uso del léxico evaluativo, los distintos patrones gramaticales juegan un factor determinante en la expresión de la evaluación. Labov (1972; 1985) fue uno de los pioneros en la detección de patrones gramaticales. Este evidenció la existencia de elementos lingüísticos con valores evaluativos, como ya se ha mencionado, clasificando estos en cuatro:

- (a) Intensificadores: como gestos o fonología expresiva, cuantificadores, repeticiones y expresiones rituales.
- (b) Comparadores: como los negativos, futuros, modales, cuasi modales, preguntas, imperativos, cláusulas disyuntivas, superlativos y comparativos.
- (c) Correlativos: incluyendo progresivos, partículas adjuntas, y atributivos.
- (d) Explicativos: como las partículas de subordinación “while”, “though”, “since” o “because”.

Stubbs (1986) discute sobre la gramática evaluativa a partir de la gramática modal o el punto de vista. Toma como punto de partida una exploración concisa de la interacción entre léxico, gramática y conocimiento pragmático que permite evidenciar algunos de los elementos evaluativos presentes en la expresión del hablante:

- (a) Expresiones del origen de las proposiciones
- (b) Sintagmas que limitan el compromiso como “if I can”.
- (c) Formas de ser explícito a través de enunciados performativos o el uso consciente de la vaguedad durante el acto de habla.
- (d) La elección de aspectos simples o progresivos de los verbos cognitivos.
- (e) El significado modal de los private verbs, que expresan estados (p. ej., “believe”) o acciones intelectuales (p. ej., “discover”) que no son observables.
- (f) Conectores lógicos: (p. ej., “and”, “but”, “or”, “if”, “because”...)
- (g) El tiempo pasado indicando lejanía temporal (p. ej., “I did wonder if I might ask you a question”)
- (h) Referencias al futuro
- (i) *Tag questions*.

Biber y Finnegan (1989) realizaron un listado de los llamados marcadores de postura (*stance markers*), una serie de elementos léxicos seleccionados de tipos de palabra y otras categorías que se dividen en *afecto* (de carácter positivo o negativo) y *evidencialidad* (certeza y duda) (Tabla 1):

<i>Afecto</i>	
<i>Positivo</i>	<i>Negativo</i>
Adjetivos: “I feel FORTUNATE”, “It is AMAZING that...”	Adjetivos: “I am SHOCKED”, “it seems unnatural...”
Verbos: “I ENJOY...”, “it really PLEASES me that...”	Verbos: “I DREAD...”, “It EMBARRASSES me that...”
Adverbios: “happily”, “conveniently”, “luckily”...	Adverbios: “disturbingly”, “sadly”
<b>Evidencialidad</b>	
<b>Certeza</b>	<b>Duda</b>
Adjetivos: “impossible”, “obvious”, “true”...	Adjetivos: “alleged”, “dubious”, “uncertain”
Verbos: “to CONCLUDE...”, “This DEMONSTRATES that...”	Verbos: “I ASSUME that...”, “this INDICATES that...”
Adverbios: “assuredly”, “indeed”, “without doubt...”	Adverbios: “allegedly”, “perhaps”, “supposedly”
Enfáticos: “of course”, “really”, “so + ADJ”	Hedges: “at about”, “maybe”, “sort of”
Modales predictivos: “will”, “shall”	Modales de posibilidad: “might”, “could”
	Modales de necesidad: “ought”, “should”

**Tabla 1** Resumen de categorías de afecto y evidencialidad. Biber y Finnegan (1989)

### 3.5.2 Evaluación y textualidad

La evaluación está presente a lo largo de todas las secciones de un texto, y no en secciones puntuales, ya que este permite estructurar y aportar coherencia narrativa. Labov (1972) consideró que los patrones evaluativos proponen relaciones intertextuales entre las distintas cláusulas del texto y estableció una lista de características lingüísticas asociadas con la evaluación en la narrativa, destacando seis elementos: resumen, orientación, acción complicación, evaluación, resultado y coda. La evaluación sería, entonces, el medio principal del relato pues gracias a ella que el narrador puede expresar la razón de ser de una narrativa, esto es, el motivo de relatar algo (Hunston, 2011).

Hoey (1983) detectó la existencia de patrones evaluativos claramente identificables dentro del texto, diseñando un esquema textual que típicamente responde al modelo *Situación-Problema-Respuesta-Evaluación* (*Situation-Problem-Response-Evaluation*). El foco en estos estudios está en la forma en la que se reconoce la evaluación dentro del texto: mientras que el tipo de evaluación es irrelevante para Labov, para Hoey consiste simplemente en clasificar los segmentos de evaluación como positivos o negativos (Thompson, 2014). En la misma línea, Lemke (1998) considera que la expresión de la evaluación que se realiza en textos extensos y cohesivos no está construida solo a partir de cláusulas individuales aisladas. Los fenómenos de lenguaje evaluativo extienden sus recursos semánticos a nivel textual debido a la tendencia de este tipo de

lenguaje hacia realizaciones prosódicas, es decir, tienden a distribuirse no solo a través de la cláusula, sino entre los límites de las cláusulas y las oraciones. Las partículas evaluativas, se vinculan a través del texto para formar una *cadena* en virtud de sus relaciones semánticas entre sí de manera que pueden propagarse de una estructura a otra p. ej., a partir de la correferencia o la sinonimia.

Hasta ahora se han observado dos tipos generales de evaluación: modalidad y afecto. Como era de esperar, la gramática de la modalidad tiene mayor facilidad de detección que la del afecto. La mayoría de las expresiones de modalidad se realizan a través de las unidades modales (destacando los verbos), mientras que los patrones gramaticales del afecto ocurren de manera menos evidente en el plano léxico, generalmente mediante sustantivos y adjetivos, los cuales suelen tener características más difusas. Además, la evaluación *modal* usa generalmente proposiciones; en cambio, la evaluación *afectiva* suele realizarse en relación con entidades. Thompson y Hunston (2001) proponen tres clasificaciones de las características lingüísticas que sirven de guía en la detección de segmentos evaluativos en el texto:

1. La evaluación es considerada como *comparador* del objeto que debe ser evaluado y que siempre tiene una vara de medir como referencia (los comparadores). Estos incluyen: “just”, “only”, “at least”, expresiones de negatividad de tipo morfológico, como los morfemas derivativos “un-”, así como otros afijos, o unidades gramaticales como “not”, “never”, “hardly”... y unidades léxicas como “fail”, o “lack”.
2. La evaluación es subjetiva: los marcadores de subjetividad conforman un grupo amplio como modales y otros marcadores de certeza. Estos comprenderían adjetivos, adverbios y conjunciones y estructuras de cláusula marcada, p. ej., los patrones gramaticales de carácter impersonal que comienzan con “it” (“it is well considered that...”).
3. La evaluación contiene valor: los marcadores de valor son elementos léxicos que se usan para evaluar como los indicadores de las metas y su cumplimiento (“what is good”, “what achieves our goals”).

### 3.5.3 Identificación de patrones en el género periodístico

Una de las razones de que existan múltiples taxonomías en el campo de la evaluación se debe a que “el acto de evaluar algo” se puede hacer desde múltiples parámetros. El género



de las noticias tiene unas prioridades bien marcadas cuando se quiere expresar la deseabilidad de una evaluación. Al fin y al cabo, las evaluaciones bueno/malo son elementos amplios que dependen del sistema de valores y preferencias que haya detrás del texto. Por ejemplo, el hecho de que en algunas situaciones: p. ej., que una experiencia estética sea más importante que la falta del confort (para un alpinista, por ejemplo, la recompensa sería el paisaje, a pesar del enorme esfuerzo o la climatología). El género de un texto sería el encargado de priorizar las evaluaciones de certeza, las evaluaciones del tipo bueno/malo o las evaluaciones del tipo de importancia/expectativa. Las segundas irían más enfocadas a discusiones más *mundanas* (mundo real) mientras que las últimas son de carácter abstracto e irían más *orientadas a texto* (académico).

Bednarek (2006), revisado en Bednarek y Caple (2012), propuso un marco específico para este género que permitiría detectar segmentos evaluativos dentro del texto a través de un conjunto de *parámetros evaluativos*. Estos parámetros crean un marco conceptual para estudiar de manera concisa los marcadores evaluativos de un discurso altamente evaluativo como el periodístico. Este se divide en diez ejes:

- (a) Importancia (*un/importance*): en él se responde a la pregunta de “¿cuán importante es esto?”. Por ejemplo:

This is a MAJOR event.

- (b) Comprensibilidad (*in/comprehensibility*): en él, los actores de la noticia evalúan si la calidad de la información recibida es buena, o por el contrario existe vaguedad o si la misma está más o menos implícita. Por ejemplo:

It's impossible to PIN DOWN the reasons for this success.

- (c) Posibilidad, capacidad (*im/possibility, in/ability*): relacionada con la modalidad deóntica y se preocupa de qué es (o no es) posible cuando se valora una situación o un evento. Por ejemplo:

They've not BEEN ABLE TO make a go of it financially.

- (d) Necesidad (*un/necessity*): relacionada con la modalidad deóntica y dinámica, es decir, construcciones, verbos modales, verbos, adverbios y adjetivos que expresan qué es (o no es) necesario de manera que se explica la consideración sobre lo que se debe o no se debe hacer. Por ejemplo:

The dispute between Belarus and Russia SHOULD act as a wake up call to the EU commission.

- (e) Emotividad (*emotivity*): hace referencia a la valoración por la fuente de opinión de una situación como positiva o negativa a partir de valores estéticos o morales de deseabilidad y tiene un amplio trasfondo ideológico. Esta característica es la más relacionada con este trabajo y va de la mano con el parámetro de Actitud de la teoría de la valoración. Este eje es clave en la noticia periodística, y central en nuestra tesis, pues a partir de él las cabeceras editoriales construyen las historias en línea con su público objetivo de manera que sugieren a sus lectores qué es deseable según una serie de valores.
- (f) Autenticidad (*un/genuiness* o *in/autenticity*): aquí el reportero valora la autenticidad de los elementos de la noticia. Por ejemplo:

But there was a question of how much was GENUINE weakness and how much was due to these factors”.

- (g) Parámetro de fiabilidad (*reliability*): está directamente relacionado con la modalidad epistémica (modales como “will” o “could”) y hace referencia a determinados elementos que evalúan la probabilidad de que algo ocurra. Este es el caso de:

The bank is NOT LIKELY to be the only one affected by the credit crunch.

- (h) Expectativa (*un/expectedness*): como su nombre indica expresa cuán esperado era un determinado evento incluyendo conjunciones (“but”, “yet”), adverbios (“although”, “even”, etc), etc. de manera que se modula el texto mostrando lo que no cumple con lo deseado. En términos generales es similar a nivel teórico al conjunto de negación de los *contextual valence shifters* de Polanyi y Zaenen (2006).
- (i) Evidencialidad (*evidentiality*): hace referencia a las expresiones que indican las bases en las que se fundamentan las enunciaciones y la información. Como se verá con más detenimiento en el apartado de lenguaje periodístico del Capítulo 4, existen una serie de construcciones (p. ej., el estilo indirecto) que indican la fuente desde donde la fuente de opinión obtiene la información de manera que se dé mayor credibilidad a lo enunciado.
- (j) Estado mental (*mental state*): en este marco se evalúan las opiniones concernientes al sentir de la gente con respecto a una determinada situación: emoción, volición, creencias, expectativas, etc. de manera que se aumente la

noticiabilidad al compartir dichos sentimientos. Hablamos aquí de expresiones como: “growing frustration”, “refusing”, “doubts that”, etc.

Este marco de estudio es muy conciso a la hora de estudiar los distintos factores evaluativos de las noticias y sirve como buen punto de partida para conocer qué elementos evaluativos influyen en la redacción de estas. Dicha clasificación no distingue entre factores léxicos y gramaticales al ampliar para este fin la línea de estudios de los patrones evaluativos de Hunston (p. ej., Thomson y Hunston, 2000) ni tampoco entra a discutir en profundidad los difusos límites entre evaluación y opinión que, paralelamente, otros autores (principalmente Liu, 2015) han discutido con vistas a realizar estudios de análisis de sentimiento. Sin embargo, sí pone sobre la mesa la dificultad que tiene discernir de forma aséptica entre enunciaciones evaluativas y factuales, y hace constar, coincidiendo con ejemplos sobre la economía, que existen enunciaciones que aparentemente no son evaluativas y que sin embargo contienen claramente una polaridad. Este es el caso de enunciaciones como “petrol prices were raised”, formalmente factual, pero que sin embargo aporta un sentimiento que dependerá de quién sea el receptor de esta. El ejemplo se considerará positivo en el contexto de una economía petrolera como es la del Ecuador, pero se considerará negativo en una economía importadora neta de recursos energéticos como es la española.

Teniendo en cuenta todos los marcos de evaluación, podemos concluir que, además de los marcadores evaluativos lexicogramaticales clásicos relacionados con la expresión de la modalidad y el afecto, es necesario tener en cuenta a la evaluación como un elemento fundamental en la organización un texto. Además, es necesario relacionar estos elementos con los marcadores que comparan la realidad con elementos subjetivos de deseabilidad que son relativos al individuo o a la construcción de un marco evaluativo global de una comunidad de hablantes.

### **3.6 Gramática local: forma y función de la evaluación**

La gramática local (*local grammar*) hace referencia a descripciones de áreas particulares del lenguaje en uso como pueden ser, por ejemplo, las características de los titulares de periódico en inglés, o la estructura que tienen las definiciones en un diccionario. Uno de los trabajos fundacionales sobre la influencia de la fraseología en el lenguaje fue el realizado por Harold Palmer (1933). Si bien este es más recordado en el campo de la didáctica de la lengua, fue el primero en insistir sobre los vínculos semánticos que van

más allá de los límites de la cláusula. Ya en la década de 1950, el británico J. R. Firth (1957) le dio el sentido moderno al término *colocación* y sentó las bases para el estudio de las frecuencias de coocurrencia, lo que permitió un estudio más empírico de la lingüística a partir de la búsqueda de patrones lexicogramaticales.

El desarrollo de la lingüística de corpus, especialmente a partir de la década de 1980, abrió un enorme abanico de posibilidades a partir de las nuevas técnicas computacionales, que permitían realizar consultas de lenguaje real y observaciones estadísticas de una manera más rápida y sencilla. El estudio de amplios corpus textuales permitió aportar evidencias de carácter cuantitativo y detectar configuraciones estructurales (colocaciones, coligaciones, etc.) de un carácter más amplio que el obtenido mediante introspección (Moreno-Ortiz y Pérez-Hernández, 2014).

El trabajo de Sinclair y el equipo Cobuild (1991), uno de los principales promotores de la lingüística de corpus, trató dos enfoques gramaticales opuestos para describir la formación del significado en la lengua. Muchas de las descripciones gramaticales tradicionales tratan los aspectos léxicos y semánticos de una manera totalmente separada de la estructura sintáctica, quizá caricaturizadas en el clásico ejemplo chomskiano “Colorless green ideas sleep furiously” (Chomsky, 1957). Esta oración carente de significado semántico podría, discutiblemente, llegar a ser considerada gramatical y aceptable bajo una perspectiva que considere la sintaxis como una estructura de tipo *slot-and-filler* y solo tenga en cuenta los tipos de palabra como ejes organizadores del lenguaje. Los enfoques basados en corpus, a través de la significancia estadística, evidencian la no aleatoriedad de la coselección léxica, de forma que hay unidades léxicas que ocurren preferentemente juntas. Esto se ha llamado principio idiomático (*idiom principle*). Sobre este fenómeno, los autores (Partington, 1998; Hoey, 2005) han considerado los distintos fenómenos de la colocación como un fenómeno psicolingüístico que permean todo el lenguaje y del que hemos tenido evidencias a partir del estudio estadístico de corpus. El principio idiomático considera que en el lexicón de los hablantes existe gran serie de oraciones semiconstruidas junto a las palabras, y que estas se unen a través de distintos niveles de fijación en una línea ascendente que veremos a continuación:

- *Colocación*: sería el nivel más estricto de coocurrencia, de manera que los nodos se colocan fijamente de determinada. De esta manera hablamos de “black and white”, “salt and pepper” o “quantitative easing”.
- *Coligación*: asociación estadística fuerte de una palabra y una categoría gramatical particular. Mientras que las colocaciones son a nivel léxico, las

coligaciones lo son a nivel sintáctico. Asimismo, también se da el caso de que este tipo de palabra evite la coocurrencia con una función gramatical concreta. Un ejemplo de coligación sería, la clase de verbos de percepción, como “hear”, “notice”, “see” coligan con la secuencia del objeto + el infinitivo o verbo con forma -ing. Por ejemplo:

We heard people talk/talking.

- *Preferencia semántica*: cuando una determinada palabra se coloca con un conjunto léxico, parte de la oración o lemas semánticamente relacionados. Por ejemplo, cuando la palabra “large” normalmente coocurre con unidades de cantidad y tamaño.
- *Prosodia semántica* (prosodia del discurso o asociación semántica): expresa la actitud del hablante. En relación con esto, la gramática de patrones (Hunston, 2011; Hunston y Francis, 1999) se ha desarrollado a partir del estudio de distintos patrones que relacionan forma y significado en el lenguaje, enfocándose en el estudio de distintos patrones evaluativos. Esta considera que la relación entre forma y significado no es arbitraria, de manera que los distintos significados de una palabra polisémica se distinguen por los distintos patrones gramaticales con los que con frecuencia coocurre dicha palabra. Además, las palabras próximas semánticamente revelan patrones gramaticales (a las que Hoey llama *conjuntos semánticos*) más allá de las colocaciones, para llegar a formar unidades fraseológicas a partir de una serie de elementos. Por ejemplo, indica que “consequence” atrae adjetivos que proporcionan una asociación de carácter lógico (causa-efecto). Estos se dividen en subconjuntos: necesidad, etapas del proceso lógico, qué tan natural es, cuán directo es y cuán esperado es dicho proceso.

Esta jerarquía de la idiomatidad no solo explica la organización del *cognos* léxico, sino que además nos desvela cómo ocurren los cambios lingüísticos a través de giros y grietas en las colocaciones iniciales a partir de nuevos contextos de utilización que pueden motivarse por cambios sociales o históricos. Coincidiendo con Hunston (2014, p. 12), las colocaciones trascienden su papel de meras fijaciones entre varias unidades: “collocation, concordance lines and wider context explicate the meaning, cultural or ideological significance and function of the words and phrases so identified”. Como es la intención de este trabajo, se puede evidenciar la evolución del sentimiento de

una unidad léxica a través de la observación estadística de los elementos que componen sus asociaciones semánticas. Por ejemplo, O'Halloran y Coffin (2004) evidencian que una sucesión de cotextos negativos para una palabra en un medio aparentemente neutro suele suponer, previsiblemente, que se realice una lectura negativa de la misma (Hoey, 2005).

Existe una larga tradición en el estudio del lenguaje evaluativo y su relación con los distintos patrones lexicogramaticales, p. ej., la amplia gama de estudios sobre la postura realizados por Biber et al. (p. ej., 1999). Más recientemente, otros estudios de corpus (como es el caso de la amplia labor realizada por Hunston, 2011) se enfocaron en el estudio de los distintos patrones gramaticales que tiene el lenguaje evaluativo. Moreno-Ortiz y Pérez-Hernández (2014) revisan algunos estudios relacionados como el caso de Sinclair (1991), que observó que “set in” se utiliza frecuentemente para describir situaciones desagradables, o el caso de Stubbs (1995; 2001), quien observó las colocaciones de conjuntos de palabras con una connotación positiva o negativa y los componentes evaluativos. El autor, a través de su trabajo con corpus determinó que la unidad “job” coocurre tanto con unidades positivas como negativas, mientras que observó que “career” suele relacionarse con palabras de índole positiva.

El estudio de la *prosodia* o la *asociación semántica* es central para el estudio cualitativo de este trabajo. Los posfirthianos Sinclair (1987) y Louw (1993) fueron los primeros en utilizar el término, que hace referencia a expresiones asociadas mayoritariamente con significados relacionados con el sentimiento: positivas, negativas o neutras. Louw (1993) definió la prosodia semántica como “a consistent aura of meaning with which a form is imbued by its collocates” (p. 157) de forma que se mantiene la *armonía* evaluativa y ayuda a mantener la comprensibilidad del oyente. Por ejemplo: los verbos “cause” o “happen” suelen relacionarse con léxico con sentimiento negativo. Siguiendo el trabajo de Alcaraz Mármol y Soto Almela (2016), la prosodia semántica se concibe como un significado extendido en el que un elemento que parece denotativamente neutro puede tener un efecto positivo o negativo a partir de las asociaciones semánticas de las palabras que frecuentemente aparecen con ella. Por lo tanto, el concepto va más allá del mero significado denotativo de un diccionario y es un significado adicional que es capaz de indicar asociaciones emocionales e incluso sobre el significado pragmático o actitudinal.

Se entiende la prosodia semántica como una propiedad del léxico que no es estática, sino variable en el tiempo. Bublitz (1996), o más recientemente, Morley y

Partington (2009) consideran que este fenómeno puede afectar a palabras cercanas pues está basado en un conjunto de colocaciones que pueden tener un matiz semántico positivo o negativo que se va adquiriendo a lo largo del tiempo y que puede ser observable mediante una metodología basada en corpus. Stewart (2010, citado por Alcaraz Mármol et al., 2016) estudió este fenómeno dentro de un marco diacrónico, donde la prosodia semántica evoluciona contagiándose del significado de sus colocaciones a través del tiempo: “prosodic meaning may be attached, be taken on, be taken, be acquired, be accorded, be carried or be held; it may imbue, colour, taint, rub off, be reflected or spill over; further, prosodic meaning is an aura, a halo, a shade or a hue, and is the result of infection or contagion” (p. 43). En esta línea, McEnery y Hardie (2012) señalan el hecho de que una unidad tenga una prosodia positiva o negativa está relacionado con sus coocurrencias: “words or phrases are said to have a negative or positive semantic prosody if they typically co-occur with units that have a negative or positive meaning” (p. 135).

El estudio de patrones de este tipo es muy interesante cuando se implementan soluciones técnicas orientadas al estudio del análisis de sentimiento. Esto concierne a la creación de lexicones con el sentimiento palabras y en muchos casos sus colocaciones: las expresiones multipalabra<sup>20</sup>. Las críticas a este modelo giran principalmente, según Moreno-Ortiz y Pérez-Hernández (2014), al desarrollo ad-hoc del mismo ya que no ofrece explicaciones teóricas sobre la base cognitiva que lo sustenta. Otro de los problemas es el asunto del dominio, ya que las prosodias se desarrollan en un contexto y un tiempo concreto, pues los hablantes necesitan un intervalo temporal y una costumbre de uso para darle aplicaciones de carácter más creativo, como por ejemplo el uso de la ironía.

Por otro lado, han sido varios los intentos realizados para complementar la gramática de patrones (Hunston y Francis, 2000) con otros marcos teóricos, como es el caso de la teoría de la valoración de Martin o el proyecto semántico *FrameNet* de Fillmore et al. (2002). Hunston (2011) estudia, critica y complementa de manera eminentemente práctica el modelo teórico de Martin a partir de un análisis basado en el corpus de distintos

---

<sup>20</sup> De acuerdo con Sag et al. (2002): “Multiword expressions (MWEs) are lexical items that: (a) can be decomposed into multiple lexemes; and (b) display lexical, syntactic, semantic, pragmatic and/or statistical idiomaticity”. Mientras, en el procesamiento del lenguaje natural, los *n-grams* o *n-gramas* hacen referencia a una secuencia contigua de *n* elementos (bigramas, trigramas...) de una muestra determinada de texto (generalmente palabras, aunque también fonemas, sílabas, letras...), sin la necesidad de que exista una relación idiomática. Ambos conceptos son clave en el proceso de extracción automática terminológica.



adjetivos en inglés y sus distintos patrones léxicos. Este trabajo, igualmente, propone técnicas relacionadas con la fraseología que son útiles para conocer si determinadas construcciones se relacionan con palabras con un sentimiento positivo o negativo, además de añadir factores que determinan distintos grados de intensidad, como es el caso de construcciones como “on the verge of + NP”.

Para concluir con la revisión, uno de los trabajos teóricos más completos es el realizado por Bednarek (2008), quien realiza un conciso estudio sobre la relación completa de los patrones de términos emocionales, que combina los patrones de Hunston con doce marcos emoción de *FrameNet* para crear lo que denomina *emotion profiles* (perfiles de emoción). En esta misma línea encontramos el sistema *SentiFrameNet* (Ruppenhofer y Rehbein, 2012). Este categoriza las oraciones a partir de una recopilación de carácter sintáctico-semántico de expresiones evaluativas a partir de la clasificación de *frames* semánticos en categorías como polaridad, intensidad y afectación. Los modelos basados en *FrameNet* capturan tanto construcciones como fraseología de carácter evaluativo, de manera que se puede observar la representación profunda de la expresión de la evaluación inscrita en la estructura verbal (Benamara, et al., 2017).

### 3.7 La teoría de la valoración

Los profesores J. R. Martin y P. White desarrollaron la teoría de la valoración (*Appraisal Theory*) en Sydney durante los años 90 y la primera década del siglo XXI. El grueso del trasfondo teórico se expone en su obra *The language of evaluation: appraisal in English* (Martin y White, 2005), que recopila y amplía lo expuesto en trabajos anteriores (p. ej., Iedema, Feez, y White, 1993). Este enfoque profundiza y amplía las miras tradicionales del estudio del lenguaje evaluativo que, como hemos revisado, se había venido encargando principalmente de analizar las distintas facetas de la modalidad, en particular bajo el doble paraguas de la *modalidad* y la *evidencialidad*. De hecho, se trata de un desarrollo complejo de la lingüística sistémico-funcional de Halliday (2014), para el análisis de los usos evaluativos del lenguaje a partir de una visión dialógica del discurso. A través de tres ejes distintos (*actitud*, *compromiso* y *gradación*) se expande el engranaje sistémico-funcional de la metafunción interpersonal en los que se mueve la postura intersubjetiva del hablante. Este marco de estudio del lenguaje evaluativo está escrito desde y para la lengua inglesa y aclaran que no se debe asumir que las categorías descritas existan de tal manera en otras lenguas, aunque se han llevado a cabo trabajos



comparativos (Thomson y White, 2008) para observar cómo funciona la teoría en el ámbito de otras lenguas (White, 2015).<sup>21</sup>

La gran extensión y concisión de la teoría de la valoración podría servir como soporte teórico lingüístico para los estudios de análisis de sentimiento, por lo que merece la pena su estudio y conocimiento. Esta teoría aporta un análisis preciso y detallado de los factores por los que los textos aportan evaluaciones positivas o negativas; la intensidad y lo explícitas o implícitas que sean determinadas expresiones y, finalmente, si lo enunciado se alinea con enunciaciones anteriores e interlocutores presentes o potenciales. Debido a su naturaleza, es complicado implementar soluciones automatizadas que abarquen la amplitud y nivel de detalle (Liu, 2015) de manera que se cubran todos sus sistemas a través del procesamiento del lenguaje natural. Sin embargo, sus sistemas están presentes en los sistemas de análisis de sentimiento en mayor o menor medida, principalmente la detección del léxico y patrones evaluativos, así como los aspectos contextuales que gradúan el nivel de intensidad de las enunciaciones.

La valoración es un término de alto rango que contiene la totalidad de los usos evaluativos del lenguaje. La novedad, en comparación con la revisión bibliográfica anterior, recae en la construcción discursiva de la actitud y de la postura intersubjetiva. Se entiende que el objetivo del lenguaje evaluativo no es otro que establecer vínculos con otras personas a través de un sistema de elecciones lexicogramaticales y discursivas. Además de analizar los medios mediante los cuales los hablantes codifican sus actitudes de manera explícita, también se observan los recursos lingüísticos que activan de manera indirecta las posturas evaluativas. Por ello, los hablantes no solo expresan opiniones propias, porque para expresarlas se tendrá en cuenta los puntos de vista de otros interlocutores ya sean reales o potenciales (Kaplan, 2007). El resultado de este proceso no es otro que el posicionamiento de los receptores. El otro aspecto que opera en la evaluación es el tratamiento de la intensidad en el discurso y el nivel de precisión. Este permite describir la manera en la que los hablantes aumentan o disminuyen la fuerza de

---

<sup>21</sup> La adaptación al castellano de los términos usados en este documento está basada en la traducción realizada por Ghio de White (2003), disponible en <http://www.grammatics.com/appraisal>

sus aseveraciones y la forma que tienen los hablantes para afinar o difuminar las categorizaciones semánticas con las que operan.

### 3.7.1 La teoría de la valoración situada en el modelo sistémico funcional

Para comprender los distintos parámetros en los que se estudia el lenguaje evaluativo se deben definir algunos conceptos clave para la lingüística sistémico-funcional. En esta línea, el lenguaje se considera un amplio sistema de opciones disponibles en las que el hablante tiene que realizar una elección consciente, por lo que cada sintagma emitido es resultado de la selección entre un conjunto de opciones estructurales. De allí parte el adjetivo *sistémico* de su nombre: el objetivo de la teoría es reconstruir el sistema semiótico de opciones disponibles a partir de los enunciados (es decir, la estructura) emitidos por los hablantes de una lengua (Eggins 2004, Halliday y Matthiesen, 2014; Martin 1992).

La metafunción es el elemento principal aquí, pues es la función en su máximo nivel de abstracción que se pone de manifiesto en la organización del lenguaje (Eggins, 2004). Las metafunciones se comprenden como un recurso para cartografiar los significados ideacionales, interpersonales y textuales. En resumidas cuentas:

- Los recursos *ideacionales* se dedican a interpretar la experiencia: qué ocurre, qué le hace tal a cuál, cuándo, dónde, cómo y por qué.
- Los recursos *interpersonales* negocian las relaciones sociales: cómo interactuamos unos con otros. Esto incluye los sentimientos que el hablante quiere compartir.
- Los recursos *textuales* se preocupan del flujo de información: las maneras en las que los significados interpersonales e ideacionales se distribuyen en ondas de semiosis, incluyendo interconexiones entre ondas de información, y entre lenguajes y modalidades de asistencia.

Con vistas a posicionar el espacio que la teoría de la valoración ocupa en la lingüística sistémico-funcional, es necesario conocer los distintos niveles de abstracción en los que el lenguaje se realiza. Cuando se dice “realización” se debe entender como el proceso por el que el lenguaje “toma cuerpo”, considerándose al lenguaje como un sistema semiótico estratificado en tres capas de codificación repartidas en una escala de diferentes niveles de abstracción: (a) fonología/grafología, (b) lexicogramática y (c) semántica del discurso.

La teoría de la valoración se desarrolla en el tercer plano mencionado y es responsable de observar las unidades supraclausales encargadas de reflejar los distintos aspectos contextuales de la organización del discurso. Martin y White (2005) tienen en cuenta tres razones:

- (a) La expresión de la *actitud* empapa transversalmente el discurso trascendiendo los límites de la oración, especialmente cuando se producen fenómenos como la amplificación.
- (b) Una actitud determinada puede llevarse a cabo no solo a partir de una única categoría, sino de una serie de categorías gramaticales relacionadas: Por ejemplo, el caso, “interesting” / “interested” / ”interestingly” (epíteto / proceso / adjunto):
  - a. An INTERESTING proposal was made.
  - b. The proposal was INTERESTING.
  - c. INTERESTINGLY, the proposal was made.
- (c) Es necesario concebir a la metáfora gramatical como el significado interpretado de manera distinta por medio de una construcción gramatical. Dicha metáfora añade tensión extra a partir del uso estilístico de las palabras, alcanzando a atribuir y gradar las distintas opiniones. Por ejemplo:
  - a. PERHAPS his playing style might be different.
  - b. MAYBE this playing style would be different.

En términos gramaticales ilustrativos, una descripción se puede realizar a partir de la modalidad, principalmente a través de adverbios o verbos modales. Las distintas alternativas de enunciación posibilitan un *juego* lexicogramático de manera que las elecciones del hablante muestran distintos grados de certeza:

- a. I SUSPECT the proposal was not interesting enough.
- b. I THINK the proposal was not interesting enough.
- c. I KNOW the proposal was not interesting enough.

Por estos motivos, la tradición sistémico-funcional considera de poca ayuda la tradicional arbitrariedad entre contenido y forma de expresión como dictaba la Escuela de Ginebra (*significado* y *significante*). En palabras de Halliday (1975) “the internal

organization of language is not arbitrary but embodies a positive reflection of the functions that language has evolved to serve in the life of social man” (p. 26). Teniendo esto en cuenta, además, la lexicogramática sería más un recurso de construcción de significado que un conjunto universal de formas. Como conclusión, la lingüística sistémica pone en primer plano la noción de que la lengua es un sistema de elecciones por el que un hablante selecciona una forma determinada de lenguaje en detrimento del conjunto de sus alternativas.

### 3.7.2 Sistema general de la teoría de la valoración

El sistema general de la teoría de la valoración está dividido en tres subtipos principales, que son interactivos entre sí: *actitud*, *compromiso* y *gradación*. La *actitud* es el dominio relacionado con los sentimientos e incluye las reacciones emocionales, juicios de comportamiento y la evaluación de las cosas. Por otro lado, el *compromiso* se encarga de la relación entre las actitudes y el juego de voces interpersonal que hay alrededor de las opiniones expresadas en el discurso. Finalmente, la *gradación* se atiene a los fenómenos de escala a partir de los cuales las sensaciones se atenúan o amplifican y las categorías se precisan o se difuminan (Martin y White, 2005).

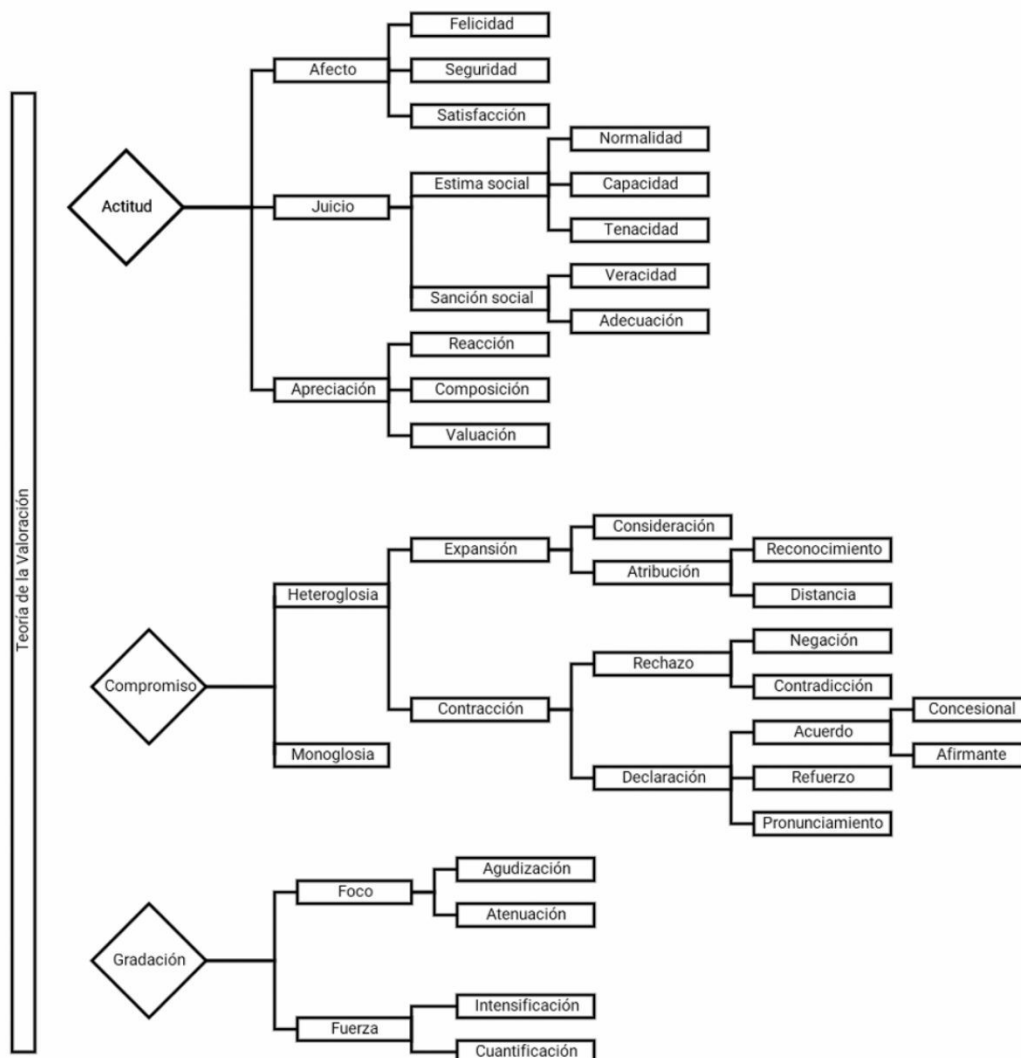
El sistema de *actitud* se divide en tres regiones del sentimiento: *afecto*, *juicio* y *apreciación*.

- El *afecto* se encarga de los recursos que se utilizan en el intercambio verbal para interpretar las reacciones emocionales.
- El *juicio* se dedica a estudiar los recursos utilizados para evaluar distintos comportamientos de acuerdo con principios morales y normativos.
- La *apreciación* observa los recursos para interpretar el valor de las cosas, incluidos los fenómenos naturales y la semiosis.

El *compromiso* está relacionado con la perspectiva de Bajtin (1934) y analiza el significado interpersonal de las enunciaciones desde una perspectiva social y dialógica. Se preocupa de las maneras en las que recursos como la proyección, la modalidad, la polaridad o la concesión se utilizan para posicionar al hablante con respecto a la posición de valor que se expresa como anticipo a las potenciales contraargumentaciones de esa posición de valor.

La *gradación* se encarga de analizar la gradabilidad. Ajusta el grado de una evaluación, la fuerza de un sentimiento. Esto se materializa en intensificación a partir de

distintas estrategias: p. ej., morfología comparativa y superlativa, repetición, además de varias características grafológicas (p. ej., subrayados, negritas...) y fonológicas (p. ej., subir el volumen de voz o dar gravedad al tono).



**Figura 1** Sistema general de la teoría de la valoración

### 3.7.3 Actitud

La *actitud* (*Attitude*) es el núcleo principal dentro del sistema de valoración y está estrechamente relacionado con el estudio del sentimiento, mostrando distintos sistemas de positividad o negatividad. En efecto, expresa la mayor carga de significado subjetivo y está dispuesta para ser modulada por los otros dos sistemas de valoración: la *graduación* y el *compromiso*. El sistema de actitud se encuentra tradicionalmente vinculado a tres

regiones semánticas claramente diferenciadas. Estas son la emoción, la ética y la estética (Martin y White, 2005).

- El *afecto* (*affect*) se encarga de registrar la primera región e implica sentimientos positivos o negativos como estar feliz, triste, aburrido o entretenido.
- El *juicio* (*judgement*) se encarga de las actitudes que un hablante tiene hacia los comportamientos humanos, los cuales admiramos o criticamos, alabamos o condenamos.
- La *apreciación* (*appreciation*) incluye evaluaciones de tipo semiótico o de fenómenos naturales, de acuerdo (o no) con las formas en las que se valoren en un campo determinado.

Según los autores, los significados actitudinales tienen a extenderse y “dar color” a la fase del discurso en las que los hablantes se posicionan en el discurso y proporcionan afecto (emociones, reaccionar a comportamientos), juicio (ética, evaluar comportamientos) o apreciación (estética, evaluar texto/proceso, fenómenos).

### **Afecto**

La actitud, como sistema semántico de discurso, se realiza mediante una serie de estructuras gramaticales relacionadas con la respuesta emocional y la disposición. Los elementos de afecto aportan una forma fácilmente distinguible de reflejar la adopción de una postura frente a determinado fenómeno, es decir, si esto le afectó emocionalmente o le pareció apropiado o inapropiado (White, 2003). Por ejemplo, para el afecto estas materializaciones implican la modificación de los participantes y procesos, así como los procesos mentales afectivos y de comportamiento.

Las respuestas afectivas están cargadas emocionalmente conforme a las convenciones, más o menos universales, de una cultura y reflejan la empatía y solidaridad entre el hablante y su audiencia (Bednarek, 2008). Mediante el uso de elementos afectivos en su discurso, el hablante insta a sus interlocutores a compartir o comprender el sentimiento y espera que estos actúen en consecuencia. Si esto se lleva a cabo, entonces se produce un aumento en la solidaridad o empatía entre hablante y oyente y, además, se considera que es más probable que este último esté dispuesto a recibir otras posturas expresadas en la enunciación. De lo contrario, si la respuesta emotiva no es aceptada (p. ej., porque se considere el valor afectivo como inadecuado o disfuncional), existen más probabilidades de que la solidaridad disminuya y se produzca un desacuerdo ideológico (White, 2003).

El afecto, a su vez, se subdivide en tres instancias: cualidad, proceso y comentario. Los ejemplos en esta sección son los propuestos por Martin y White (2005) en la Tabla 2.

<i>Categoría</i>	<i>Funcion</i>	<i>Ejemplo</i>
Afecto como cualidad	Describir participantes Atribuido a participantes Forma de los procesos	“a sad captain” - epíteto “the captain was sad” - Atributo “the captain left sadly” - circunstancia
Afecto como proceso	Afectivo-mental  Afectivo conductual	“his departure upset him” Proceso efectivo “he missed them” Proceso medio “the captain wept”
Afecto como comentario	Desiderativo	“sadly, he had to go” Complemento circunstancial de modo

**Tabla 2** Tipos de afecto según Martin y White (2005)

La mejor manera de clasificar emociones es a través de la cartografía del terreno, entendido como un sistema de opuestos. Similar al *objeto de opinión* y al *origen de opinión* de Liu (2015), Martin y White (2005) llamaron *emocionando (emoter)* al participante consciente de la emoción y *desencadenante (trigger)* al fenómeno responsable de la misma. A partir de ahí, trazaron seis preguntas para su clasificación:

1. Sentimiento: ¿se interpretan dichos sentimientos culturalmente como positivo (algo que se goza) o negativo (a ser evitado)?
  - a. Afecto positivo: “the captain was happy”.
  - b. Afecto negativo: “the captain was sad”.
2. ¿Se materializan estos sentimientos a partir de una *emergencia de la emoción* que implique algún tipo de manifestación paralingüística / extralingüística manifiesta o se experimenta como un tipo de disposición emotiva o como un proceso mental en activo?
  - a. Emergencia conductual: “the captain wept”.
  - b. Proceso / estado mental: “the captain disliked leaving” / “the captain was sad”.
3. ¿Se interpretan los sentimientos como algo que se dirige o reacciona con algún desencadenante emocional concreto?
  - a. Es en reacción a otro fenómeno: “the captain disliked leaving”.
  - b. *Humor* sin direccionamiento: “the captain was sad”.

4. Cómo se gradan los sentimientos hacia un extremo u otro de una escala de intensidad (alto, medio, bajo).
5. ¿Implican estos sentimientos intencionalidad (en contraposición a reacción) con respecto a un estímulo *irrealis*? El afecto *irrealis* (Tabla 3) siempre implica la existencia de un *desencadenante* (p. ej., se le tiene miedo “a algo”).

<i>Sentimiento</i>	<i>Emergencia conductual</i>	<i>Disposición</i>
Miedo (-)	“tremble” “shudder” “cower”	“wary” “fearful” “terrorized”
Deseo (+)	“suggest” “request” “demand”	“miss” “long for” “yearn for”

**Tabla 3** Ejemplos de afecto irrealis (adaptado de Martin y White, 2005)

6. Existen varios conjuntos principales relacionados con la in/felicidad e in/seguridad. Mientras que in/felicidad cubre las emociones relacionadas con “asuntos del corazón” (amor, odio, tristeza, etc); la in/seguridad cubre emociones relativas con el bienestar ecosocial (ansiedad, miedo, confianza, etc).

Los ejemplos que aquí se aportan no son exhaustivos, aunque ofrecen una amplia panorámica de la gama de significados afectivos existente. Además, ponen sobre la mesa que la selección léxica implica gradación en la mayoría de los casos. A la hora de detectar patrones distintivos de este sistema, es interesante determinar el proceso atributivo relacional con un participante consciente que implica el verbo “feel”: “I feel” / “It makes me feel”.

- a. {PERSON feels affect about something}: “I feel happy.”
- b. {it makes PERSON feel affect that [proposición]}:  
“It makes me happy that they have come.”



<i>Sentimiento</i>	<i>Emergencia conductual</i>	<i>Disposición o estado mental</i>
<b>Infelicidad</b>		
“misery” [humor: “dentro de mí”]	“whimper” “cry” “wail”	“down” [baja] “sad” [media] “miserable” [alta]
“antipathy” [sentimiento dirigido: “a ti”]	“rubbish” “abuse” “revile”	“dislike” “hate” “abhor”
<b>Felicidad</b>		
“cheer”	“chuckle” “laugh” “rejoice”	“cheerful” “boyant” “jubilant”
“affect”	“shake hands” “hug” “embrace”	“be fond of” “love” “adore”

**Tabla 4** Ejemplos de in/felicidad (adaptado de Martin y White, 2005).

### Juicio

El sistema completo de juicio (*judgement*) fue expuesto por primera vez por Iedema, Feez y White (1994) y evalúa verbalmente las actitudes del hablante hacia el resto o *estimación social* y la *sanción social* o comportamiento humano. Martin y White (2005) diferencian estas dos subcategorías a partir de dos casos hipotéticos: cuando se busca ayuda para un juicio de *estimación social* negativo se necesitaría la ayuda de un psicólogo, mientras que, para una *sanción social* negativa sería necesario reclamar los servicios de un abogado o un guía espiritual.

Los juicios de *estimación social* tienen relación con lo usuales o inusuales que son los comportamientos de un sujeto con respecto a los estándares sociales preestablecidos. La estima social está relacionada con la cultura interna (conversación, humor, etc.) de un grupo social y es esencial para estrechar lazos sociales. Su ruptura, en las sociedades que se aclaman como modernas, no suele afectar a la legalidad o moralidad. Sin embargo, sí existen patrones deseables que se alientan desde la sociedad, en detrimento de otros, considerados indeseables. Se divide en tres categorías:

1. *Grado de normalidad*, cuán inusual es alguien.
2. *Capacidad*, cuán capaces son determinados individuos.
3. *Tenacidad*, cuán responsable es alguien o cuán bienintencionada es esa persona.

Los juicios de *sanción social* están más relacionados con la cultura escrita como son los decretos, reglamentos establecidos por instituciones morales y el estado y que impliquen penas y castigos dado el caso de su incumplimiento. Se divide en dos categorías:

- *Veracidad*, cuánto se acerca a la verdad esta persona o cuán honesta es la persona
- *Adecuación*, cuán irreprochable es el comportamiento de esta persona.

Al igual que con el sistema del afecto, podemos reconocer evaluaciones positivas y negativas, es decir, hechos admirables o descartables por el hablante o conjunto de hablantes. Por ejemplo, cuando se hace uso contextual del lenguaje, el significado actitudinal de una palabra puede cambiar la valencia dependiendo de la comunidad de hablantes a la que vaya dirigida, es decir, su dominio específico. Por ejemplo “the slow food movement”, en contraposición a “fast food”, sería positivo, mientras que, para muchos dominios “slow” es un adjetivo considerado no deseable (Martin y White, 2005).

Los parámetros para organizar el juicio reflejan las distinciones gramaticales en el sistema modal. El grado de *normalidad* estaría relacionado con los modales de frecuencia, la *capacidad* con la aptitud modal, la *tenacidad* con los modales de inclinación, la *veracidad* con los modales de probabilidad, y la *adecuación* los modales de obligación. Por ejemplo, los modales de probabilidad suelen estar relacionados con juicios de *veracidad* que son gradados con verbos modales a través de la selección léxica:

- (1) a. He’s naughty.
- b. He’s CERTAINLY naughty.
- c. It’s TRUE he’s naughty.

Finalmente, otro eje de estudio del juicio se basa en torno a las estrategias de selección léxica que se han seguido. Se divide en dos categorías *juicio explícito* (o juicio inscrito) y *juicio implícito* (o indicios de juicio). La evaluación se presenta explícitamente utilizando un elemento léxico que carga el juicio de valor, p. ej.: “lazily”, “smartly”, etc (White, 2003). En otros casos, los juicios de valor se presentan a través de *indicios de juicio* del hablante. Estas enunciaciones son superficialmente ideacionales, pero sin embargo tienen la capacidad de evocar en el oyente respuestas enjuiciadoras. De esta forma, un hablante puede emitir un juicio de estimación social, utilizando palabras como “inaction” para referirse a las acciones económicas de un banco, o, evocar ese valor a partir de estrategias semánticas utilizando indicios como en “RBS for instance did not set aside any money for bonuses in third quarter.” Por supuesto, en esta última enunciación no aparece ningún elemento léxico evaluativo, pero su lectura detallada sí tiene el potencial de evocar evaluaciones relacionadas con la incompetencia en los lectores especializados. Estos deben tener una serie de conocimientos previos para inferir el juicio negativo de la situación.

Es posible observar patrones que permitan observar la gramática local del juicio. En este caso, es útil determinar un proceso atributivo relacional en el corpus de manera que adscriba una actitud relacionada con el comportamiento de alguien.

- a. [it was judgement for person/of person to do that] → IT was silly OF/FOR THEM TO DO THAT.
- b. [(for person) to do that was judgement] → (FOR THEM) TO DO THAT WAS silly.

### Apreciación

La *apreciación* es el tercer sistema de la *actitud*, por el cual se realizan evaluaciones de productos, procesos o acciones. Es decir, mientras que el sistema *juicio* suele valorar los comportamientos de las personas, el sistema *apreciación* evalúa entes inertes o constructos sociales abstractos, como en este trabajo son los planes y las políticas económicas. Los seres humanos también pueden ser evaluados mediante la apreciación si no se evalúan sus comportamientos éticos sino cuando son tomados como entidades no físicas en lugar de ser considerados como participantes en la sociedad de los que se espera un comportamiento determinado (Martin y White, 2005; Kaplan, 2004, White, 2003). Este sistema es similar al sistema de análisis de sentimiento a nivel entidad/atributo que se observará en el Capítulo 5 (Liu, 2015). De hecho, la apreciación se puede desgranar en muchas categorías que dependen de la naturaleza de la entidad evaluada. Así, se puede estudiar la reacción del hablante sobre las cosas que le llaman la atención y cómo a su vez esta se puede dividir en términos con características propias. Por ejemplo, las categorías de equilibrio y complejidad, el valor de una entidad, o cuantificar cuán innovadora o auténtica esta sea, etc.

En relación íntima con los procesos mentales, Martin y White (2005) consideran la observación de estos fenómenos:

- a) *Reacción*: evalúa la manera en la que la estética de algo nos afecta o impacta (“captivating”, “boring”, “stunning”, etc). Relacionado con la metafunción interpersonal.
- b) *Composición*: se evalúa la forma que tiene la entidad evaluada a través de la percepción (“harmonious”, “balanced”, “curvy”, “irregular”...). Relacionado con la metafunción textual.

- c) *Valuación*: se evalúa en cuanto al valor que se da a través de la cognición, a las opiniones consideradas del hablante (“profound”, “original”, “worthwhile”, etc.). Está relacionado con la metafunción ideacional.

Al igual que el *afecto* y el *juicio*, en la apreciación se reconoce claramente la polaridad en las evaluaciones positivas o negativas. Es decir, se otorgan propiedades deseables o no a dichas entidades abstractas. Algunos ejemplos en la Tabla 5.

	<i>Positivo</i>	<i>Negativo</i>
Reacción: impacto “¿me llama la atención?”	“arresting”, “captivating”, “engaging”...	dull”, “boring”, “tedious”, “dry”....
Reacción: calidad “¿me ha gustado?”	“okay”, “fine”, “good ”...	“bad”, “yuk”, “nasty”...
Composición: equilibrio “¿conjuntaba bien?”	“balanced”, “harmonious”, “symmetrical”...	“unbalanced”, “discordant”, “irregular”...
Composición: complejidad “¿era difícil de seguir?”	“simple”, “pure”, “lucid”, “clear”, “precise”...	“ornate”, “extravagant”, “byzantine”, “arcane”...
Valuación: “¿mereció la pena?”	“penetrating”, “profound”, “innovative”, “exceptional”...	“shallow”, “reductive”, “insignificant”...

**Tabla 5** Resumen del sistema de apreciación (Martin y White, 2005), traducción.

En cuanto a su gramática local, existen algunos patrones de apreciación para la lengua inglesa. Por ejemplo, el léxico relacionado con la apreciación se puede detectar a partir de un proceso mental que adscribe una *actitud* a una entidad: “I consider...”, “I see it as...”.

- a. [person consider something appreciation] → I CONSIDER IT beautiful
- b. [person see something as appreciation] → THEY SEE IT AS beautiful

**Realizaciones indirectas en el sistema de actitud**

Durante la presente exposición solo se han mostrado realizaciones evaluativas que están inscritas de manera directa en el texto. Sin embargo, existe una amplia variedad de significados ideacionales que invocan la evaluación de manera indirecta, es decir, los significados emocionales se producen a pesar de la ausencia total de léxico actitudinal que nos indique directamente el sentimiento (Martin y White, 2005).

El análisis de la evaluación invocada en un texto es crucial para especificar la postura que el investigador está teniendo durante la revisión del texto, ya sea esta de carácter táctico, complaciente o resistente. Por lo general, el emisor de un mensaje realiza de manera natural su posición y la postura que quiere que los lectores compartan con él.



Por otro lado, en el plano actitudinal, las inscripciones actúan como señales que indican al lector cómo leer las selecciones ideacionales que les rodean.

- (a) El uso de símiles o metáforas comparativas se puede usar para provocar afecto o juicio.
  - (2) They treated us like sheep.
- (b) La intensificación simple a través de variaciones del léxico también es otra estrategia para cambiar la intensidad de la valencia del afecto. Por ejemplo: *break* puede ser sustituido por términos como “smash”, “damage”, “dismantle”.
- (c) Las palabras organizadoras del texto como p. ej., “however”, pueden cambiar los valores actitudinales de sus cláusulas próximas. Sirven, así, como indicadores de factores no predichos, y por lo tanto negativos, que sirvan para alertar al oyente de que los valores actitudinales (positivo/negativo) en el texto están cambiando.

El hecho de reconocer la actitud inscrita e invocada significa que es posible la existencia de un código doble en las categorías linderas de actitud. Primero, cuando los actores son juzgados cuando llevan a cabo su acción, hablaríamos de un caso de apreciación invocada. Por otro lado, es pertinente volver a mencionar que cuando una actividad se aprecia de manera explícita como si fuera un ente / objeto, se invocaría un juicio sobre quien lo haya cumplido. Véase este contraste en el ejemplo propuesto por Martin y White (2005): en el primer caso “he proved a fascinating player” y en el segundo “it was a fascinating innings”.

Los significados apreciativos (“the film was very sad”) están muy cerca en términos semánticos de los procesos mentales (“the film moved me to tears”). Martin (2003) postula una distinción importante entre *valoración inscrita* y *evocada*. La importancia del estudio de la distinción entre valoración inscrita y evocada aportará información esencial para comprender los factores que desencadenan las inferencias en un texto y la interpretación del oyente:

- (d) La *valoración inscrita* se expresa de manera explícita en el texto (“a bright kid”, “a vicious kid”).
- (e) La *valoración evocada* se proyecta a través de la referencia a los eventos o estados que son apremiados (“a kid who reads a lot”) o desdeñados (“a kid who tears the wings off butterflies”).

Finalmente, existen también elementos paraléxicos que trascienden la actitud. A lo largo de esta sección, el foco central ha estado en los elementos léxicos gradables que conformarían el estudio lenguaje evaluativo. Sin embargo, Martin y White (2005) consideran reseñable el uso de la blasfemia o los *tacos*, maldiciones o interjecciones malsonantes pues estos no constituyen léxico gradable. En algunos textos, es claramente una amplificación del sentimiento; sin embargo, en otros acumulan sentimientos intensos en su significado de forma que es difícil clasificarlos como afecto, juicio o apreciación. Este es el caso de:

MY GOD, we had a ball doing it though, didn't we?

### 3.7.4 Compromiso

El *compromiso* es el segundo sistema de la teoría de la valoración que define los valores dialógicos entre el hablante y el oyente del texto, en relación con la metafunción interpersonal del lenguaje. Martin y White (2005) consideran la interacción verbal como un motor y una realidad básica en el lenguaje, en la que toda comunicación es esencialmente dialógica y que, de una manera más o menos abierta, incluye una postura determinada y una actitud. Al hacer una enunciación el hablante no parte de cero, sino que debe afrontar la existencia de enunciaciones previas de otros hablantes, los cuales han tenido una postura determinada en un universo ideológico concreto dentro de una comunidad de hablantes que comparten creencias y valores. Por esto, es necesario establecer un marco que caracterice las distintas posibilidades que se abren cuando durante un intercambio textual los hablantes adoptan una postura con respecto a las instancias expuestas en el texto y a los interlocutores a los que se hace referencia, de manera directa o indirecta. Para su estudio, los autores toman como punto de partida la dimensión dialógica de Bajtin (1934) con vistas a determinar si la posición del hablante se puede considerar *expansiva* y respaldada por la comunidad de hablantes o si, por el contrario, esta va a ser *contractiva* y contestada, discutida o reprimida. Además de la perspectiva dialógica, es esencial comprender el discurso del hablante en relación con sus interlocutores desde tres perspectivas: (a) deseabilidad, (b) alineamiento y (c) solidaridad.

La *expansión* y la *contracción* son los dos sistemas principales del *compromiso*. Dentro del sistema de expresiones dialógicamente expansivas se encuentran dos sistemas: la *consideración* (*entertain*) y *atribuir* (*attribute*). Dentro del sistema de expresiones dialógicamente contractivas se encuentran el *rechazo* (*disclaim*) y la *declaración*

(*proclaim*). Además de aportar una herramienta para caracterizar las estrategias de uso del significado interpersonal del hablante, se tiene en cuenta el trasfondo no verbal del diálogo: el contexto que forma el telón de fondo de variantes de la lengua y puntos de vista alternativos sobre los que se construye el texto (Kaplan, 2004; Martin y White, 2005).

Según el pensamiento bajtiniano, la *heteroglosia* corresponde a las distintas variantes de la lengua que compiten entre sí dentro de una comunidad lingüística determinada. En esta fricción constante se reflejan las enunciaciones a partir de factores sociológicos (edad, clase, étnia, género), sociolingüísticos (dialectos, registros y estilos) y factores relacionados con el dominio (lenguas de especialidad, sociolectos). Mediante el uso de una variedad de la lengua, se vislumbran las grietas que de manera simbólica evidencian la lucha por la hegemonía entre los discursos oficiales e insurgentes, entre discursos dirigidos al exterior y al interior de un grupo. La heteroglosia se encuentra en la escala microlingüística del enunciado. Mientras que las oraciones no están dirigidas aparentemente a nadie, los enunciados tienen un destinatario identificado. Por ello, la enunciación de un emisor cualquiera se caracterizará por estar siempre orientada hacia un interlocutor y que dentro de sus selecciones lingüísticas se incluirá una polifonía a partir del reconocimiento de múltiples discursos o el uso de fragmentos previos. Asimismo, los enunciados pueden ser una respuesta a una enunciación previa, o una anticipación a la respuesta del interlocutor (Kaplan, 2004; 2007). En la línea de Kristeva (1986, citada por Kaplan, 2004) ningún enunciado es totalmente novedoso, pues responde a otros enunciados pasados, de manera que se cimienta sobre ellos y los reelabora. Esta nueva elaboración se incluye en el repertorio colectivo y pasa a disponibilidad de ser reutilizada en enunciaciones futuras.

En contraposición a lo anterior, la *monoglosia* representa a enunciados emitidos de manera no dialógica en las que el emisor pretende reflejar de manera única su opinión. En un enunciado de carácter monoglósico el autor no ofrece posiciones alternativas en el discurso. Se producen omisiones implícitas facilita que el destinatario del mensaje se alinee con el discurso del emisor. Así, se pretendería anular la otredad, mediante distintos grados de violencia, para reducir la multiplicidad de discursos a una voz predominante y uniformadora (Altamirano, 2002; Martin y White, 2005).

Es necesario proponer un marco de posicionamiento intersubjetivo al necesitar un modelo de clasificación para los efectos dialogísticos claves, ya que estos están asociados con dichos significados. Primero, se debe tener en cuenta el rol en los procesos de



significación por los que el hablante enfrenta el texto que refleja sus opiniones y creencias con las distintas posiciones de valor asociadas en una comunidad de hablantes. Obviamente, las posiciones de valor que emite el hablante pueden ser o no deseables en mayor o menor grado por la comunidad de oyentes (Martin y White, 2005).

El segundo eje esencial es la negociación de alineamiento / no alineamiento con lo enunciado. Aquí se establece la relación en la que hablante y oyente (real o putativo) apoyan o no lo que se expone el texto. Un hablante que se refiere a un oyente putativo, como sería en el caso de un locutor en un medio de masas o un columnista de prensa, debe encontrar la manera de retratar el oyente modelo con el que el hablante estará más o menos alineado y por el que se anticipa un punto de vista con el que cree que va a estar de acuerdo, Por el contrario, el hablante anticipará estrategias ante una postura antagónica y generará estrategias para ganarse a su audiencia y sumarla al punto de vista que expone en su enunciación (Martin y White, 2005).

El tercer elemento dialógico clave es la *solidaridad*. Esta supera al alineamiento (o su ausencia), y trata la tolerancia ante puntos de vista divergentes. Cuando se establece una conexión entre lo expuesto en la posición del hablante y la recepción del oyente aumenta la solidaridad entre ellos y, como consecuencia, la posibilidad de lograr acuerdo ideológico. Aquí se entiende como la expansión del discurso dialógico para dar cabida a voces externas al diálogo y a puntos de vista alternativos al propuesto. Estos puntos de vista no son reprimidos, sino reconocidos; pues aún en desacuerdo, se les da validez y legitimidad (Martin y White, 2005).

## **Dimensiones dialógicas expansivas**

### **Consideración**

El sistema de consideración (*entertain*) hace referencia a las cadenas de palabras por las que la voz autoral indica que su posición es solo una de entre un número de posiciones posibles y que, por tanto, en mayor o menor grado, hace espacio dialógico para también introducir las posibilidades que plantea su proposición. Dentro del dominio de la semántica, coincide con la modalidad epistémica y la evidencialidad y la escuela sistémica lo ha estudiado en el sistema de modales de probabilidad. Su selección lexicosemántica incluye auxiliares modales (“may”, “might”, “could”...), adjuntos modales (“perhaps”, “probably”...), atributos modales (“it’s posible that...”, “it’s likely that...”). Además de la clara función dialógica que tienen los modales, los autores hacen referencia a las locuciones de modalidad, p. ej., los casos en los que se emplea una



proyección mental-verbal (primer ejemplo). Los verbos evidenciales como *seem*, *suggest* (segundo ejemplo) también entrarían dentro de las realizaciones léxicas de este sistema, pues la proposición se interpreta como casual y una consecuencia subjetiva de un proceso de deducción por parte del hablante (Martin y White, 2005).

I BELIEVE they have betrayed our people.

This mismatch SEEMS worse than it was ten years ago.

Por último, dentro del sistema de consideración se incluirían las *pseudo preguntas* que se realizan de manera retórica con el fin de abrir debate más que con la función de demandar información o servicios. A este tipo de preguntas se le ha llamado *pregunta expositora*.

Is Tara on a downhill spiral to her bad old ways?

### Atribución

Dentro de las expresiones dialógicamente expansivas (Martin y White, 2005), la atribución (*attribute*) separa de manera clara el punto de vista de la voz autoral de la posición de la fuente externa. Esto puede ilustrarse con un hablante exponiendo a un público una conversación de otras personas que tuvo lugar el año anterior, o una presentadora de noticias relatando una conferencia política de manera imparcial. Especialmente en este último ejemplo, la presentadora usa un estilo más impersonal o imparcial ya que se pretende transmitir una mayor sensación de veracidad con las ideas expuestas a través de distintas estrategias. La primera es la atribución de las fuentes de opinión con un alto nivel de relevancia en un campo (p. ej., los célebres “científicos de Harvard” que exponen su opinión en televisión sobre algún hallazgo relacionado en el campo la astrofísica) o se hace referencia a una multitud de personas que comparten una determinada opinión (p. ej., “La mayoría de los españoles consideran la corrupción y el desempleo como uno de los principales problemas”).

Normalmente, la atribución se realiza en el discurso mediante construcciones gramaticales que utilizan el estilo directo o el estilo indirecto y utilizan verbos comunicativos como *say*, verbos mentales como *think* (o sus nominalizaciones) o complementos circunstanciales. Por ejemplo:

Mr. Rajoy said that his goal was to arrange new Catalan elections within six months.

There is a common thought both in business and academia that the technology adoption process is a key component of success.

Otras instancias de atribución son aquellas que no tienen una voz autoral específica y se atribuyen a rumores a través de elementos léxicos como *reportedly* o patrones como *it is said that*:

The US President REPORTEDLY exposed "very basic misunderstandings" of how EU trade works during a meeting.

Dentro del subsistema de atribución, existen dos subcategorías: *reconocimiento* (*acknowledge*) y *distancia* (*distance*), las cuales se realizan a partir de distintos elementos. En el primer caso, se utilizan verbos como “say”, “report”, “declare”, “announce”, “believe” o “think”. En el caso de la distancia, se utilizan verbos que marquen especialmente la separación y exoneración de la responsabilidad con respecto del punto de vista de la voz emisora de la opinión. Se destaca aquí el verbo *claim*.

- a. Varoufakis has explained how the EU has overpowered Britain in the Brexit negotiations.
- b. Varoufakis claimed EU leaders were “emerged in a fantasy”.

### **Dimensiones dialógicas contractivas**

Mientras que los valores de consideración y atribución se encargan de abrir el espacio dialógico a otras voces alternativas, las categorías correspondientes al sistema de enunciaciones dialógicamente contractivas (*contract*) son el *rechazo* (*disclaim*) y la *declaración* (*proclaim*). Estas restringen la existencia de otras voces, ya sea de una manera absoluta o de manera progresiva. El rechazo, se refiere a los significados por los cual la alternativa dialógica se refuta directamente o es suplida por la proposición de la voz autoral. Por otro lado, en la declaración las alternativas dialógicas son confrontadas y desafiadas a través de la interpolación (Martin y White, 2005).

### **Rechazo**

Bajo el término *rechazo* (*disclaim*), se cubren las formulaciones por las que una expresión previa o una proposición alternativa es totalmente rechazada o reemplazada. Según Martin y White (2005), existen dos categorías principales de rechazo: *negación* (*deny*) y *contradicción* (*counter*). La negación es el recurso que se utiliza para introducir una posición alternativa en el diálogo y, desde ahí, reconocerla o rechazarla. En estos términos

dialógicos, el negativo no es la mera inversión del positivo, ya que la expresión negativa incluye per se la expresión positiva dentro de su enunciación. Mientras tanto, la situación opuesta no tiene lugar. Por ejemplo, la negación “There is nothing wrong with democratic socialism”, responde a la una creencia previa de que “There IS something wrong with democratic socialism”.

Existen variaciones en este mecanismo con respecto al oyente putativo de un texto y el alineamiento del hablante con este. Por ejemplo, si la expresión se llevara a cabo a través de los medios de masas, lo más probable es que la falta de alineación no sea en relación con el oyente, sino con un tercero al que a menudo se hace referencia de manera explícita. En otros casos, el oyente putativo sí es el receptor de la negación. Si esto es así, el emisor generalmente tendría en cuenta que sus receptores, o parte de ellos, poseen algún tipo de pensamiento o creencia contraria a lo expresado. Por último, los autores hacen referencia a un factor al que llaman *paradigma existencial (existential paradigm)*. Según este fenómeno, un hablante experto se dirige al público a rebatir malas comprensiones de forma constructiva más que antagónica, con la intención de aumentar los niveles de solidaridad por parte de la comunidad oyente.

La segunda subcategoría del rechazo es la *contradicción (counter)*. Esta contempla formulaciones que representan la proposición actual, de manera que reemplazan la proposición de otras voces textuales que, según el hablante, eran esperadas en su lugar. Aunque la relación de esta categoría con la alineación hablante-oyente es bastante similar a lo que ocurre en la negación, aquí esta tiene un efecto opuesto a lo generalmente esperado. En este caso, esta función alinea las visiones entre el hablante y sus oyentes para aumentar la solidaridad entre los interlocutores, ya que ocurre dentro de un sistema de creencias compartido. En el siguiente ejemplo, el emisor se muestra de una manera solidaria tan sorprendido como el destinatario putativo:

Surprisingly, euro bonds issued by a Germany-less euro zone would still compare favorably with those of the US, UK, and Japanese bonds.

Este fenómeno también ocurre en las cláusulas adversativas, donde la contradicción se realiza (a) a partir de conjunciones y conectores como “although”, “however” y “yet”, (b) mediante de complementos circunstanciales como “surprisingly” o (c) a través de complementos que denoten hechos contrarios a las expectativas como “even”, “only”, “just” o “still”. Por ejemplo:

EVEN THOUGH we are getting divorced, Bruce and I are still best friends.

## Declaración

La *declaración* (*proclaim*) es la segunda característica de la contracción dialógica. Se trata de todas las formulaciones en las que se realiza una enunciación que repentinamente rechaza o se superpone a todas las proposiciones contrarias, de forma que se limita el campo de alternativas dialógicas en un intercambio verbal. Martin y White (2005), identifican tres tipos de declaración: *acuerdo* (*concur*), *pronunciamento* (*pronounce*) y *refuerzo* (*endorse*).

Primero, el *acuerdo* implica formulaciones que anuncia un emisor que anuncia abiertamente estar de acuerdo o coincidir en conocimiento con un receptor. Se apoya en el uso de locuciones como “of course”, “naturally”, “not surprisingly” o “certainly”.

OF COURSE, subprime mortgages were hotbeds for fraud during the boom years.

La relación de acuerdo también puede llevarse a cabo a través de preguntas retóricas que no tienen que ser respondidas porque la respuesta es obvia e inevitable, muestra de que hablante y oyente, como en el caso de:

Should we stop this argument now?

Haciendo un análisis más refinado, existen dos tipos de acuerdo: acuerdo concesional (“Admittedly ... but”, “I accept that... however...”) y acuerdo afirmante (“naturally”, “obviously”). La funcionalidad de los valores del acuerdo se complica en este caso pues los autores muestran que suelen ocurrir como un precursor de una contradicción. Esto viene motivado como ejercicio retórico pues en él se hace una concesión al interlocutor de forma que alivia la tensión previa a un contraargumento dirigido hacia un lector putativo reticente de manera que se mantenga un umbral mínimo de solidaridad (Martin y White, 2005).

La segunda categoría de la declaración es el *refuerzo* (*endorsement*) donde se referencian las proposiciones externas que la voz autoral la entiende como totalmente correctas y legítimas. Utiliza procesos verbales (o sus nominalizaciones) como “show”, “prove”, “pronounce”, o “point out”. En contraposición a las atribuciones, las cuales hacen uso del estilo indirecto para disociar la proposición de la voz autoral, aquí la voz autoral se fusiona con la voz externa y asume todas sus responsabilidades. Un ejemplo sería el siguiente:

Five studies show that dependence is associated with greater inequality.

La tercera y última categoría dentro de este sistema es el *pronunciamento* (*pronounce*), y se refiere a las fórmulas que recalcan las intervenciones autorales o interpolaciones. Una de las más frecuentes es “The fact of the matter is...” que remarca el valor y garantía del autor sobre su proposición y, en ocasiones, implica una posible resistencia por parte del lector putativo contra la que el autor subraya la confianza en su declaración. En la línea del subsistema negación y contradicción, los autores consideran que existe una visión heteroglosa entre los lectores putativos y, como respuesta, hacen uso de la voz autoral para restringir esa diversidad. Un ejemplo sería el propuesto a continuación:

THE FACT OF THE MATTER is that Greece is now in this mess, and has to play by the rules.

Las consecuencias de las relaciones hablante-oyente asociadas con el uso de declaraciones variarán según si el desafío a la postura enunciada está dirigida al interlocutor o a una tercera persona ajena al intercambio lingüístico. Cuando se da el primer caso, el hablante suele establecer distintas estrategias (verbales o no verbales) que modulan y aportan un plus de solidaridad para afrontar la desalineación entre los pares. Ocurre lo contrario en la segunda situación planteada, el hablante construye la solidaridad con la segunda persona en detrimento de la tercera con vistas a maximizar el alineamiento del interlocutor en detrimento de las personas o entes referidos. Esta última estrategia *nosotros-ellos* se utiliza de manera frecuente en la propaganda política y en los editoriales periodísticos.

Las realizaciones lexicogramaticales en el pronunciamento son bastante más diversas que las otras clasificaciones vistas en este apartado. En cierto modo, esta es análoga a las distintas categorías de modalidad (p. ej., Biber et al., 1999) ya que estos se mueven en torno a dos ejes claramente diferenciados: subjetivo vs. objetivo y explícito vs. implícito. El primer eje cuestiona si la voz autoral expresa claramente su postura subjetiva (“I THINK he’s not a good person”) o si se ofusca mediante recursos distintos (“IT’S PROBABLE THAT he may not be a good person...”). El segundo eje implícito-explícito diferencia entre el uso de modales, el cual se codifica mediante la cláusula principal de la oración (“THE FACT OF THE MATTER IS that it was his worst address”) (Martin y White, 2005).

Para finalizar, los autores señalan que hay formulaciones en el pronunciamento que se agrupan con evaluaciones de alta probabilidad a lo cual se le ha llamado

intensificador (*booster* o *up-toner*), como en “He MUST be lying”. Dependiendo del tono de esta expresión, no obstante, esta puede ser entendida como consideración, pudiendo contraer o expandir el discurso (Martin y White, 2005). Este es el caso del siguiente contraste:

He MUST be lying, don't you think?

He MUST be lying.

### 3.7.5 Gradación

El tercer sistema de la teoría de la valoración es la *gradación* (*gradation*). Esta observa la fuerza de los significados y cómo estos escalan ascendente o descendentemente de manera que se superpone al resto de sistemas de la teoría de la valoración. Además de ser una propiedad general de los distintos valores de actitud y compromiso, tiene un rol central en la asignación de valencias de sentimiento. Mediante este sistema se aumentan o disminuyen los grados de positividad o negatividad y están en relación con la implementación de los *modificadores contextuales de la valencia* (*contextual valence shifters*) de Polanyi y Zaenen (2006), que se verán en el Capítulo 5.

Según Martin y White (2005), la gradación opera en dos ejes de escalabilidad: fuerza y foco. El primer eje gira en torno a la intensidad o la cantidad. La segunda gira alrededor de la proximidad que los elementos referidos con respecto a su prototipo y con la precisión a la hora de entrar dentro de una categoría. La gradación de la intensidad o cantidad opera en categorías que implican evaluaciones escalares a través de clinas de, por ejemplo, tamaño, grosor, proximidad, etc.

#### Foco

En el esquema de la gradación, se denomina *foco* (*focus*) al sistema de alineamiento de una entidad con su prototipo. El *foco* se aplica a las subcategorías que no son escalables, es decir, aquellas que no pueden utilizar unidades de medida.

This is a REAL campaign of hate.

I'm vaguely aware that there's SOME SORT OF crisis.

A través del foco es posible afinar la especificidad, de forma que el prototipo sea similar a la entidad real o, por el contrario, esta se atenúe. Según esto último, esta caracterizará la entidad al coincidir con el patrón de la categoría de manera parcial o de

manera marginal. Esto puede llevarse a cabo a partir de *hedges* como “kind of” o “sort of” por ejemplo:

I felt SORT OF bitter.

Cuando el término gradado mediante el *foco* es un término que no es considerado perteneciente al sistema de actitud (p. ej., “jazz music”, “husband”, “father”) existe una fuerte tendencia a recibir propiedades actitudinales y polaridad a partir de la clina de prototypicalidad. Por ejemplo, las instancias de agudización muestran una evaluación positiva (p. ej., a real politico, a true politico) mientras que las instancias de marginalidad – atenuación denotan una evaluación negativa (p. ej., “jazz of sorts”, “a sanctuary of a kind”). Esta actitud está sujeta por la semántica específica de la categoría gradada junto con influencias contextuales como puede ser la prosodia semántica de los valores actitudinales que están inscritos en el texto (Martin y White, 2005).

Por otro lado, cuando la entidad gradada es abiertamente actitudinal, el efecto causado en los oyentes variará dependiendo de si esta se agudiza o se atenúa. En el caso de la agudización actitudinal, el efecto es indicar una implicación máxima partir de la voz autoral (ya sea negativa o positiva) con vistas a alinear al lector con la postura defendida (p. ej., “a real brat”, “a real wonder”). Cuando la agudización tiene un sentimiento negativo, el efecto es indicar un descenso de la inversión del hablante en la posición de valor y, desde aquí, ofrecer un gesto conciliador que mantenga la solidaridad con quienes tengan visiones contrarias (p. ej., “bordering the unreasonable”). En cambio, cuando se da el caso de que los valores actitudinales positivos se atenúan, este contraste suele causar sorpresa. Por ejemplo, cuando la entidad tiene una polaridad positiva mientras que a su vez está rodeado de otras valoraciones negativas suele generar problemas en la comprensión inicial y, en ocasiones, en los niveles de solidaridad entre hablante y oyente. Por ejemplo:

His lack of etiquette and candor was resulting kind of sexy  
among the audience.

### Fuerza

La *fuerza* (*force*) es la segunda categoría del sistema de gradación y cubre el grado de intensidad y la cantidad de una evaluación. Los primeros procesos se conocen como *intensificación* (*intensification*) y pueden operar en calidades (p. ej., “slightly foolish /

extremely foolish”), procesos (p. ej., “This slightly saddened them” / “This greatly saddened them”) o modales de probabilidad, frecuencia, tendencia y obligación (“It’s just possible that” / “It’s very possible that”). El segundo tipo de evaluación en este sistema se denomina *cuantificación* (*quantification*) y se aplica a entidades en lugar de cualidades y procesos. Aportan información de número (“few miles” / “many miles”) y entidades clasificadas como incontables en referencias a clasificaciones de tamaño, peso, distribución o proximidad (p. ej., “a small amount” / “a big amount, nearby mountain” / “distant mountain”) (Martin y White, 2005).

Las intensificaciones se dividen en dos clases: *implícito* (*isolating*) y *explícito* (*infusing*) (White, 2003). La distinción se muestra cuando el ascenso o descenso es llevado a cabo por un elemento aislado e individual que tan solo realiza la función de establecer el nivel de intensidad, o si el sentido de ascender o descender está infundido con un significado que sirva a otra función semántica. Las realizaciones implícitas se ejemplifican de esta manera:

- Premodificación de un adjetivo (p. ej., “a bit miserable”, “somewhat miserable”, “relatively miserable”, ... “extremely miserable”) o premodificación de un adverbio (p. ej., “slightly abruptly”, “somewhat abruptly”, ... “very abruptly”).
- El *escalado* localizado o relativo con respecto a la intensidad se lleva a cabo a través de comparativos y superlativos (p. ej., “less / least miserable”).
- La *maximización* es el extremo superior de intensidad en una escala de intensificación (p. ej., los adverbios “utterly miserable”, “perfectly happy” o los adjetivos “red hot”).
- Colocaciones: son los casos en los que la lexicalización no usa partículas gramaticales sino léxicas (p. ej., “ice cold”, “brand new”, “crystal clear”) o que incluyan además un trasfondo actitudinal (“deliriously happy”, “ridiculously easy”).

En la *maximización implícita* no se usan partículas que conlleven el sentido de ascenso o descenso, sino que esta está incluida en el término. Varios términos independientes relacionados constituirían una escala de intensidad. Las subclasificaciones de este ítem son las siguientes:

- Calidad: “warm”, “hot”, “scalding”.
- Proceso: “I glanced over the manuscript”, “I looked over the manuscript”, “I scrutinised the manuscript”.



- Modalidad: “possible”, “probable”, “certain”.

Es pertinente explicar que las distintas formas de repetición son otra forma de intensificar. Esta puede ser a través de reiteraciones del mismo elemento léxico (“It’s hot hot hot”) o a partir de una cadena de términos relacionados semánticamente (“immature, irresponsible, disgraceful and misleading address”).

Para concluir, es posible utilizar significados figurativos (p. ej., metáforas y símiles) para intensificar los procesos. Obviamente, como en el caso de la lexicalización, se produce también una pérdida de significado léxico en mayor o menor grado. Esto puede ocurrir con la intensificación explícita (a) e implícita (b):

- a. He came out like a jack in box.
- b. Prices have sky-rocketed.

La siguiente clasificación propuesta por los autores sería, entonces, la cuantificación. Esta implica escalar con respecto a la cantidad (p. ej., tamaño, peso, fuerza, número) y cómo se distribuye una entidad en el espacio y el tiempo (p. ej., cómo se distribuye, cuánto dura) y proximidad en el tiempo y el espacio (qué cercanía, cuán reciente). La entidad cuantificada puede ser concreta (“a great shark”) o abstracta (“a great problema”). Además, la cuantificación gradúa con respecto a número (“few” / “many”...), masa (“small” / “big”, “heavy” / “light”) y proximidad (“near” / “far”) y de distribución (“long-lasting”, “wide-spread”).

Si bien, la cuantificación se realiza por lo general de manera explícita (“many”, “large”, “recent”, “long lasting”), existen cuantificaciones implícitas. A diferencia de lo que ocurre en el subsistema de intensificación, la cantidad no está incluida en un modificador, sino en el núcleo del sintagma nominal per se a través de unidades léxicas, generalmente colocaciones (número: “A THRONG OF digital images”; masa: “he’s a MOUNTAIN of a man”; extensión: “I see a PAUCITY of talent in this country”).

Para concluir, hay que destacar que el sistema de fuerza interactúa con la actitud en el sentido de que sube o baja el “volumen” de dicha actitud. De esta manera, la prosodia semántica generada causa que ambas categorías se asocien en total consonancia a la alineación y la solidaridad del texto. Por ello, se entiende que el aumento del tono de una entidad se infiere como mayor compromiso con la comunidad de valores (“The extremely rough decisión”). Lo opuesto, bajar el tono, supondrían un nivel menor de alineamiento con la posición de valor expresada (“Such an irrelevant decisión”).

### 3.7.6 Procesamiento de la valoración y análisis de sentimiento

En cuanto a las aplicaciones del sistema de valoración, han existido intentos de su sistematización casi desde el momento de su publicación. Esta clasificación aportaría más información que los sistemas clásicos de análisis de sentimiento y permitiría conocer aspectos interpersonales de la evaluación, p. ej., si una evaluación hace referencia a sentimientos personales, o si los sentimientos son compartidos por la comunidad, etc. Estudios como los de Taboada y Grieve (2004) y Whitelaw et al. (2005) propusieron una clasificación de adjetivos que se clasifican dentro del primer nivel del sistema de *actitud*. Estos tienen el fin de analizar textos y posteriormente determinar la subcategoría predominante en estos.

Varios estudios de Bloom y Argamon (p. ej., 2010) trabajaron en la detección de objetos en la teoría de la valoración, de manera que su sistema no solo extraía los adjetivos según su subcategoría de *actitud*, sino que asignaba un objeto a cada uno de ellos. Así, podía determinarse de manera automática quién emitía la valoración y hacia qué o quién se hace referencia (Benamara et al., 2017).

El sistema de valoración es realmente rico y detallado de manera que se puede obtener una gran cantidad de información a *partir* del análisis de un texto y clasificarlos según su categoría de actitud (*afecto, juicio y apreciación*) de la misma manera que los sistemas de análisis de sentimiento a nivel entidad pueden ofrecer descripciones de las distintas categorías de un texto. El sistema de actitud es un conjunto teórico que explica con mayor nivel de detalle las actitudes del hablante en torno a la existencia o no de subjetividad y si estas valoraciones se hacen a nivel ético (en torno a una escala individual de creencias), o moral (compartido según las normas sociales).

El sistema de compromiso permite conocer más sobre la óptica dialógica del discurso evaluativo a través de las realizaciones lexicogramaticales del problema de la voz, facilitando el conocimiento de más detalles sobre la voz emisora de la opinión y su relación en torno al objeto de opinión, explicitando factores sumamente importantes para la estimación del sentimiento de un texto. Esto es relevante, p. ej., en el caso de la negación en relación con el análisis de sentimiento. El conocimiento de los engranajes internos de la expresión de la evaluación puede ser muy provechoso en la extracción de patrones locales propios de los textos evaluativos. Más allá de los sistemas de actitud, los distintos subsistemas de compromiso permiten conocer más detalles sobre el alineamiento de la voz con las opiniones o el origen inicial de las mismas. Además, el sistema de

gradación aporta una aproximación estructurada al problema de la asignación de puntuaciones de sentimiento que vayan más allá de la mera clasificación categórica (positivo/negativo) de la polaridad y que permitan clasificar la actitud desde una lógica escalar que tenga en cuenta la perspectiva del emisor de la opinión. Por ejemplo, detectar niveles de convencimiento sobre lo expresado en la opinión.

Muchas de las expresiones relacionadas con *fuerza* y *foco* permiten calibrar las puntuaciones en el análisis de sentimiento. El sistema de gradación se implementa en todos los sistemas de análisis de sentimiento que integren reglas contextuales de intensificación y atenuación como las propuestas por Polanyi y Zaenen (2006). Por ejemplo, estas reglas están implementadas muy eficientemente en sistemas como *Lingmotif*, el cual es utilizado en esta tesis.

El modelo propuesto por la teoría de la valoración ha aportado un trasfondo teórico a los estudios de análisis de sentimiento y su comprensión es esencial para cualquier trabajo relacionado con esta disciplina. Ya que se trata de una aproximación teórica, minuciosa a la vez que compleja del estudio del lenguaje evaluativo, no conocemos sistemas de clasificación que aprovechen directamente todas las posibilidades de esta teoría con el fin de generar un clasificador automático del sentimiento que clasifique todos estos parámetros. Sin embargo, con mayor o menor detalle, sus tres sistemas están presentes en cualquier implementación de minería de opiniones.

### **3.8 El estudio de la opinión desde la perspectiva del análisis de sentimiento**

Una vez conocidos los principales enfoques que afrontan el lenguaje evaluativo, es relevante para este trabajo buscar direcciones que permitan aprovechar las oportunidades de automatización del estudio del lenguaje evaluativo que aportan los algoritmos de análisis de sentimiento. Esta tarea comenzó, de manera pionera a abarcarse desde la literatura académica en la década de 1990. Ya entonces se proponía la búsqueda de modelos que estudien la opinión a partir no solo del léxico sino mediante el estudio de los contextos, por ejemplo, a partir de criterios de direccionalidad como Hearst (1992), la búsqueda de segmentos subjetivos (Wiebe, 1994) o la tonalidad de los adjetivos en los textos y sus coocurrencias (Hatzivassiloglou y McKeown, 1997).

Para el problema del análisis de sentimiento, se parte aquí del estudio de la opinión de Liu (2012; 2015), que estudia las contribuciones semánticas y pragmáticas que las

distintas cláusulas u oraciones aportan al discurso al relacionar un contexto de entrada y un contexto de salida.

### 3.8.1 Aproximación a la expresión de la opinión

La opinión se entiende como un concepto amplio, y su estudio abarca, por un lado, una aplicación eminentemente práctica de los postulados teóricos expuestos hasta ahora (modalidad, evaluación, evidencialidad, valoración, etc.) y, por otro, como un sistema simplificado de información que asocia al sujeto y el objeto de la opinión. Por una mera cuestión técnica, es necesario aclarar que aquí el término *sentimiento* se usa para hacer referencia exclusiva a la sensación positiva o negativa que subyace a un determinado enunciado.

El sentimiento y la opinión son, no sorprendentemente, subjetivos y, en contraste, Liu (2015) considera que el sentimiento tiene relación con la relatividad de la experiencia, pues la opinión de distintas personas estará basada en la introspección de los fenómenos relacionados con la experiencia personal y la perspectiva de observación. Dibujando una analogía, el hecho considerado afortunado de encontrar un billete en la calle implica que la persona que lo perdió, por lo general, no se va a alegrar.

Otro factor muy importante tras la opinión es, de nuevo, la ideología: los anteojos con los que las personas observamos la realidad teniendo en cuenta los intereses, deseos, esperanzas y gustos internos de cada uno en torno a lo establecido. El teórico con mayor vigencia en la discusión del concepto de ideología Slavoj Žižek (2012), lo expone así:

'Ideology' can designate anything from a contemplative attitude that misrecognises its dependence on social reality to an action-orientated set of beliefs, from the indispensable medium through which individuals live out their relations to a social structure to false ideas which legitimate a dominant political power. (p. 3)

Aunque muchos de los valores que conjugan el conjunto de creencias son compartidos, la ideología no solo puede ser irracional y caprichosa, sino que es infinitamente diversa y difiere necesariamente de individuo en individuo. Por ello, para el estudio de la opinión es preferible contar con las distintas opiniones de múltiples personas antes que la opinión de un solo individuo. La novedad aquí reside en que los trabajos de procesamiento del lenguaje natural sean, quizá, la mejor forma de acceder a grandes bolsas de datos que incluyan experiencias desde múltiples puntos de vista ideológicos y, por ende, aproximarse a la opinión desde un enfoque cuantitativo.

Según Liu (2015), aunque las opiniones y sentimientos son de naturaleza subjetiva, al expresar el punto de vista de una o varias personas, por lo general, es necesario desvincularlas del marco conceptual de las emociones. Esto se debe a que las personas pueden expresar felicidad o tristeza por determinados hechos factuales y, por otro lado, la expresión de la opinión puede no contener sentimientos. Por ejemplo:

- a. I hit the jackpot while playing a \$1 Wheel of Fortune slot machine.
- b. I think that Tesla's revenue will go up next month.

La opinión puede clasificarse en dos categorías: (a) opiniones factuales y (b) opiniones emocionales. Las *opiniones factuales (fact-implied opinion)* o racionales son aquellas que no expresan emociones y provienen del razonamiento, las creencias tangibles y las actitudes utilitarias. Además, el sentimiento emocional proviene de respuestas no tangibles y sensaciones que profundizan en los estados mentales, así como la perspectiva psicológica de las personas. La opinión factual trata la deseabilidad de las informaciones basadas en hechos evidenciables. Además, cada dominio específico tiene su propia lista de hechos deseables e indeseables, así como un sistema estructural determinado. Existen, entonces, algunos patrones que pueden utilizarse para detectar el sentimiento de las opiniones factuales. Esta categoría se divide en dos subtipos (Liu, 2015):

1. Opinión factual personal (*personal fact-implied opinion*): se basa en la experiencia personal de un individuo y no se diferencian de las opiniones subjetivas. Por ejemplo:

I'm so distraught that Endesa's profit went up.

Esta categoría incluye además a las metaopiniones, esto es, las opiniones sobre las opiniones. En ellas, el objeto de una metaopinión no es sino otra opinión (O'Leary, 2011).

I love Microsoft bashing.

2. Opinión factual no personal (*nonpersonal fact-implied opinion*): esta opinión es la encontrada frecuentemente de informes o noticias, se ajusta a un guion acordado de manera colectiva y no aporta la opinión de una persona en específico:

Tesla's revenue went up by 30%.

Obviamente, este tipo de información no tiene portadores específicos de opinión, aunque suelen citar a sus fuentes de información para aportar autoridad a la enunciación. Aun así, tratamos este tipo de oraciones como opinión por dos razones: (a) cada una de las oraciones indica un estado deseable o indeseable para las entidades que se basan en el saber común, (b) las personas que emiten dicha información posiblemente expresan opiniones positivas o negativas de manera implícita. Por poner un ejemplo frecuente en nuestros días, las personas comparten opiniones y enlaces de noticias en las redes sociales (p. ej., Twitter o Facebook) porque el mero hecho de que esa información tenga lugar les provoca algún tipo de emoción positiva que respalda la opinión expuesta. Si esto no fuera así, matizarían esta acción mediante la expresión de una opinión sobre la noticia expuesta. Asimismo, cuando se aportan opiniones factuales de carácter no personal, se debe tener en cuenta la posible falta de alineación entre interlocutores. Por ejemplo, lo que puede ser dado por positivo o negativo para el primero puede significar lo opuesto para el segundo, como en el caso de los siguientes ejemplos:

This shoe size is too small for me.

Google's profits went up by 30%.

En este caso la talla de los zapatos solo era pequeña en relación con el tamaño del emisor de la opinión, pero no lo es para todo el mundo. Asimismo, en la segunda opinión, los lectores pueden o no compartir esta buena perspectiva dependiendo de distintos factores, p. ej., ser accionistas de la competencia o tener una ideología contraria a los grandes beneficios de las corporaciones privadas (Liu, 2015).

### 3.8.2 Opinión y análisis de sentimiento

Para afrontar de manera sistemática el problema del análisis de la opinión, Liu (2015) establece inicialmente una cuádrupla que incluye los factores formales que constituyen una opinión: entidad, atributo, sentimiento, emisor de opinión y tiempo ( $e, a, s, h, t$ ). Una opinión consta en dos elementos claramente diferenciables: un objeto (*target*)  $g$  y un sentimiento  $s$  sobre el cual se expresa una opinión. El objeto de opinión  $g$  es una entidad ( $e$ ) sobre la que se expresa el sentimiento y está comprendida en distintos aspectos o partes ( $a$ ). Tanto las entidades, como sus partes, pueden tener un valor de  $s$ : un sentimiento etiquetado de manera categórica como positivo, negativo o neutro, o cuantificado numéricamente mediante una valencia dentro de una escala. Para resolver las complejidades que presenta la asignación de sentimiento a un objeto, las distintas frases

o palabras que expresan sentimiento, así como determinados objetos de opinión, tienen relaciones sintácticas observables mediante gramáticas locales. La primera dificultad es de clasificación y estructuración de datos, ya que es necesario determinar si existen uno o múltiples objetos de opinión y dentro de cada objeto, si existen uno o varios atributos. Posteriormente, se pueden diseñar distintas estrategias para relacionar objeto y sentimiento de manera que se puedan sistematizar algoritmos para llevar a cabo esta tarea.

El sentimiento  $s$ , asigna una puntuación que determina la orientación semántica de manera que le asigna un valor en forma de puntuación a determinado aspecto o entidad. Los elementos que constan en los resultados del clasificador tras la asignación de polaridad o valencia se representan mediante la tupla  $(y, o, i)$  donde  $y$  es el tipo de sentimiento,  $o$  es la orientación del sentimiento e  $i$  es la intensidad del sentimiento.

En cuanto a la información de sentimiento resultante, existen dos enfoques de sensibilidad principales dependiendo del nivel de granularidad de la intensidad del sentimiento. Los valores asignados de sentimiento pueden ser de carácter polarizado, frecuentemente denominado en la literatura como enfoque *Thumbs up / Thumbs down*: positivo, negativo, y, en ocasiones, neutro (Pang, Lee, y Vaithyanathan, 2002) o gradado, denominado en la literatura enfoque *Seeing Stars*, de manera que presente una puntuación basada en una escala, por ejemplo, en una escala  $[-2, 2]$ , -2 sería el polo negativo, 1 ligeramente negativo, 0 neutro, 1 ligeramente positivo y 2 el polo positivo (Pang y Lee, 2005). La polaridad y la intensidad de un elemento léxico pueden modularse al interactuar con otros elementos contextuales de manera que se determina a partir de un proceso de composición que captura la manera en la que las expresiones de opinión interactúan entre sí a través de la aplicación de distintos operadores de intensificación, atenuación o inversión (Polanyi y Zaenen, 2006). Brevemente, un ejemplo de inversión sería “He’s not a sympathetic person” en el que la polaridad positiva se invierte con la negación. Otro factor interesante es que determinadas palabras consideradas neutras, en combinación con otras, pueden cambiar su polaridad. Por ejemplo, las unidades léxicas “steady” y “job” tienen un sentimiento neutro, pero en conjunto, suelen ser consideradas como positivas (Hass y Versley, 2015). Este último punto será abordado con más profundidad en el capítulo dedicado a los sistemas de análisis de sentimiento.

Mientras que el sentimiento es sencillo cuando es explícito, este es difícil de detectar (incluso para una revisión manual) cuando está implícito. Es posible encontrar evaluaciones irónicas que aporten el sentimiento contrario al expresado literalmente a las llamadas *expectativas contrariadas* (*thwarted expectations*) en las que varias opiniones



negativas encadenadas concluyen con una evaluación positiva, de manera que provoque una reacción sorpresiva en el oyente, como en el caso expuesto por Benamara et al. (2017):

The characters are unpleasant. The scenario is totally absurd.  
The decoration seems to be made of cardboard. But, all these  
elements make the charm of this TV series.

El emisor de opinión (*h*, *opinion holder*) sería el sujeto que expresa su opinión sobre una entidad. Para muchos sistemas, este valor es importante cuando se pretende seleccionar la información, pues dependiendo de quién sea la persona u organización que aporte esta opinión permitirá valorar su influencia y credibilidad. En textos como críticas y reseñas de productos, la fuente de opinión suele ser el autor de la publicación; en otros, como las noticias o las narraciones, determinar el emisor de la opinión suele ser una tarea de gran complejidad que además, en bastantes casos, es de carácter irrelevante.

Finalmente, el tiempo de la opinión (*t*): referencia a la fecha y hora de publicación de la opinión. Como es el caso en este trabajo, el tiempo es esencial en el análisis de series temporales de sentimiento y así observar las tendencias y la evolución del sentimiento de una entidad a lo largo de un intervalo determinado. Así, se pueden realizar distintas metodologías de trabajo p. ej., observar su posible estacionalidad y establecer relaciones causales entre el sentimiento y fenómenos externos (Farzindar e Inkpen, 2015; Fernández-Cruz y Baixauli-Pérez, 2018).

Pueden añadirse nuevas dimensiones de complejidad a esta cuádrupla de análisis de la opinión al observar de manera individual las distintas características del objeto de opinión, a partir del esquema del *análisis de sentimiento basado en aspecto* o *en características* (*aspect-based sentiment analysis* o *feature-based sentiment analysis*) (Hu y Liu, 2004; Liu, 2010). De esta forma, se define a la opinión como una quintupla (*e*, *a*, *s*, *h*, *t*) donde *e* corresponde a la entidad del objeto de sentimiento, *a* refiere a cada uno de los aspectos o partes de una entidad, *s* es el sentimiento de una opinión sobre un aspecto *a* dentro de una entidad *e*, *h* es el emisor de la opinión y *t* es el tiempo en el que se ha hecho dicha opinión. El sentimiento *s* puede ser positivo, negativo o neutro o tener una puntuación dentro de una escala. Cuando una opinión califica al conjunto de la entidad, se utiliza el aspecto especial GENERAL. Para aclarar, tanto *e* como *a* representan al objeto *g* expresado en el modelo.



El factor *entidad* (*e*) hace referencia al global de un producto, servicio, tema de discusión, persona, organización, asunto o evento que es evaluado. Las entidades son definidas mediante la tupla  $e: (T, W)$ , donde  $T$  es definido como una ontología de las distintas partes y subpartes que componen la entidad, mientras que  $W$  es un conjunto de atributos que pueden describir a la entidad. Además, cada parte o división puede tener su propio juego jerárquico de atributos. Liu (2015) propone este ejemplo para ilustrar esto: una cámara de fotos réflex tiene un conjunto de partes (objetivo, batería, lente) y a su vez estos tendrían un subconjunto de atributos o aspectos p. ej., en la batería, sería la *vida de la batería, tiempo de carga, capacidad...* Las partes entran dentro de una descripción jerárquica de entidades basada en una relación parte-todo. En el ejemplo, el nodo raíz es el nombre de la entidad como sería la *cámara Nikon* y el resto de los nodos correspondería a sus partes y subpartes.

Otros elementos que profundizan en la exhaustividad del estudio y clasificación de la opinión son los siguientes:

- Motivo de la opinión (*reason for opinion*): evidencian los motivos para que la opinión sea positiva o negativa. La importancia de su análisis recae en la aportación de soluciones para los problemas planteados. Por ejemplo:

Revenue cuts pension because it is 'too high'.

- Calificador de la opinión (*qualifier of opinion*): limita o modifica el significado de la opinión. Pongamos como ejemplo el titular:

There is no penalty too harsh for a banker.

Aquí, se usa un sentimiento negativo para una condena. Sin embargo, no se hace referencia a todas las penas pues el motivo por el que el sentimiento es negativo se debe al uso de *too harsh*, el cual solo aplica *for a banker*.

Aunque si tenemos en cuenta todo lo anterior, se llega a hablar de hasta ocho factores, la tarea de describir la opinión puede ser simplificada. Un alto nivel de jerarquización suele ser poco práctico y presumiblemente ineficiente en el dominio general. Un ejemplo extremo de lo expuesto sería el modelo de Salas-Zárate et al. (2016), que propone un sistema ontológico mediante el estándar *Web Ontology Language* para analizar las entidades del dominio financiero con, entre otras, 247 clases y 38 propiedades de objeto. Alcanzar un nivel exhaustivo de detalle supone una tarea altamente compleja

y, al contrario de lo que se podría esperar, dada la creatividad del lenguaje natural, no produce grandes beneficios y su aplicación supondría menores cotas de eficiencia.

Liu (2015) propone simplificar las ontologías en el análisis de sentimiento de manera que la estructura lógica de partes y atributos profundice sólo en dos niveles y que este nivel inferior sea denominado aspecto, el cual hará referencia ambos. En un árbol simplificado, el nodo raíz hará referencia a la entidad, y el segundo nivel incluiría a los distintos aspectos de la entidad. Por otro lado, unos niveles muy altos de simplificación en el análisis pueden tener como resultado una pérdida significativa de información. Este caso se da, por ejemplo, en el análisis de la oración propuesta por el autor:

The INK of this printer is EXPENSIVE.

Según el autor, aquí el adjetivo *expensive* es una evaluación con sentimiento negativo en relación con el atributo “precio” del aspecto *ink*. Primero, se debe entender la relación expresiva tinta-impresora como una relación anidada parte-todo. En esta oración no se hace mención del precio de la impresora como una entidad global, sino a una de sus partes, por lo que no debería atribuirse sentimiento negativo al aspecto “precio” de la impresora. En caso contrario, si en un estudio dado solo fuera relevante el aspecto *ink*, entonces este debería ser tratado como una entidad fuera de la jerarquía de *printer*.

Como conclusión, Benamara et al. (2017) consideran que este modelo cubre los elementos principales para afrontar de manera práctica implementaciones dedicadas al tratamiento del lenguaje evaluativo. Desde un punto de vista lingüístico, especialmente si se compara con el modelo de Martin y White (2005), este modelo netamente formal puede resultar estático dada la dificultad de tratar el lenguaje evaluativo en escenarios en los que muchos de los elementos están inscritos en el contexto y, por lo tanto, no verbalizados.

### 3.8.3 El *Nuevo modelo dinámico del lenguaje evaluativo*

Benamara et al. (2017) proponen un nuevo modelo evaluativo basado en el paradigma de la semántica dinámica de Kamp y Reyle (1993), Liu (2012, 2015) o el modelo de Contextual Valence Shifters de Polanyi y Zaenen (2006) con vistas a detectar y formalizar los aspectos pragmáticos y discursivos de la evaluación. Este modelo de evaluación, parte de los modelos lingüísticos anteriores para buscarle una implementación teórica sencilla que sirva como adaptador de estos a la disciplina de análisis de sentimiento.

El modelo se representa como un sistema de componentes  $\Omega$  y  $\mathfrak{S}$ , siendo el primero similar a la definición de opinión de Liu (2015), al que le añade un conjunto de

reglas contextuales. El primero de los elementos es  $\Omega = (e, a, s, h)$ , el cual es una cuádrupla que representa a las propiedades del lenguaje evaluativo (entidad, aspecto, sentimiento y portador), en el caso que el factor tiempo ( $t$ ) sea necesario para el análisis, puede añadirse a la cuádrupla. El sentimiento  $s$  se subcompone además de una cuádrupla ( $span, category, pol, val$ ): expresión evaluativa, categoría semántica, polaridad y valencia. Aquí el sentimiento  $s$  se estructura en términos similares a los de Liu, sin embargo, este solo se ciñe a tramos textuales compuestos solo por partes explícitamente subjetivas: adjetivos, verbos, sustantivos, adverbios y emoticones, y excluye a los operadores locales a nivel oracional. Los autores fundamentan esto con el objetivo de separar el valor descontextualizado de  $s$  de su interpretación contextualizada, considerada parte de las propiedades extrínsecas de evaluación (como se verá a continuación en  $\mathfrak{S}$ ). Además, la polaridad ( $pol$ ) incluye cualquier tipo de clasificación: p. ej., positivo-negativo, pro/contra, positivo/neutro/negativo, o cualquier número de estrellas en una escala.

El elemento de este modelo corresponde al conjunto de funciones  $\mathfrak{S} = \{F_1, \dots, F_n\}$  que reúne las propiedades extrínsecas de la evaluación, ajustando o adaptando los valores previos de  $\Omega$  cuanto esta se da en un contexto determinado tal que

$$\forall F_i \in \mathfrak{S}, F_i: \Omega \rightarrow \text{Actualiza}(\Omega).$$

**Figura 2** Nuevo modelo dinámico del lenguaje evaluativo (Benamara et al., 2017)

Este conjunto de funciones es abstracto y puede adaptarse según la necesidad del sistema de evaluación añadiendo u omitiendo las funciones necesarias que se incluirán en  $\mathfrak{S}$ , de modo que los usuarios puede especificarlos según las necesidades y las técnicas de procesamiento del lenguaje natural disponibles.

Este conjunto de funciones actuará como operadores de actualización complejos en los niveles de organización del discurso (oración, documento, extradocumental). La interpretación de estas funciones reflejará la influencia que tienen en las expresiones evaluativas. De esta manera, a nivel oracional, se puede diseñar una función para reflejar la negación o modalidad. A nivel documental, las funciones tendrán en cuenta las relaciones argumentativas para actualizar la polaridad que previamente tenía una determinada evaluación. Los autores proponen el siguiente ejemplo:

What a great animated movie. I was so scared the whole time that I didn't even move out of my seat.

Para poner un ejemplo de cómo funciona este sistema, en la primera oración “great” actualiza el sentimiento de “movie”, de manera que la tupla  $\Omega = (e, a, s, h)$  quedaría así:

$$\Omega_1 = \text{Actualiza}(\Omega_1) = (\text{movie}, \_, (\text{great}, -, +, 1), \text{hablante})$$

**Figura 3** Modelo de actualización del discurso evaluativo (Benamara et al., 2017)

En la segunda oración, originalmente quedaría así:  $\Omega_2 = (\text{movie}, -, (\text{scared}, \_, -, 1), \text{hablante})$ . Sin embargo, esta segunda oración habría que actualizarla tres veces: la primera por la intensificación que el adverbio *so* hace al sentimiento de *scared*, la segunda de carácter discursivo ya que la primera oración nos indica que el conjunto es positivo y una tercera actualización de carácter pragmático, que nos indica una evaluación positiva adicional. Se propone aquí una ejemplificación de la instanciación del análisis a tres niveles distintos (oracional, discursivo y pragmático), de manera que permite realizar análisis de sentimiento con sensibilidad a distintos dominios y contextos. Este proceso se llevaría a cabo de la siguiente manera:

$$\Omega_2 = \text{Actualiza}_{\text{oración}}(\Omega_2) = (\text{movie}, \_, (\text{so scared}, \_, -, 2), \text{hablante}).$$

$$\Omega_2 = \text{Actualiza}_{\text{discurso}}(\Omega_2) = (\text{movie}, \_, (\text{so scared}, \_, +, 2), \text{hablante}).$$

$$\Omega_2 = \text{Actualiza}_{\text{pragmát.}}(\Omega_2) = (\text{movie}, \_, \{(\text{so scared}, \_, +, 2), (\text{didn't move...seat}, \_, +, 2)\}, \text{hablante}).$$

**Figura 4** Modelo de instanciación de la evaluación según los distintos niveles lingüísticos (Benamara et al., 2017)

Este modelo es eminentemente descriptivo e idealista, de manera que se anticipa a futuros trabajos de análisis de sentimiento. Según la revisión que aquí se expone, no existe ningún sistema que lleve a cabo un análisis de evaluación a más de dos niveles. Los autores, proponen dos direcciones de exploración para mejorar el conocimiento de los distintos contextos: (a) aumentar el léxico de manera que se disponga de más información contextual y (b) combinar las distintas teorías lingüísticas con un enfoque de análisis de datos de carácter principalmente cuantitativo. De esta manera, se pueden obtener nuevas introspecciones que, a su vez, mejorarían la eficiencia de los sistemas de análisis de sentimiento.

### **3.9 Conclusiones**

En este capítulo se ha resumido de manera integral el estado del arte en el estudio del lenguaje evaluativo desde un punto de vista eminentemente lingüístico. Tras una caracterización de la evaluación en distintas disciplinas, hemos realizado un recorrido por las últimas cuatro décadas de desarrollo teórico, el cual coincide con el desarrollo de la lingüística de corpus y del cambio de paradigma textual de la lingüística. Por tanto, se han recopilado los principales enfoques teóricos que permiten acercarse a un entendimiento estructurado del funcionamiento del lenguaje evaluativo.

La evaluación sería entonces una función realizada por el texto a través de una acción explícita o implícita a través del uso de una amplia variedad de recursos lingüísticos. El estudio de la evaluación, afecto o valoración, entre otros términos, es una cuestión multidisciplinar que implica, además del estudio del lenguaje, conocimientos de psicología o sociología. La relevancia que ha tomado este campo de trabajo es indiscutible, ya que las teorías de la evaluación han servido de base teórica de cualquier estudio computacional de la opinión.

Asimismo, se ha realizado un recorrido general del desarrollo sistémico-funcional de la teoría de valoración por Martin y White (2005). En él, desde una perspectiva predominantemente interpersonal, se observan las acciones dentro del discurso, en particular las acciones de expresar actitud, posicionamiento y alineamiento con respecto a una audiencia determinada o un interlocutor, así como sus distintos niveles de gradación.

El discurso y los fenómenos sintácticos, textuales e intertextuales juegan un rol crucial en la resolución del problema del análisis de sentimiento. Sin embargo, enfoques tan refinados como el de la teoría de la valoración están destinados a llevarse a cabo de una manera predominantemente cualitativa a través de trabajos de análisis de discurso realizados por expertos de manera muy minuciosa. Debido a su naturaleza, es complicado implementar soluciones automatizadas (Liu, 2015) que permitan su estudio a través del procesamiento del lenguaje natural. Sin embargo, el estudio del léxico afectivo, los distintos patrones evaluativos y determinados aspectos contextuales deben incorporarse a los sistemas de análisis de sentimiento. El análisis de sentimiento no es necesariamente una meta en sí misma, sino una herramienta que permite realizar estudios de mayor calibre. Por esto, se considera aquí que el estudio textual de la evaluación puede combinarse con distintos métodos estadísticos para lograr obtener una mayor cantidad de

datos que aporten nuevas perspectivas para analizar la realidad discursiva de una manera más eficiente.



## 4. Lenguajes especializados y terminología

« El mundo era tan reciente que muchas cosas carecían de nombre, y para mencionarlas había que señalarlas con el dedo. »  
Gabriel García Márquez (1927-2014), *Cien años de soledad* (1982).

### 4.1 Introducción

Los textos del dominio general modulan el lenguaje mediante el uso de un léxico con un bajo nivel de especialización, de manera que el hablante procedente de cualquier espectro puede llegar a obtener sin mayor esfuerzo una visión general de un determinado asunto o materia de trabajo. Sin embargo, cuando el texto va dirigido a una comunidad experta, se requiere utilizar un dominio específico del lenguaje con la inclusión de elementos léxicos altamente especializados. Estos textos son ricos en terminología, palabras específicas de un determinado dominio que son utilizadas para, principalmente, conseguir mayores niveles de precisión y evitar ambigüedades que puedan dar lugar a errores en la comprensión. Dado que la muestra analizada en este trabajo está constituida por textos especializados, es pertinente y necesario estudiar el tratamiento y metodologías que tales textos requieren. Este capítulo trata los principales factores relacionados con los lenguajes especializados, la terminología y la gestión terminográfica. Por otro lado, se revisarán las características de los dos dominios principales de trabajo: el lenguaje periodístico y el lenguaje económico-financiero.

Cualquiera de los campos específicos más conocidos, como los revisados en esta tesis, u otros como la medicina y la ingeniería disponen de amplios lexicones y recursos lingüísticos que son específicos de cada campo y que, en mayor medida, no son comprensibles en toda su amplitud por el público no especializado. El objetivo subyacente y primordial de los lenguajes especializados y la terminología no es otro que tender puentes entre especialistas hablantes de esta o distintas lenguas y la resolución de situaciones comunicativas en situaciones especializadas (Cabré, 2003).

Los lenguajes especializados han oscilado entre servir a las necesidades de trabajo cotidianas y el profundo debate que ha suscitado la disciplina y sus cambios



paradigmáticos a lo largo de los tiempos. Asimismo, la terminología, como rama del conocimiento que estudia y clasifica el lenguaje especializado, no es una cuestión reciente, y su desarrollo metodológico ha corrido paralelamente junto al desarrollo de la ciencia y la tecnología desde el siglo XVIII. En *Las palabras y las cosas*, Foucault (1966/2005) muestra que hasta la actualidad el lenguaje científico ha operado de un modo particularmente diferente del empleado por el lenguaje general, en cualquier caso menos ambiguo y polisémico. El argumento central de la obra es que todas las épocas han tenido condiciones subyacentes que constituyeron los elementos que eran aceptables en el discurso, destacando el discurso científico. Afirma, además, que dichas condiciones del discurso han ido variando a lo largo del tiempo, acompañando a los cambios de paradigma. En la obra pone ejemplos de la interdisciplinariedad de estos cambios donde muestra paralelismos en el uso del lenguaje durante el desarrollo de tres disciplinas distintas: la lingüística, la biología y la economía (Aussenac-Gilles y Condamines, 2007).

La *tradición discursiva* es un concepto clave que destacadamente fue iniciado y desarrollado por la escuela germánica de lingüística románica a partir del trabajo de Coseriu (1955). Esta propone tres niveles del habla, entendidos como requisito básico para estudiar el lenguaje: (a) nivel universal del habla en general, común a todos los seres humanos, que permite la comunicación a través de signos lingüísticos, (b) nivel histórico, en el que la lengua es un sistema de significación desarrollado culturalmente a través de la historia y (c) textos o discursos concretos. Estos tres niveles se llevan a cabo de forma simultánea al hablar una lengua, por lo que las lenguas no pueden entenderse como un mero conjunto estructurado de signos sino como la producción de textos. Se destacan aquí las relecturas llevadas a cabo de los trabajos de Koch (1997) y Österreicher (1997), donde se define el concepto de la tradición discursiva mediante la reduplicación del nivel histórico. A partir de esto, cualquier enunciación en su nivel histórico debe cumplir con dos factores: el sistema lexicogramatical de la lengua y la tradición en la producción del discurso. Esta tradición adquiere carácter de signo a través de su recurrencia y reutilización, y es esencial para comprender los textos, de manera que, por ejemplo, no solo basta con que estos cumplan con las condiciones esperadas de gramaticalidad en su producción, sino que estarán estructurados de una determinada forma y harán uso de giros, formulaciones y estructuras deseables. De esta manera, las noticias económicas disponen de un formato y un uso del léxico y giros gramaticales marcados por la tradición (Kabatek, 2005).

Históricamente, pero especialmente tras el desarrollo de la terminología, el estudio de las lenguas de especialidad ha sido un campo de trabajo interdisciplinario vital para el óptimo rendimiento de las disciplinas y estrechamente relacionado con la lingüística, la lógica, la ontología y la ciencia de la información. Como cualquier componente del saber humano, la terminología está influenciada por los dominios específicos con los que trabaja. A diferencia del trabajo lexicográfico, que en principio procura aproximarse a la recopilación de la totalidad de las palabras, la terminología parte de la necesidad de delimitar el léxico a subconjuntos que constituyen el vocabulario de las lenguas especializadas. Ya en el siglo XVIII se realizaron clasificaciones estructuradas del vocabulario específico con vistas a fijar las denominaciones de los conceptos propios de las disciplinas, como es el caso de Lavoisier y Berthold en química, o la taxonomía de Linneo que sigue siendo hoy en día la base de la clasificación de los seres vivos en biología. Desde el siglo XX, el desarrollo de la ingeniería propone desarrollos en diccionarios técnicos, destacando el exhaustivo trabajo llevado a cabo durante la recopilación del *Technolexicon* elaborado desde 1907 por la *Asociación de Ingenieros Alemanes* que en gran medida acompañó a la fundamentación teórica desarrollada por Eugen Wüster, referente en el estudio, definición y normalización del estudio de la terminología. No obstante, ha sido en los últimos decenios del siglo XX cuando la terminología ha realizado un ejercicio de autorreflexión en cuanto a principios y métodos, para alcanzar así la categoría de disciplina científica. Sorprendentemente, el interés por la terminología desde una perspectiva lingüística no llega sino hasta finales de los años 60 al cambiar del paradigma lingüístico formal al paradigma funcional. Se supera el estudio exclusivo de la lengua como un sistema estructural y se pasa a estudiar además como un instrumento de comunicación. La lengua de especialidad se convierte entonces en un yacimiento importante de trabajos lingüísticos, al ser los términos y su puesta en escena elementos fundamentales para una comunicación especializada correcta y eficiente (Santamaría Pérez, 2009).

A diferencia del lenguaje general y la relación clásica entre significante y significado, la terminología se enfoca en sistematizar los principios que se utilizan para designar conceptos abstractos según patrones o principios generales y realizarlos a partir de una serie de términos que constituyen la representación lingüística de una red más o menos estructurada de conceptos. El término sería entonces el elemento utilizado para referenciar a la unidad tangible de estudio de la terminología. No obstante, existen

diversas posturas en torno a la definición de término. En efecto, Aldestein (1998, p. 1) enumera algunas de ellas:

- (a) unidad lingüística de uso en dominios de especialidad o propia de los lenguajes de especialidad;
- (b) unidad léxica propia de los dominios de especialidad;
- (c) unidad semiótica de los dominios especializados (tanto unidades lingüísticas como no lingüísticas);
- (d) parte formal de una unidad semiótica de especialidad;
- (e) parte formal (o parte significante) de una unidad lingüística de especialidad.

La palabra *terminología*, y las diferentes posiciones sobre su condición multifacética como campo de trabajo autónomo, es polisémica, y su definición es dependiente del punto de su estudio. Según la recopilación de Pérez Hernández (2002), la terminología puede considerarse desde puntos de vista no excluyentes y observar este conocimiento como (a) un recurso al que acceder, (b) un conjunto de metodologías y procedimientos, (c) un elemento clave en el flujo de comunicación, (d) una serie de organizaciones que intervienen en su creación, establecimiento y difusión o (e) una disciplina académica.

Dados los múltiples factores disciplinares que la terminología debe tener en cuenta, ¿estaríamos hablando de la terminología como una ciencia autónoma, un subcampo de trabajo de otra ciencia, o de una disciplina transversal? Existen tres posturas distintas ante la concepción disciplinar de la terminología (Santamaría Pérez, 2009):

- a. La terminología como disciplina autónoma y autosuficiente: la terminología parte de sus propias bases, con vínculos a otros campos (p. ej., Wüster, 1998).
- b. La terminología como parte de otra disciplina: para algunos (p. ej., Sager, 1990) entra dentro del estudio de la lingüística de la mano de la lexicografía.
- c. La terminología como materia interdisciplinar: al configurar su propia especificidad hace uso selectivo de elementos de las otras materias para construir su propio ámbito científico.

Teniendo en cuenta lo anterior observamos la terminología como un campo de trabajo líquido que adopta distintas formas según sea su campo de estudio continente. Destacamos las cinco perspectivas principales, basándonos en la recopilación inicial realizada por Pérez Hernández (2002) la cual se amplía en el siguiente listado:

1. En relación con la filosofía: se entiende la terminología como una serie de unidades cognitivas utilizadas para representar el conocimiento. El propósito de

estas unidades es doble. Por un lado, son unidades de conocimiento, pequeños ojos de buey para mirar al mundo en perspectiva; por otro lado, son unidades de representación, que aproximan a las estructuras organizativas del mundo especializado (Cabré, 1995). Aquí entendemos los términos no como elementos eminentemente asépticos cargados de complejas relaciones de significado que se estructuran dentro de una taxonomía organizada alrededor de un marco referencial, sino como entes dotados de pleno significado que permiten enunciar expresiones de carácter evaluativo en relación con las posturas y discursos sobre la naturaleza del ser humano y su relación con el mundo.

2. En relación con la lingüística, la terminología es una parte del estudio del léxico especializado a partir de criterios temáticos o pragmáticos, la terminología se entiende como una disciplina lingüística en cuando estudia al lenguaje en uso. Para la lingüística, los términos no son distintos a las palabras, ya que componen un subconjunto dentro del módulo léxico de la gramática de una lengua. A partir de aquí se pueden considerar sus múltiples aplicaciones, partiendo de disciplinas educativas como enseñanza de idiomas para fines específicos, al trabajo conjunto con subdisciplinas computacionales, como es en nuestro caso el análisis de sentimiento de dominio específico dada la necesidad de diseñar y utilizar bancos de datos de manera estructurada diversos o aplicar técnicas de extracción terminológica para el etiquetado de elementos léxicos con características sentimentales específicas (Sager, 1990; Santamaría Pérez, 2009).
3. En relación con las políticas de planificación lingüística, la terminología es un ámbito donde se puede intervenir para reafirmar normalizar el uso de los términos y la pervivencia de una lengua, garantizando su modernización continuidad como medio de expresión, especialmente ante el fenómeno de globalización lingüística. Esta planificación terminológica puede venir desde la proposición o desde la regulación. En el primer caso están organismos como el *Comité de Terminología del Instituto de Ingeniería en España*<sup>22</sup>. Un caso paradigmático de legislación en la normalización lingüística lo constituye la la provincia canadiense de Quebec,

---

<sup>22</sup> “El objetivo fundamental [...] es la adecuación de la terminología técnica, en particular. Buscar el vocablo más adecuado para los neologismos que exige la constante evolución de la técnica.” Obtenido en <https://iies.es/comites-tecnicos/comite-de-terminologia/>

que regula su política lingüística a través de la Ley 101, que encarga a la *Oficina Quebequesa de la Lengua Francesa* proponer directrices para la elaboración de terminología (Streicher-Arseneault, 2013).

4. En relación con los estudios científico-técnicos, la terminología refleja formalmente diversos mapas conceptuales, donde organizan y distribuyen saberes y términos con vistas a facilitar la transmisión de la expresión de ideas. A partir de aquí, se permite la elaboración de documentación especializada comprensible y reproducible posteriormente en forma de p. ej., tesauros, glosarios y diccionarios técnicos.
5. En relación con el usuario general, la terminología consiste en una serie de unidades de comunicación, útiles y procedimientos que han de ser evaluados en función de criterios de economía, de precisión y de adecuación de su expresión.

Cabré (1993) considera la terminología como una ciencia que se define “en relación con otras materias, de las que toma prestados un conjunto específico de conceptos” (p.71). En esta línea, Pérez Hernández (2002) considera que el desempeño de las tareas propias de la terminología no debe entenderse como un ente aislado del conocimiento que representan ni de la función comunicativa del léxico, por lo que la terminología como disciplina realizaría una labor inconclusa si no realizara estas tareas para transmitir y reproducir de manera eficiente el conocimiento especializado mediante el uso de la lengua especializada como herramienta orgánica.

Para afrontar esta labor, Cabré (2002) propone enfocarla desde el ámbito de la lingüística para implementarse con sencillez una teoría que describa los términos como unidades que, dada una realidad discursiva determinada, adquieran un valor especializado. Como Temmerman (2000) o Faber (2012), explica la terminología desde una perspectiva lingüístico-cognitiva y con un propósito comunicativo que trasciende a una aséptica organización conceptual. Esta visión previa, contrasta con una perspectiva mucho más relacionada con la naturaleza del lenguaje y sus características. Por ejemplo, los lenguajes de especialidad también pueden estudiarse desde el punto de vista del lenguaje evaluativo de manera que hoy en día se puede afirmar con claridad que muchos elementos del léxico especializado incluyen características evaluativas dentro de su significado y que, además, este depende en muchos casos del dominio específico (Moreno-Ortiz y Fernández-Cruz, 2015). En este sentido, podemos poner ejemplos típicos del dominio económico como los términos “growth” y “expansión”, que serían considerados portadores de una connotación positiva, en contraste con otros dominios

también especializados como la astrofísica, la hidráulica o la demografía: “the expansion of the universe / water / population”.

La transición de la consideración de la terminología como una disciplina dominada por la ingeniería y las ciencias naturales a retroalimentarse con la lingüística vino motivada por múltiples factores. Cabe destacar el cambio paradigmático que supuso el paso de la lingüística desde el prescriptivismo hacia el descriptivismo, de manera que se abrió el abanico de estudio hace designaciones y variantes que no estaban descritas de manera sistemática. Por otro lado, la lexicología y la lexicografía realizaron una aproximación a la terminología que, además de abrir nuevas posibilidades en el estudio del léxico, vino al menos parcialmente fomentado por el interés editorial.

El desafío a los patrones terminológicos grecolatinos en las lenguas europeas desde otras lenguas del mundo generó nuevas discusiones sobre cómo generar patrones de creación terminológica desde otras perspectivas e influencias (Sager, 1990). Especialmente relevante fue la revolución ocurrida a finales de la década de 1980 a partir de la informatización del proceso terminológico, pasando de una recopilación manual que se servía del papel como medio de trabajo principal al salto cuantitativo y cualitativo que ofreció la incorporación de las bases de datos como plataforma de trabajo. En el siglo XXI esta tendencia se ha generalizado gracias a la facilidad (y necesidad) de compartir el conocimiento a través de Internet o al desarrollo de herramientas de la lingüística computacional, junto al enorme crecimiento de la capacidad de procesamiento, que ha ampliado enormemente las posibilidades: recopilar corpus actualizados, técnicas de extracción terminológica (semi)automatizada altamente eficientes, reconocimiento automático de entidades mediante técnicas novedosas como el aprendizaje automático se suman a la enorme cantidad de recursos lingüísticos especializados que pueden encontrarse en la red. Aun estando asentados muchos de los principios y discusiones sobre la terminología como rama del conocimiento, junto con otros tantos elementos en discusión, las posibilidades de trabajo de recopilación, clasificación y análisis terminológico hoy en día son mayores que nunca y ofrecen nuevas perspectivas de reflexión y debate.

Al superar la idea inicial de una disciplina centrada en el estudio y clasificación de los conceptos, la terminología contemporánea recoge la tensión entre la eficiencia de la comunicación y la necesidad de una eficiente representación de conceptos con precisión dentro de un dominio. Por ello el desarrollo funcional de la terminología aumenta conforme aumenta la especialización del discurso, de manera que se produce

mayor sistematización en la estructura de los términos, no solo en su función de crear relaciones estables entre el elemento léxico y el significado, sino también en forma lingüística, de manera que en muchos dominios se acerca a la rigidez y sistematización de los lenguajes artificiales. A su vez, a medida que la terminología comparte su forma lingüística con el vocabulario general, esta tiende hacia el uso de la flexibilidad total del lenguaje natural y en su capacidad de desarrollar relaciones similares entre léxico y significado del discurso de lo cotidiano (Kageura, 2002).

Para concluir, la investigación contemporánea en terminología ha ampliado su alcance, abarcando desde aspectos fonéticos al desarrollo de aspectos más formales y conceptuales, incorporando estudios diacrónicos y sincrónicos. El uso de las nuevas tecnologías para el procesamiento terminológico rompe con la concepción rígida de la terminología, observa las relaciones conceptuales de una manera más natural y permite observar la diversidad lingüística en contexto, poniendo en valor la pluralidad y diversidad de la mano de la emergente y amplia gama de oportunidades de desarrollo teórico y técnico en el campo. Sin embargo, cualquier revisión sobre la terminología queda incompleta sin conocer los detalles de los fundamentos de la teoría general de la terminología desarrollada por Wüster y la Escuela de Viena a lo largo del siglo XX ya los que los subsiguientes desarrollos teóricos en terminología que toman como punto de partida la crítica a los planteamientos de esta escuela (Cabré, 1995; 2000; Kageura, 2002).

## 4.2 Los lenguajes especializados

Los lenguajes especializados son denominados de muchas maneras en la literatura: lenguajes de especialidad, lenguajes especializados y lenguajes con propósitos específicos, los cuales en ocasiones se utilizan de manera indistinta o se superponen. El estudio de los lenguajes especializados aparece al final de la década de 1960, a partir de una necesidad eminentemente didáctica de formar a profesionales en la comprensión de materiales y la comunicación técnica en lenguas extranjeras. Hoy en día las especialidades de dominio son diversas y están en constante evolución adquiriendo sus aplicaciones un carácter cada día más interdisciplinar (Gambier, 2016). Las lenguas de especialidad se aglutinan en torno a su tendencia elitizante y restrictiva que los ubica como “tipos mixtos a medio camino entre los lenguajes artificiales y las lenguas naturales” (Cabré, 1993).

¿Cómo determinar si un texto es o no especializado, y dado el caso, cuánto más (o menos) *especializado* está con respecto a otro? Establecer el carácter de *lo especial*



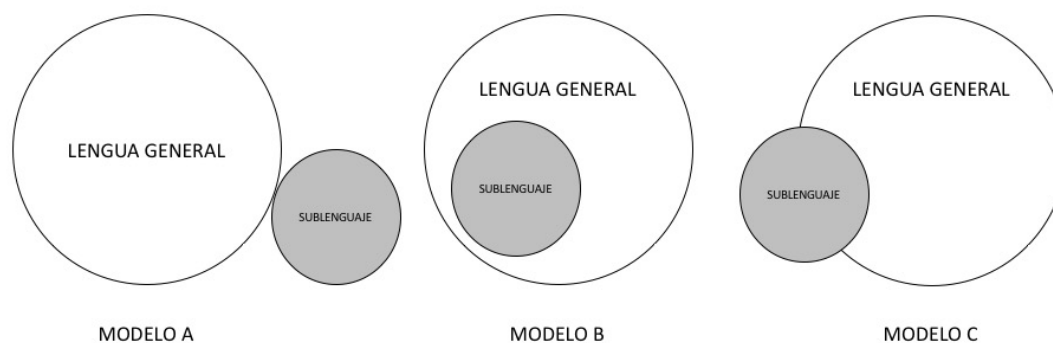
permitiría delimitar ciertas áreas fronterizas entre el lenguaje general, raramente definido, y el lenguaje de especialidad definido desde la oposición entre los dominios de especialidad por los que se materializan o por su grado de especialización, difícilmente objetivables. La investigación contemporánea considera que el estudio de los conjuntos *lengua general* y *lengua de especialidad* deben ser abordados como una línea continua: los elementos lingüísticos puestos en juego en el discurso especializado no pueden tener diferencias fundamentales de aquellos utilizados en el habla ordinaria y, dado el caso, se estaría hablando en términos de frecuencia de los elementos (Gautier, 2014). El anglista Michel Petit (2010) propone una definición precisa de especialidad tras haber abordado esta problemática epistemológica en muchos de sus trabajos:

Nous appellerons domaine spécialisé tout secteur de la société constitué autour et en vue de l'exercice d'une activité principale qui, par sa nature, sa finalité et ses modalités particulières ainsi que par les compétences particulières qu'elle met en jeu chez ses acteurs, définit la place reconnaissable de ce secteur au sein de la société et d'un ensemble de ses autres secteurs et détermine sa composition et son organisation spécifiques. (p. 9)

Una vez abordada la cuestión de la especialidad, es pertinente afrontar la consideración del estatus de lenguaje especializado en relación con el lenguaje general. Existen fricciones en los argumentos en torno a la consideración de las características compartidas entre la lengua general y las lenguas especializadas (Figura 5), principalmente: (a) el trato de estos los subconjuntos como lenguas de derecho propio (Hoffmann, 1998); (b) como variantes en los usos de la lengua general con un léxico específico (Rondeau, 1983); (c) subconjuntos pragmáticos del lenguaje general (Sager, 1990). En torno a esto, se puede afirmar que las lenguas de especialidad son lenguas *restrictivas* ya que se adquieren de forma consciente pero *flexibles* al incorporar al discurso elementos procedentes de los lenguajes artificiales que incluyen conceptos exclusivos a un dominio. Los textos especializados suelen realizarse, con mayor frecuencia, a través del medio escrito, y gracias a esto pueden ser estudiados y regulados con mayor facilidad (Montero Martínez, 2003).

¿Son los lenguajes especializados un simple uso especializado del léxico o, como el lenguaje evaluativo, opera a distintos niveles? Aunque existieron opiniones, p. ej., Hoffmann (1998), que consideran la presencia de terminología como condición suficiente y necesaria para determinar el nivel de especialidad de un texto, existen múltiples voces que elevan la discusión a niveles superiores. A tal efecto, Gautier (2014) considera que la terminología no debe dejarse atrás, pero que en ningún caso debe tomarse como un criterio definitivo del estudio de las lenguas de especialidad.





**Figura 5** Modelos de relación entre lengua general y lenguajes especializados (Santamaría Pérez, 2009)

De hecho, la bibliografía (p. ej., Cabré, 1998; Ciapuscio, 1998; Coseriu, 1977) discute sobre las variables funcionales y sociales, situación comunicativa y objeto comunicativo. A efectos prácticos, la mayoría de estos factores se pueden abarcar fácilmente mediante el análisis estructural y lingüístico, que como resultado determinará del grado de especialidad de un texto a partir de la descripción lingüística del mismo. Ciapuscio (1998) reflexiona que la superación del término como elemento central debe tener en cuenta la amplia gama de elementos oracionales y supraoracionales (texto y elementos pragmáticos) relativos al texto especializado. Hace dos décadas, Cabré (1993) expuso que los lenguajes especializados componen un tipo específico de registro, una variedad lingüística que se adapta a los usos comunicativos en el campo de trabajo cada disciplina. Al respecto, Gambier (2016) considera que, aunque hay una ambigüedad en la relación con la lengua general, en ningún caso se podría hablar de una lengua aparte, sino de una terminología y sintaxis específicas, estructuradas en una organización discursiva que busca la objetividad en una comunicación entre especialistas. Por último, Lerat (1995) puso sobre la mesa el factor cognitivo, ya que considera la lengua de especialidad como una lengua natural concebida como vector de los conocimientos especializados. Todas las áreas especializadas, así como las partes simbólicas no lingüísticas, las ciencias, los oficios y profesiones, son ámbitos de conocimiento especializado que requiere el uso de unidades y reglas específicas para cumplir con la necesidad comunicativa:

Une langue spécialisée ne se réduit pas à une terminologie : elle utilise des dénominations spécialisées (les termes), y compris des symboles non linguistiques, dans des énoncés mobilisant les ressources ordinaires d'une langue donnée (p. 21).

Las distintas corrientes coinciden en tres características: (a) estos lenguajes son grupos especializados por su propia naturaleza temática, su historia o el uso de la misma por sus usuarios (*tradición discursiva*); (b) las características de dichos lenguajes tienen

una relación estrecha por lo que no se puede hablar de fenómenos aislados sino de elementos dependientes entre ellos; (c) la función primordial del lenguaje es la comunicativa, con características descriptivas y clasificativas (Cabré, 1993). Cabré (1998) considera que la distinción aparentemente clara de las lenguas especializadas esconde una serie de rugosidades que requieren mayor precisión en el análisis. Por ello, además de la cuestión de los lenguajes de especialidad, se discutirán en dos conceptos más: la *noción de comunicación de especialidad*, y la noción de *texto especializado*.

En relación con la *comunicación de especialidad*, el lenguaje especializado se describe a partir de unas temáticas, unos interlocutores y de unas situaciones particulares de comunicación determinados sobre los que gira el discurso. El factor cognitivo entra en juego ya que se requiere un conocimiento propio para poner el resto de factores en juego. Este sublenguaje será un subconjunto de la lengua global y tratará una parte muy tematizada de la realidad que generalmente estructurará la percepción de esta, de manera que una gramática especializada clasifica y relaciona los términos de dicha lengua (Harris, 1986, citado en Montero Martínez, 2003). El uso del adjetivo *especializado* revela la esencia graduable del concepto y le permite ir acompañado de cuantificadores progresivos (“más/menos/algo/poco/muy”). Por lo tanto, el lenguaje utilizado por los profesionales puede utilizar varios códigos al mismo tiempo.

El aprendizaje de los contenidos supone la adquisición y adecuación explícitas a una serie de formas expresivas propias de la disciplina (Cabré, 1998; Ciapuscio, 2010). Las situaciones de comunicación tienen interlocutores que juegan papeles claramente delimitados (tareas rutinarias de una labor, transmisión de saberes) a través de roles sociales estratificados por nivel de especialización o jerarquía laboral. Por ejemplo, el rol comunicativo entre el médico y el enfermo o la comunicación entre el médico de familia y el médico especialista. Las jerarquías dentro del dominio jurídico tendrán en cuenta el uso lingüístico del juez, el fiscal, los abogados defensores y acusadores, el denunciante y el acusado.

En estos casos, como cualquier interacción, el uso de la lengua dependerá de las normas retóricas de producción y reproducción. La *comunicación de especialidad*, entonces, responderá a convenciones retóricas claras: *mecanismos intertextuales* (referencias a textos anteriores o discusiones generales dentro de una disciplina), *cotexto* (relaciones significantes y referencias a partes previas del enunciado), *pretexto* (motivación del enunciado), *subtexto* (alusiones, presuposiciones, estereotipos,

connotaciones) y *peritexto* (microenunciaciones anticipatorias durante el transcurso del enunciado) (Gambier, 2016).

En cuanto a los aspectos funcionales de los lenguajes especializados, por lo general, el objetivo de comunicar información relativa a un tema especializado se consigue a través de fórmulas textuales como la descripción, la definición, la clasificación, la enumeración, el cálculo, el razonamiento, la argumentación, la citación, la referencia, etc. A diferencia del uso del lenguaje general, los elementos comunicativos aquí tienden habitualmente a utilizar formas no personales y, aparentemente, poco subjetivas (aunque como podemos ver en la sección relativa a la teoría de la valoración, esto es solo en apariencia). Serían ejemplos de recursos comunicativos diferenciadores los siguientes: uso de verbos en presente, primera persona del plural, ausencia de exclamaciones y otros recursos expresivos, uso de oraciones cortas, ausencia de redundancia, entre otros (Montero Martínez, 2002).

Es en el momento del aprendizaje de la temática cuando se adquiere una serie estratificada de relaciones semióticas de cara a su conceptualización, clasificación y comunicación. Comunicar especialidades requiere niveles más altos de abstracción y precisión, por lo que se requiere una selección más precisa a nivel gramatical, semántico y sintáctico cuando se produce la comunicación. Los interlocutores, esencialmente los emisores, son los agentes especializadores del lenguaje. Sobre esto, Hofmann (1998) propuso un eje cartesiano de gradación o *continuum* en la especialización de la lengua que se remonta a la escuela de la estilística funcional checa (Havránek, 1964; Benes, 1969).

El primer factor de este eje será la *variación vertical*, la cual se compone de (a) la *intención* y (b) el *nivel de especialización* del emisor. Aquí generalmente será un especialista que hará uso de esta variedad del discurso para transmitir un mensaje lingüístico relativo a un tema pertinente. La intención del hablante se manifiesta entonces en el nivel de especialización del mensaje, a partir del reconocimiento de las características del conocimiento del receptor. En este caso, el conocimiento científico puede ser expresado mediante un lenguaje trivial, más próximo al conocimiento general, y alejado de la estructura conceptual propia del dominio o un lenguaje altamente especializado próximo a los lenguajes artificiales o semiartificiales (Sager, 1990), accesible solo a personas con un conocimiento profundo de la disciplina. Aquí debe diferenciarse entre un *discurso especializado* dirigido a los especialistas, que implica distintos niveles en la especialización; un *discurso didáctico*, con un objetivo

especializante en la enseñanza y el *discurso divulgativo*, que va dirigido hacia el público general. Condamines (1994) considera que el uso de la terminología en mayor o menor grado está condicionado bidireccionalmente por los distintos niveles de competencia. Al ser el discurso un ente meramente social regido por principios cooperativos, la comunicación especializada requerirá la adaptación del uso de terminología y formulaciones especializadas a cada nivel discursivo. El segundo factor se denomina *variación horizontal* y tiene relación con (a) la *temática* especializada que define a los distintos lenguajes (p. ej., dentro de la medicina estarían la otorrinolaringología y la endocrinología) y (b) la *perspectiva*, es decir, la óptica desde la cual se aborda el mensaje.

La tercera cuestión propuesta por Cabré (1998) es la naturaleza de los textos especializados. La escalada del nivel de estudio ha seguido un camino rectilíneo en términos históricos: palabra, sintagma y finalmente, situación, la cual ha podido llevarse a cabo a partir del desarrollo de la lingüística textual en los años 1980, que parte de la consideración integral del texto como objeto multiescalar. De aquí surgen dos líneas: la primera hace uso central del corpus especializado para constatar las introspecciones y observaciones puntuales (*corpus-based analysis*), y para detectar estructuras a partir del estudio sistematizado del corpus (*corpus-driven analysis*). La segunda teoriza sobre el funcionamiento de los textos, observando las fuentes de manera directa con fines como la enseñanza de lenguas o la redacción técnica (Gautier, 2014). A lo largo de la literatura, han existido múltiples trabajos en los que se han descrito las características más relevantes de las producciones especializadas en conjunto, además de las especificidades de cada género en la funcionalidad de sus textos. De acuerdo con Rondeau (1983), las dos características principales que se dan en los textos especializados ocurren en los planos textual, léxico y pragmático. Estos textos, globalmente, además de las características propias de cada temática, tienen una configuración específica muy diferenciada de la configuración de textos de otros géneros, como pueden ser los textos de carácter epistolar, literario, publicitario, etc. Además, teniendo esto último en cuenta, el criterio temático no sólo debe tener en cuenta a la materia en sí, sino también la perspectiva u *óptica* desde la que se observa a un objeto dentro de la especialidad de manera que se podría contrastar entre la producción textual de las ciencias exactas en contraste con las humanidades, o todas las ciencias, en contraste con los campos de actividad especializada donde se podrían establecer diversas perspectivas de abordaje del campo especializado, como por ejemplo si se trata a la disciplina desde la producción, o por el contrario se trata desde el

punto de vista del consumo, ya que la configuración de los textos suele variar significativamente según sea la perspectiva.

Los textos especializados, además de contener una serie de palabras funcionales propias del lenguaje común, contienen elementos léxicos particulares de la temática que tratan. Dentro del eje textual del estudio de las lenguas especializadas hay que tener en cuenta un segundo factor crucial para el estudio: el léxico que presentan. Desde una óptica léxica, los textos especializados lo son también desde la terminología que utilizan. No es posible hablar del léxico del lenguaje general y del especializado como compartimentos estancos, pues la terminología juega un papel fundamental en la caracterización del lenguaje especializado a la hora de establecer y clasificar sus distintos tipos, así como la difusa frontera con los textos generalistas. Pérez Hernández (2002) considera que se produce un constante trasvase de unidades entre la lengua general y las lenguas especializadas, y que esto es la causa principal de la imposibilidad de trazar una línea divisoria rotunda entre las palabras (como unidades léxicas de la lengua general) y los términos (unidades de los lenguajes de especialidad).

Por esto, las lenguas de especialidad presentan distintos grados de opacidad, especialización y apertura estructural que pueden graduarse de tres formas según la cantidad de terminología utilizada (a) léxico común a textos generales y especializados; (b) léxico fronterizo entre la lengua común y la especializada; (c) léxico claramente específico del texto especializado. Estos niveles dependerán no solo del tipo de temática, sino que, como se ha visto en la descripción de la variación horizontal, dependerá asimismo del nivel de especialidad del texto y de la selección léxica del emisor en relación con sus destinatarios (Cabré, 1998; Montero Martínez, 2002).

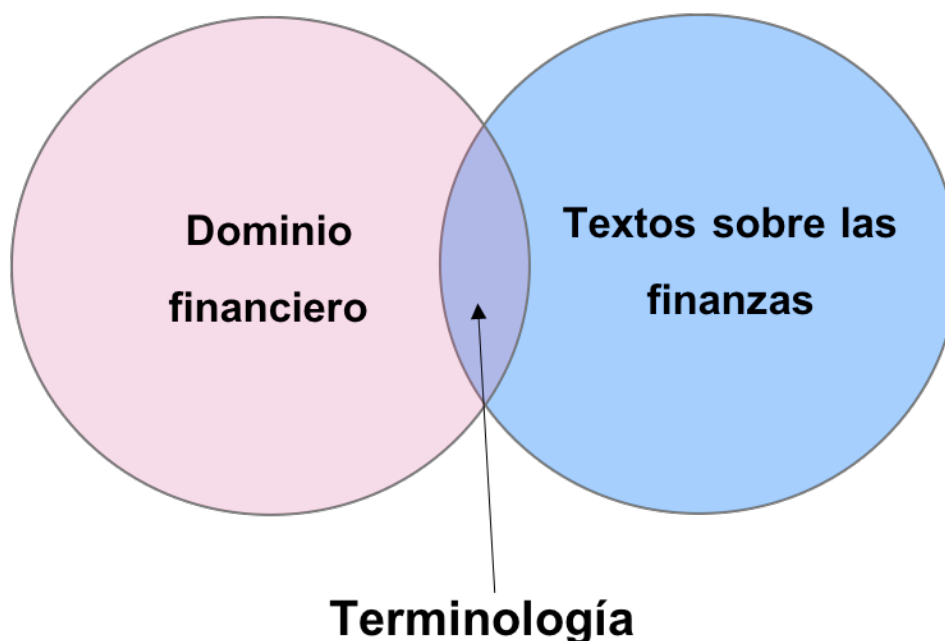
La porosidad de esta frontera entre el lenguaje especializado y el general se observa con los siguientes procesos de ida y vuelta en la especialización: la *terminologización* y la *vulgarización*. Los procesos de *terminologización* atribuyen al léxico de nuevos contenidos conceptuales a una unidad léxica preestablecida, como el caso de *rescate* o *recorte* en el lenguaje económico. El proceso inverso es incluso aún más frecuente debido a una mayor influencia de la ciencia y la tecnología en nuestras vidas. Este fenómeno se conoce como la *divulgación* o *vulgarización*. De esta manera, los medios de comunicación han popularizado el uso de términos en la lengua general sin que el hablante no especializado los comprenda en su totalidad ni los emplee de manera precisa (Pitch y Arntz, 1995; Santamaría Pérez, 2009).

Como paso intermedio, Gautier (2014) estudia la selección de los distintos elementos fijos que se utilizan sistemáticamente en los textos especializados (giros pragmáticos, fijación del esquema textual o los elementos estilístico-sintácticos) que permiten salir de los enfoques centrados exclusivamente en el léxico y ascender a niveles estructurales superiores. De esta manera, es posible iniciar la observación de las realidades extralingüísticas de la comunicación especializada.

El nivel superior sería, según Cabré, la *óptica pragmática* que complementa la caracterización de los textos de especialidad. Tras revisar la relación dentro de los géneros o la selección terminológica, un lector reconoce que está ante un texto especializado por otros factores periféricos a la lingüística. Un ejemplo de esto será la gran diferencia para el lector que existe en el grado de comprensión interlingüística. Al leer un texto especializado, un lector no iniciado puede llegar a afrontar un nivel de dificultad lectora similar al que supondría la comprensión de un texto en una lengua extranjera que no domina. En términos inversos, Cabré pone el ejemplo de los profesionales familiarizados con textos especializados en otras lenguas distintas a su lengua materna que, sin embargo, no dominan su uso el discurso general. Según Gambier (2016), en los factores pragmáticos externos jugarían un rol principal en el uso de las lenguas especializadas ya que las lenguas de especialidad son usadas e inferidas en situaciones particulares pues se articulan a partir de los actos del lenguaje y las terminologías que remiten a realidades fuera del ámbito lingüístico: objetos, procedimientos, procesos, etc. En otras palabras, los lenguajes especializados tienden a ser tautológicos y autorreferenciales. Así, se puede llegar a explicar por la redundante y sistemática realidad textual de los documentos especializados o el uso poco creativo de la terminología. Esto es un factor esencial en la asimilación y comprensión de los lenguajes especializados y se retornaría al eje cartesiano de variación comunicativa de los textos especializados de Hoffman (1998) en el que no solo la perspectiva, sino que el grado de especialización y sistematicidad de los textos serán factores correlacionales en el momento de ser comprendidos por sus destinatarios.

Como cierre de este apartado, es necesario concluir que, aunque existen lenguajes altamente especializados, no es posible hablar de lenguas separadas del lenguaje general, por lo que esta última tiene una base fundamental en las distintas aplicaciones. Es necesario, entonces, tener en cuenta los criterios expuestos aquí para trazar una frontera más o menos porosa entre el lenguaje general y la lengua de especialidad a partir de criterios eminentemente pragmáticos que se amolden al uso requerido. En esta zona de frontera (Figura 6) se sitúa el discurso de la prensa generalista en la sección de economía,

por ejemplo, y en ella se produce una *traducción* para el público que hace uso de un lenguaje generalista rico en terminología que se materializa en los textos sobre las finanzas, y se distingue del discurso altamente especializado emitido por ministerios, bolsa o bancos centrales y privados. La terminología sería por tanto el campo léxico común a ambos dominios.



**Figura 6** Puntos de convergencia léxica entre el dominio financiero y los textos divulgativos sobre las finanzas.

En el caso del enfoque seleccionado para el análisis de sentimiento, esto se hace patente desde la misma base teórica que asume el software utilizado, *Lingmotif* (Moreno-Ortiz, 2017). Dicho software incluye un lexicón de sentimiento de carácter general que incluye un léxico y unas reglas gramaticales específicas y además, en caso de que se requiera analizar un texto de especialidad, se ofrece la posibilidad de emplear un lexicón especializado que incluya un diccionario de terminología del dominio específico junto con su sentimiento en ese dominio. Ambos léxicos trabajarán en conjunta de forma que el lexicón especializado prevalece sobre el general. Se opta en este trabajo por el pragmatismo en el tratamiento de los distintos dominios específicos como una parte especial (temática) de la lengua general que se pone en uso en determinados contextos y se atiene en igualdad de condiciones a los principios gramaticales de la lengua.



### 4.3. Teorías de la terminología

#### 4.3.1 La teoría general de la terminología de Wüster

Aunque aquí consideramos la terminología como una disciplina dentro de la lingüística, centrada en el estudio y la sistematización del uso especializado del lenguaje, esta relación terminología-lingüística ha estado en disputa. Una de las propuestas terminológicas más influyentes en la literatura es la teoría general de la terminología descrita a partir del trabajo de la figura tutelar del ingeniero Eugen Wüster durante la tercera década del siglo pasado y ampliada póstumamente por sus colaboradores de la denominada Escuela de Viena. Dicha teoría dio inicio a la investigación terminológica moderna. Su principal aporte es la denominación y delimitación de conceptos de manera objetiva y objetivable, con vistas a que los esfuerzos se centren en la estandarización terminológica. Los impulsos teóricos posteriores son, en gran medida, herederos de su trabajo (Cabré, 2003). El enfoque de Wüster entiende la terminología como una disciplina estructural, considerándola un campo de estudio autónomo y transversal a otras disciplinas, como la lingüística, la lógica y la informática, a partir de la propuesta de una perspectiva sistemática de estudio y recopilación de las unidades. Resumimos brevemente cinco principios clave para la terminología wusteriana que son un punto de partida para cualquier desarrollo posterior:

- a) El *principio onomasiológico*: los términos se analizan a partir del concepto que representan. Por ende, se asume que el concepto precede a la denominación y a la definición.
- b) Los *conceptos están bien delimitados*: los conceptos de un mismo ámbito de especialidad mantienen entre sí relaciones claras de diferente tipo (Wüster, 1998). El conjunto de las relaciones entre los conceptos constituye la estructura conceptual de una materia. El significado de un término se configura mediante el lugar que ocupa en la estructura conceptual de una materia.
- c) Los conceptos, denominaciones y las definiciones terminológicas: el objetivo es el estudio desde la perspectiva de la normalización conceptual y denominativa. La formulación de las definiciones debe ser normativa. Dependiendo del espacio de comunicación estos elementos serán monolingües, dentro de una comunidad lingüística, o plurilingüe, si es internacional.



- d) *Principio de univocidad*: la normalización terminológica tiene como fin la garantía de la precisión y uso unívoco de los términos a la hora de comunicación profesionalmente mediante el uso de los términos normalizados.
- e) *Principio de sincronía*: la escuela tradicional de terminología no estudia el desarrollo del lenguaje, al ser el análisis conceptual su campo principal de trabajo, por lo que la sincronía es el único aspecto que se tiene en cuenta. Ignora la cuestión de si los conceptos evolucionan con el tiempo y si existe influencia lingüística en la conceptualización (Temmerman, 2000; Cabré, 2000).

La teoría general de la terminología toma como punto de partida el estudio de los términos a partir de los *conceptos*, con vistas a delimitar de manera precisa sus ámbitos mediante el análisis de sus relaciones y de esta forma establecer las denominaciones de manera estandarizada. Los conceptos se definen como “unit of thought constituted through abstraction on the basis of properties of a set of one or more objects” (ISO 1087-1, 2000). Los conceptos parten de una visión objetivista de la realidad en la que entidades abstractas que representan al *objeto individual* existen de manera universal y ahistórica e independientes de la realidad física, la comprensión o la imaginación y relaciones dialécticas entre humanos. Por lo tanto, el significado se entiende como un elemento objetivo que relaciona símbolos abstractos y elementos de la realidad que está aislado de cualquier tipo de realización del lenguaje natural (*parole*) (Temmerman, 2000; Sawyer, 2018).

En cuanto a la definición, la teoría general de la terminología busca definiciones que reflejen la ubicación del concepto dentro del sistema conceptual. Teniendo en cuenta su naturaleza abstracta, un determinado concepto se concibe entonces como la suma de sus características en estricta relación jerárquica con otros *conceptos*, las cuales son esenciales para definirlo. Dicho conjunto de relaciones se conoce como *intensión*. Por ejemplo, la *intensión* conceptual de una *deuda* “debt” cuenta con una *obligación* “obligation” como inversión en materias productivas, tiene una fuente “creditor”, unos plazos “timings” y unos intereses “interests”, etc. El segundo tipo de definición es a partir de la *extensión conceptual*, que hace referencia a la enumeración de la serie de conceptos jerárquicamente subordinados que comparten el mismo grado de abstracción. De manera que sea posible listar todas las especies u objetos individuales relativos a un concepto que brindan precisión o aportan características nuevas. Por ejemplo, extensiones del concepto “debt” serán los diferentes tipos de deuda p. ej., “secured debt”, “unsecured debt”, “mortgage”, “sovereign debt”, etc. La tercera sería la definición *parte-todo* que describe

la colocación de objetos individuales de manera que la revelación de las relaciones partitivas coincida con la definición de los conceptos (Wüster, 1998; Temmerman, 2000).

Los conceptos pueden organizarse en *relaciones lógicas* o *de abstracción* y en *ontologías*. Mientras que la primera clasificación estudia las distintas relaciones entre dos conceptos. Dentro de una jerarquía, los conceptos de “mortgage” y “car loan” tendrían una relación de coordinación y ambos tendrán una relación de subordinación con el concepto “debt”. De esta manera, los conceptos se pueden organizar de una manera jerárquica en relación con otros conceptos, a través de una red relacional normalizada de características que pueden compartir los distintos conceptos. Sobre el papel, los conceptos estarían organizados en torno a clasificaciones estandarizadas y reguladas. Como resultado de esta clasificación existen muchos ejemplos de uso cotidiano, como son la nomenclatura zoológica de Linneo que toma el latín como base y clasifica taxonómicamente a las especies, o la Clasificación Decimal Universal, que es, p. ej., la utilizada en la organización de las bibliotecas (Wüster, 1998).

La siguiente clave de la terminología son las *denominaciones*, entendidas como “los símbolos lingüísticos de los conceptos” (Wüster, 1998 p. 71). Si la denominación forma parte de un tecnolecto, se recomienda utilizar la palabra *término*, los cuales constan de una palabra o de un grupo de palabras. Estos se entienden como los elementos comunes que todas las personas utilizan como forma de clasificación verbal de los conocimientos y su objetivo es fijar el concepto a una realidad. Se propone una exhaustiva normalización de los términos, de manera que la variedad denominativa se decante a favor de una referencia exclusiva para evitar errores de precesión por homonimia o sinonimia. Entre los términos podrán encontrarse dos tipos principales de elementos léxicos o morfemas:

- *Palabras*: compuestas por uno o varios elementos léxicos ya sean palabras raíz, palabras compuestas (p. ej., en español “corto/circuit/ar”) o palabras derivadas.
- *Expresiones multipalabra*: constituidas por varias palabras ortográficas, en inglés suelen ser generalmente palabras compuestas que hacen referencia a una unión de conceptos p. ej., “slide way” que en ocasiones puede escribirse como una sola palabra “slideway”.

La influencia y autoridad que ha tenido la escuela de Viena durante décadas puede dar la apariencia de que esta posición es hegemónica. La influencia de esta escuela puede explicarse por la influencia de su centro de investigación *Infoterm*, avalada durante muchos años por la UNESCO. Aunque ha sido la primera gran teoría terminológica, existen críticas interesantes desde figuras muy reconocidas y perspectivas diferentes.

Se parte de que la metodología de trabajo de la terminología es onomasiológica, es decir, que parte del concepto para llegar a la denominación. En cambio, la metodología de la lexicografía es con mayor frecuencia de carácter semasiológico, esto es, se parte de un elemento léxico al que se le asigna un concepto y una definición. De esta manera es posible hablar de que, como aquí se entiende, la terminología asigna nombres a partir de una organización ontológica, pues otorga nombres a conceptos claramente dentro de un mapa conceptual preestablecido. El enfoque semasiológico trabajaría, por otro lado, con prototipos idealizados según las convenciones diatópicas y diastráticas que son descritos de manera aproximada y escalar siendo estructurados en todo caso *a posteriori* a partir de una denominación preexistente. Aunque la terminología sea una disciplina cuya tarea principal es la descripción de los conceptos, su labor es paradójicamente prescriptiva. En términos ilustrativos y explicándose de una manera poco matizada: se establecen los conceptos no a partir de la descripción de *cómo* usan los expertos los términos en el desempeño de su labor, sino a partir de una convención altamente regulada por una institución y claramente delimitada del concepto al que, en última instancia, se le ha asignado una denominación.

Autores como Weissenhofer (1995) o Ten Hacken (2015) discuten que, según sea la naturaleza del dominio, los requisitos para la creación de una terminología unívoca pueden variar por una serie de factores que van desde la estricta y necesaria definición científico-técnica al léxico del lenguaje especializado de carácter prototípico que surge del uso cotidiano del profesional. Weissenhofer, además, destaca que el rol de los términos se sitúa en el terreno de juego del uso de estos. De hecho, algunos dominios pueden exigir niveles más altos de determinación y definición que otros. Por ejemplo, los especialistas distinguen entre las situaciones donde el aspecto evaluativo juega un rol esencial y las situaciones dentro del dominio donde la definición es el objetivo principal. Este aspecto *evaluativo* de la terminología se puede mostrar a través del hecho de que los especialistas en un dominio suelen tener visiones muy diferentes, en algunos casos, sobre los conceptos centrales de su dominio. Por ejemplo, el concepto de “primitivo” en antropología sería un caso de lenguaje especializado más que una expresión estricta y regulada, de manera que los expertos de distintas corrientes antropológicas pueden, a través de la intuición y la discusión, tolerar condiciones suficientes y necesarias para su uso comunicativo aun partiendo de marcos teóricos e ideológicos en conflicto.

Temmerman (2000) critica el principio de univocidad y lo cataloga como *mantenido artificialmente* por los comités terminológicos, ya que el uso unívoco de los

términos se incumple durante el uso cotidiano. Por ejemplo, la no toda la terminología puede ser capturada a partir de descripciones o clasificaciones lógicas. Asimismo, la polisemia y la sinonimia no están ausentes en el lenguaje especializado.

Cabré (2000), por otro lado, critica desde una postura más constructiva el idealismo y asepsia de la teoría general de la terminología, al considerar que la unidad terminológica tiene interés *per se* y no únicamente desde la perspectiva de la normalización. La autora pone sobre la mesa todo el potencial comunicativo que tiene la terminología no solo en el discurso profesional, sino como se revisa en esta tesis: en el *plus* connotativo de prestigio que su uso aporta en la proyección hacia el público. La idealización se fundamenta mediante dos supuestos: primero, el conocimiento científico precede a cualquier expresión del conocimiento general; y, segundo, plantea la uniformidad del conocimiento científico y su independencia ante las realidades lingüísticas y culturales. Cabré considera la teoría insuficiente, al estar fundada en suposiciones difícilmente falsables. Así, Wüster carga las posibilidades y focos de estudio a otras disciplinas sin contemplar en ningún modo sus aspectos discursivos ni su proyección gramatical.

Tanto Cabré como Temmerman coinciden en la crítica a la consideración de los conceptos como compartimentos estancos de manera que no consideran su evolución en el transcurso del tiempo, estudiándolos de manera totalmente sincrónica. Asimismo, según la primera, los términos no tienen ninguna carga pragmática ni interpersonal, al considerarse solo dentro del registro formal profesional.

Aunque la teoría sí acepta los distintos niveles de especialización, no se tiene en cuenta en ningún caso la consideración de cualquier recomendación de uso contextual dentro de las definiciones de los términos (p. ej., el sentimiento). En relación con esto, Sager (1990) remarca la necesidad de tener en cuenta la dimensión comunicativa, de manera que los tipos de texto y las convenciones terminológicas pueden modificar la intención de un texto e influenciarlo a partir de la terminología utilizada. En nuestra línea, para ilustrar esto desde la terminología económica, la carga subjetiva al hablar de *negative growth* será claramente menor que mediante el uso discursivo del término *recession*. Aquí la perspectiva del hablante es relevante en términos discursivos de cara al uso terminológico, al ser una selección del consciente del hablante. Por tanto, es relevante considerar que existen múltiples elementos discursivos que trascienden el mero uso clínico de la terminología y que, por lo tanto, estos no deben ser simplemente obviados en su clasificación.

### 4.3.2 Las dimensiones lingüística y comunicativa de la terminología de Juan C.

#### Sager

El relevante trabajo del lexicógrafo argentino Juan Carlos Sager dentro de terminología del extinto *Instituto de Ciencia y Tecnología* de la Universidad de Mánchester se centró en la consideración de la terminología como una disciplina potencial y como una actividad práctica. Su obra más relevante *A Practical Course in Terminology Processing* (año) cuestiona la definición tradicional de la terminología como una disciplina independiente y la plantea contextualizada dentro de la lingüística aplicada, la lingüística computacional o la informática. Sager (1990, p. 3) la define principalmente como una teoría, “*a theory, i.e. the set of premises, arguments and conclusions required for explaining the relationships between concepts and terms*”. Estas prácticas serán solo medios para un fin y no propone establecer nuevos conocimientos sobre una determinada materia, de manera que se justifique ser una ciencia por mérito propio. El autor considera que, en muchos aspectos, la teoría de la terminología sería parte de una teoría de la lexicología. Esto se justifica de acuerdo con los objetivos prácticos en relación con la labor dual de los terminólogos como productores de recursos en forma de diccionarios automatizados y como asesores en el uso de terminología (Temmerman, 2000).

Sager (1990) considera tres dimensiones de la terminología: una *dimensión cognitiva* que relaciona las formas lingüísticas con el contenido conceptual, una *dimensión lingüística* que examina las formas existentes y potenciales y, especialmente novedoso, considera una *dimensión comunicativa* que añade la observación de los textos al análisis terminológico dejando atrás el estudio de los términos como meras etiquetas libres de contexto. La identificación de estos elementos parte desde el contraste entre textos a partir de la delimitación en unidades léxicas.

Comenzando con la *dimensión cognitiva*, se propone un modelo de organización espacial de los conceptos en una estructura similar a un eje multidimensional de coordenadas, las cuales dependen de la naturaleza intrínseca de cada tipo de conocimiento y de los solapamientos entre ellos. Cada uno de los elementos comprendidos entre estos ejes serían los *términos* que en colectivo conforman un *lexicón* de una disciplina especializada, cuya forma colectiva sería la *terminología* de una disciplina. A la hora de definir estándares, Sager habla de la dificultad de delimitar las definiciones de los términos sin el uso de la lengua (p. ej., mediante diagramas, matemáticas, mediciones o

imágenes). El autor consideraría *estándares* a estos últimos, mientras que a los conceptos que solo pueden ser descritos a través de la lengua, los consideraría pseudoestándares.

Sager (1990) considera la definición del concepto como axiomática, en los mismos términos que la discusión lingüística sobre la definición de palabra u oración. En cuanto a su clasificación, distingue entre cuatro tipos de conceptos:

- (a) *entidades*: que se derivan por abstracción desde objetos materiales o abstractos (entidades materiales, entidades neutrales, entidades abstractas, entidades representacionales y software).
- (b) *actividades*: procesos, operaciones y acciones realizadas por o para las entidades
- (c) *cualidades*: propiedades o disposiciones por las que diferenciamos las entidades.
- (d) *relaciones*: por las que nos establecemos entre algunos de los otros tres tipos de concepto.

Este marco teórico reconoce el problema que existe a la hora de estructurar conceptos, al no existir ningún modelo unificado de relaciones entre términos, lo cual es especialmente problemático para el diseño de bancos de términos. Existen, por tanto, relaciones complejas que no pueden ser capturadas de manera conveniente por estructuras directamente genéricas y partitivas. Ejemplos de tipo de relación son las siguientes (Temmerman, 2000):

- (a) está *causado por*: “fallout-nuclear explosión”
- (b) es un *producto de*: “paper-wood pulp”
- (c) es una *propiedad de*: “compressibility-gas”
- (d) es una *medida cuantitativa de*: “temperature-heat”
- (e) es un *instrumento de*: “computer-data procesing”.

Sager (1990) incorpora aspectos de multidimensionalidad al aplicar una *faceted classification* que indica el tipo de subclasificación de un concepto hecho de manera exhaustiva, como en el caso de los tesauros, en torno a una característica particular.

Finalmente, plantea el problema de la definición, en el cual no se puede contemplar una definición unívoca y estrictamente normalizada en todas las disciplinas por igual. Propone una serie de mínimos: (a) localizar el término dentro de la escala conceptual, (b) definición intensional que fije el significado especializado al término, (c) definición enciclopédica para aportar luz a la comprensión de los no-especialistas.

El segundo nivel sería la *dimensión lingüística*, Sager (1990) expone que la influencia del desarrollo de los enfoques basados en corpus para la recopilación léxica supuso además un cambio de paradigma en la terminología al contraponer el hecho comunicativo contextualizado con el punto de partida de la terminología clásica. Esta trabajaba a partir de lenguajes semiartificiales y altamente regulados mediante el uso de diccionarios descontextualizados que no estaban provistos aún de las funciones con las que ya contaban otros recursos léxicos más novedosos. La tarea del terminólogo raramente pasa por la tarea *onomasiológica* de denominar conceptos, sino más bien sería una tarea *semasiológica*, porque toma las palabras como punto de partida para definir su significado. La tarea del terminólogo pasa por el trabajo conjunto con un especialista, de manera que juntos puedan diseñar sistemas conceptuales para realizar colecciones de terminología coherentes con los objetivos.

En términos semánticos, este marco de trabajo supera la idea prescriptivista de la designación unívoca de conceptos para los términos y viceversa. Los términos ocurren en varios contextos lingüísticos diversos de manera que pueden existir variantes del léxico especializado dependientes del contexto. La terminología de Sager acepta con naturalidad la sinonimia y la homonimia en la expresión terminológica. Se pone en valor así la existencia de distintos niveles de representación lingüística en los entornos comunicativos especializados que naturalmente disponen distintas selecciones léxicas para un mismo concepto. Al romperse esta premisa, se hace necesario el establecimiento de criterios para la correcta identificación del nombre principal de un concepto, del cual el resto serían variantes, o problematizar sobre la definición sistemática del contexto regular donde ocurre un determinado paradigma (Sager, 1990). Este planteamiento se anticipaba ya a sistemas modernos como el *Ecolexicon* de Faber (cf. López-Rodríguez, Faber et al., 2010) que implementa de manera práctica las relaciones de sinonimia y polisemia en bases de datos terminológicas y los matices semánticos y pragmáticos que permiten realizar una selección terminológica adecuada para la práctica profesional.

El trabajo plantea también el problema de la génesis de un término. Considera este origen como orgánico, de manera que refleja el fenómeno de lengua viva y en constante evolución. Expone una teoría y puesta en práctica de la formación de términos de carácter descriptivo basado en la morfología y en los procesos clásicos de formación de palabras como patrones morfológicos, reutilización de recursos y neologismos. Asimismo, ofrece unas directrices y recomendaciones para la creación de términos.



Sobre esto, considera que el estatus pragmático de los términos es variable y se asocia con su novedad, es decir, con su asentamiento en una comunidad lingüística, su aceptabilidad, exclusividad de existencia y difusión. A partir de esto, un neologismo terminológico pasaría por varias fases hasta que las definiciones estipulativas del triángulo TÉRMINO – DEFINICIÓN – CONCEPTO se consolidasen en la comunidad lingüística o fueran descartadas. En la fase inicial se partiría de un número determinado de términos rivales en el que un término consigue una mayor predominancia, mientras que el resto queda gradualmente confinado a otros contextos. En casos como los utilizados en esta tesis, los distintos términos utilizados deben ser provistos de equivalentes en una lengua extranjera al no existir una traducción equivalente asentada, este fenómeno se conoce como *quality label*, para el cual, irónicamente, no se ha encontrado traducción al castellano. Esta “etiqueta” puede quedar restringida en el contexto en el que es utilizada a no ser que su uso se asiente y sea aceptada como término pleno. Esta etapa de nombrado se llama *proceso de terminologización*, y es esencial para la descripción de los dominios específicos.

Un agrupamiento particular de elementos léxicos como “offshore haven” (“refugio o paraíso fiscal”), que en un dominio no financiero sería una colocación libre sin mucho efecto en la comunicación, pero que en el ámbito específico debe ser tomado como una entidad cargada de significado y muy diferenciada de, por ejemplo, una colocación menos lexicalizada como *safe haven* equivalente a ‘refugio o puerto seguro’. En muchos casos, esta especial lexicalización suele llevar en inglés algún tipo de variante grafémica, como el uso de mayúsculas, cursivas o guiones (Sager, 1990).

El tercer factor, la *dimensión comunicativa*, fue la más novedosa de las dimensiones estudiadas por Sager (1990), la consecuencia de esto es el estudio de la terminología en textos. Los textos suponen una situación comunicativa que transmite información sobre una disciplina entre dos especialistas que forman parte de un grupo humano reducido y relativamente homogéneo.

La selección del lenguaje tiene una serie de convenciones lingüísticas resultado del acuerdo previo, al restringir la selección en variables fonológicas (neutralizadas al tratarse generalmente de textos escritos) y sintácticas (en muchos casos, suelen eliminar algunas formas temporales y personales). Para ello, se hará uso de un juego compartido de términos que permiten señalar una zona precisa del conocimiento pretendido por el mensaje. Aparte del elemento léxico, es posible estandarizar las estrategias de



codificación de manera que se puedan crear mecanismos de comprensión para nuevos términos derivados.

Los requisitos para una comunicación especializada exitosa se dan a partir de dos situaciones que claramente diferencian los textos especializados del lenguaje general: a) usan tipos especiales de texto (p. ej., artículos científicos o cuadernos de laboratorio), b) usan términos además de palabras, p. ej., formas de expresión lingüística que hacen referencias especializadas distintas de la referencia general de las palabras. Los términos deben aprenderse de manera separada a las palabras incluso si comparten la misma forma expresiva, ya que por lo general las definiciones incluyen matices, delimitaciones y especificaciones técnicas precisas. La eficacia de los lenguajes especializados responde, según al autor, a tres principios básicos de economía, precisión y adecuación que en gran medida se hacen eco de las máximas de Grice y que, aunque más exhaustivos en su descripción, consideramos que realmente pueden aplicarse a cualquier acto discursivo más allá del especializado.

Para concluir, hay que destacar el carácter práctico y, a la vez, visionario de la obra de Sager. La teoría de la terminología sería, en resumidas cuentas, parte de una teoría de la lexicología. La teoría terminológica tradicional identifica su enfoque como *onomasiológico*, es decir, asigna y regula los nombres de los conceptos nuevos evitando, al menos sobre el papel, la homonimia. Sin embargo, aunque existe una terminología planificada y normalizada, los casos en los que el terminólogo parte de cero son poco frecuentes. La teoría de Sager se aproxima al enfoque basado en corpus para la recogida de información, naturalizándose la concepción de los términos, que ya no son vistos como lenguajes semiartificiales sin las funciones de otros elementos léxicos. Esta nueva actitud abre paso a nuevos campos de estudios cognitivos que relacionan significado y léxico especializado, como es el caso de los métodos y enfoques de la lexicosemántica.

Sager, ya en 1990, no solo había previsto sus múltiples aplicaciones prácticas, sino que puso sobre la mesa el carácter cerrado de los desarrollos terminológicos de la época. A nuestro conocer, por primera vez, pone en valor la labor dual de los terminólogos, no solo como observadores y consejeros en el uso de terminología, sino como productores de soluciones aplicadas que trascienden la cuestión meramente técnica, sino a sus implementaciones en la vida cotidiana. Entre ellas podemos encontrar muchas tecnologías de procesamiento del lenguaje natural que, casi treinta años después, han tenido gran relevancia tanto en la investigación como en sus aplicaciones comerciales: la traducción automática, la minería de datos y los sistemas basados en conocimiento.

Proponía el desarrollo de lexicones especializados, diseñados ad-hoc mediante la reutilización de recursos léxicos para aplicaciones de procesamiento de lenguaje natural, lo cual corresponde en gran medida con el trabajo aquí expuesto. Finalmente, ya evidenció la necesidad de unir fuerzas y llegar a acuerdos de registro en los bancos de datos, en cierto modo previendo la necesidad de usar estándares abiertos que faciliten la cooperación y el intercambio científico.

#### **4.3.3 La teoría comunicativa de la terminología de Teresa Cabré**

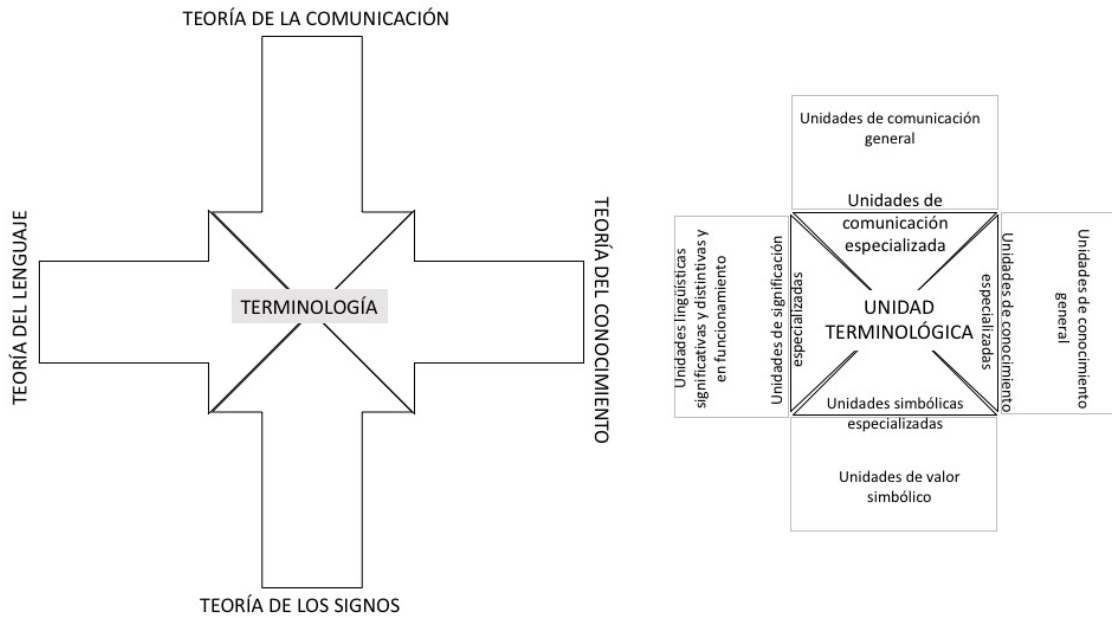
La teoría comunicativa de la terminología (TCT) propone una teoría de la terminología desde una visión eminentemente lingüística, de manera que se añade una visión social, y cognitiva al uso del lenguaje especializado (Cabré, 1998; 1999; 2000; 2003). Teniendo en cuenta el carácter especializado del conocimiento y la necesidad de su transmisión, existen varias características en la terminología que coinciden con las observadas en otras situaciones comunicativas donde se usa el lenguaje. Se considera aquí que la comunicación especializada no se desarrolla independientemente de la comunicación general; ni el conocimiento especializado es uniforme ni está separado del conocimiento general. Una teoría de la terminología debe proporcionar un marco metodológico para el estudio de unidades terminológicas. Por lo tanto, la terminología no puede estudiarse “al margen de otros signos denominativos dotados de capacidad referencial y denominativa” (Cabré, 2002), ni sin apoyarse en las teorías comunicativas y cognitivas. El objeto de la terminología son las unidades terminológicas, que tienen carácter lingüístico y se dan en el seno de una lengua natural. Esta consideración de la terminología como un fenómeno lingüístico abrió las puertas a diversos y novedosos campos de investigación terminológica: relaciones conceptuales, variación terminológica, extracción terminológica y la aplicación de distintos modelos lingüísticos en la terminología (Faber, 2009).

Las unidades terminológicas serían al mismo tiempo iguales y diferentes a las palabras de una lengua, por lo que el modo de significación y sus aspectos pragmáticos serían los elementos determinantes en la especificidad de cada término. Cabré (2003) considera las unidades terminológicas como conjuntos de condiciones derivados del contexto especializado, su área de conocimiento, su marco conceptual, significado, estructura léxica, estructura sintáctica y valencia. La especificidad del lenguaje reside, entonces, en una serie de restricciones cognitivas, sintácticas y pragmáticas, que afirman

su pertenencia a un dominio especializado (Faber y López-Rodríguez, 2012). Y finalmente, el significado de un término es considerado aquí el resultado de una negociación entre especialistas dentro del discurso a partir de la realización de predicaciones que determinan el significado de cada unidad.

Cabré (2002) propone enfocar la terminología a partir del “tratamiento multidimensional de los términos”, a lo que denominó “teoría de las puertas” (Figura 7). A partir de ella, se da la representación de posibles formas de acceder, analizar y comprender las unidades terminológicas. La unidad terminológica tendría la forma de un poliedro con un número variable de facetas. Según esto, una unidad terminológica tiene tres dimensiones: una dimensión cognitiva, una dimensión lingüística y una dimensión comunicativa. A partir de aquí, la unidad terminológica estaría en una sala a la que se accedería al abrir metafóricamente la puerta de una teoría: la teoría de los signos, la teoría del lenguaje o la teoría de la comunicación. Sin embargo, la elección de la puerta (o enfoque) no implica un rechazo de las otras dos perspectivas, que continúan residiendo en el fondo. Según la autora, la teoría comunicativa de la terminología estudia a las unidades a través de la puerta del lenguaje, pero siempre dentro del contexto general de la comunicación especializada.

Esta teoría supera el estatismo terminológico de paradigmas previos y considera a las unidades terminológicas como elementos de carácter moldeable, ya que estos pueden ser reciclados por otros dominios específicos de una forma dinámica y recursiva. Esta capacidad explica la movilidad en los ejes cartesianos propuestos por Hoffmann (1998): las unidades del léxico común pueden desplazarse hacia el léxico especializado, e incluso desde una especialidad a otra, encuadradas dentro de un marco integral de la realidad que permita variación conceptual y denominativa. Al contrario de la teoría general de la terminología de Wüster, cuyo principio de univocidad considera a este fenómeno como la existencia de varios términos homónimos, la autora considera que se trata de un solo término polisémico, es decir, que comparten el mismo origen y su significado parte de una variación en un significado anterior (Cabré, 2002).



**Figura 7** Esquema de la teoría de las puertas (Cabré, 2002)

En cuanto a sus características, se considera que las unidades terminológicas tienen las mismas propiedades morfosintácticas que las palabras. La diferencia recae en la forma en la que los términos se generan y sus modos de significación. Los términos toman como forma de base a las unidades léxicas conocidas por el hablante, que evidentemente incluyen múltiples valores cognitivos, lingüísticos (semánticos, léxicos y textuales) y pragmáticos (emisor directo o a través de un mediador lingüístico o cognitivo, destinatario y situaciones) (Cabré, 2003). Los términos son, entonces, conjuntos de condiciones derivados de estos valores cognitivos, lingüísticos y pragmáticos relacionados con los lenguajes específicos. Estas unidades terminológicas pueden (y suelen) tener una unidad normalizada utilizada en las clasificaciones formales entre sus características. Cabré (2002) razona que la consideración de una unidad léxica como *terminológica* o *especializada* no es sino uno más de sus valores pragmáticos, por lo que la tarea principal de una teoría de la terminología será la descripción formal, semántica y funcional de las unidades con dicho valor junto con las relaciones con otros términos.

A partir de esta serie de principios, se abre un amplio campo para las aplicaciones de la terminología, que va más allá de la mera descripción de las relaciones entre conceptos: desde la enseñanza de lenguas para fines específicos, la traducción asistida, el estudio de las relaciones entre profesionales al establecimiento de cánones técnicos. Cabré (2000) considera que todas las unidades terminológicas cumplen con la aplicación de la doble función de los términos: la representación del conocimiento especializado y la de transferencia de ese conocimiento.



Para aproximarse a la terminología desde la lingüística, y al igual que Sager (1990), propone realizar la misma recogida de datos que para cualquier estudio de la lengua: recopilar textos orales y escritos por los distintos actores lingüísticos. Estos tienen una estructura cognitiva organizada a partir de la gramática y compuesta por unidades de conocimiento ordenada como nodos en red. Será entonces el contexto y sus condiciones de uso el factor que materializará el significado especializado de las unidades y recibirán su valor terminológico cuando se utilicen dentro un contexto especializado. Dichas unidades de significación especializada pueden ser de dos tipos: las unidades lingüísticas pertenecientes a la lengua natural materializadas en los distintos niveles de la gramática y las unidades no lingüísticas pertenecientes a sistemas simbólicos no naturales. Para reflejar el carácter poliédrico de la terminología, además de esta descripción lingüística de tipo eminentemente gramatical, el análisis de las unidades terminológicas en este nivel debe incluir no solo reglas y condiciones formales, sino también elementos semánticos cognitivos y pragmáticos. Habiendo dado el salto de una terminología esencialmente conceptual a una perspectiva con base lingüística, esta aproximación no debe estar encorsetada en una rígida observación estructural ya que deberá tener en cuenta el hecho de que las unidades terminológicas son también entidades de comunicación y transmisión del conocimiento especializado (Cabré, 2002).

Si bien considera que es la teoría mejor posicionada para reemplazar la teoría general de la terminología, Faber (2009) critica la vaguedad de algunos conceptos y el hecho de que no proponga ningún modelo lingüístico para su análisis en profundidad. Asimismo, critica la falta de profundidad en torno a la semántica conceptual, ya que considera que se trata de una manera demasiado superficial. Aunque la TCT basa su semántica en representaciones conceptuales, no explica claramente la forma en que se crean estas representaciones, ni las describe, ni delimita sus restricciones. Si la TCT considerase que los términos tienen un significado especial dentro de una estructura, las representaciones conceptuales o las ontologías tendrían un peso mayor dentro esta la teoría, sin ser este el caso.

Para concluir, una segunda crítica de Faber (2009) es en torno a la profundidad del estudio del significado semántico en la TCT. Como se ha visto en este apartado, aquí se considera que el significado especializado se adquiere al aparecer el término en un contexto discursivo de dominio específico y que este es un añadido a la información léxica de manera que no se responde sobre la naturaleza semántica de este significado y sus componentes. De esta forma, la definición se deja a merced de la semántica léxica,

donde existen amplios disensos sobre la naturaleza del significado de las palabras y su análisis.

#### 4.3.4 La teoría sociocognitiva de la terminología de Rita Temmerman

La teoría sociocognitiva de la terminología surge como planteamiento alternativo a la terminología tradicional y parte de la obra fundacional *New Ways of Terminology Description* elaborada por Rita Temmerman en el año 2000. Propone una observación empírica de la realidad terminológica dentro del contexto de los textos especializados y mueve el foco del estudio del concepto como entidad abstracta a la terminología realizada a través del uso comunicativo mediante el estudio de los textos.

El desarrollo de la teoría de la terminología sociocognitiva sienta los cimientos de una nueva teoría de la terminología y desarrolla una metodología que plantea las herramientas y técnicas a utilizar en el análisis y estructuración de los datos de trabajo en la descripción terminográfica. Para ello, Temmerman (2000) se basa en una nueva conceptualización flexible con principios de semántica cognitiva (p. ej., Filmore, 1985) basada en la prototipicidad y en la categorización con límites difusos. La teoría sociocognitiva corre paralela a los desarrollos de esta escuela y, con diferencias, similar a la teoría comunicativa de la terminología (Cabré, 2003). A pesar de que en ambos el centro de estudio son los términos a partir de los textos, su descripción aquí se lleva a cabo a partir de la descripción de los conceptos en relación con los marcos semánticos. Los modelos metafóricos de definición, rechazados en su totalidad por el enfoque tradicional, son puestos en valor para facilitar el pensamiento y expandir la comprensión del mundo a través del lenguaje.

En cuanto a las bases humanísticas, parte desde una crítica a la visión estructuralista de la lengua, en la que considera que la teoría general de la terminología construye un metalenguaje abstracto y sin salida, propone una visión alineada con la hermenéutica y la posmodernidad. Con una visión alineada con la hermenéutica y el pensamiento deconstruccionista de Derrida (1976), rompe con la visión del concepto como ente estable, universal y atemporal y propone la deconstrucción como estrategia para revelar las capas del significado *dentro* del texto en su forma real (Montero Martínez, 2002). Según su óptica, los principios de la Escuela de Viena son idealistas, superficiales y poco prácticos. Critica la excesiva concentración en la normalización de los términos y la carencia de descripciones fidedignas del significado que parten desde una perspectiva

objetivista del mundo. Por ello, propone superar el *concepto* como foco de investigación al considerarse el significado como un hecho objetivo para así estudiar la terminología desde lo tangible, desde los textos. La terminología necesita la información que emerge de la comprensión tanto de la realidad extralingüística como del uso de los elementos léxicos.

En cuanto a la metodología, dependiendo de la naturaleza de la problemática se proponen distintas estrategias para extraer la información relevante. Generalmente, hace uso del análisis textual a partir de documentos especializados (artículos, libros, actas de congresos), así como textos lexicográficos (diccionarios, tesauros) y documentos de divulgación escritos por especialistas en una amplia gama de formatos y registros que pueden ser recogidos en corpus textuales.

Temmerman (2000) parte de los tradicionales principios wusterianos vistos en esta sección para reemplazarlos por una clasificación de cinco nuevos principios. El primero de los principios sería *las unidades de comprensión (units of understanding)*, ya que considera muy restrictiva la consideración del *concepto* como núcleo de la terminología pues la autora considera que pocos conceptos existen de una manera objetiva. El hecho diferenciador de la teoría sociocognitiva es el énfasis en una perspectiva próxima a la semántica cognitiva (Lakoff, 1988), se aproxima a la realidad a partir de los *Modelos Cognitivos Idealizados (Idealized Cognitive Models)* y propone sustituir esta noción por las *unidades de comprensión*, estructurables mediante prototipos categorizados. Este modelo de categorización se basa en la teoría de prototipos de Rosch (1978). Se vuelve aquí a la discusión sobre el punto de partida de las categorías y cuestiona si se considera un enfoque semasiológico que parta del término o un enfoque onomasiológico que tome como principio la abstracta idea de unidad de conocimiento. En contraste con la visión monolítica tradicional, Temmerman propone describir la información según tres perspectivas según las necesidades y la labor de la terminología sería tratar de describir la estrecha relación entre ellas. Las perspectivas serían las siguientes:

1. *Nominalista*: en ella la *unidad de comprensión* es el sentido del mundo.
2. *Mentalista*: cuya *unidad de comprensión* es una idea mental.
3. *Realista*: la *unidad de comprensión* es una forma externa que existe en el universo.



La terminología sociocognitiva reivindica el papel esencial que tienen las denominaciones en la comunicación y categorización al ser la transmisión comunicativa del conocimiento la principal función de la terminología. Por ello, sólo se puede estudiar a través del discurso contextualizado. Así, el *término* pasa a ser la unidad central de descripción terminológica en detrimento del *concepto*. En cuanto al significado, las características semánticas de vaguedad, polisemia y sinónima aquí serían aceptadas con naturalidad al considerar que los límites entre *unidades de comprensión* no están delimitados de una manera estricta, sino que son difusos.

El segundo principio propone que *la comprensión organiza los modelos cognitivos (understanding is sorting cognitive models)*: la terminología tradicional considera que un término solo encaja en la estructura conceptual si su definición tiene unos límites bien definidos ya sea de una manera ontológica ( $x$  es parte de  $y$ ) o lógica ( $x$  es un tipo de  $y$ ). En cambio, la teoría sociocognitiva considera que cada categoría se comprende como existente dentro de un modelo cognitivo y que la comprensión es *per se* un hecho clasificatorio. Temmerman (2000) considera que hay dos aspectos que deben tenerse en cuenta:

- (a) *Estructuras intracategoriales y estructuras intercategoriales*: en la terminología tradicional solo las características distintivas se consideraban importantes dentro de una categoría y fuera de ellas; se tenían en cuenta las categóricas lógicas y ontológicas. En la terminología sociocognitiva la existencia de una categoría no se considera independiente del lenguaje y es necesario conocer los distintos módulos de información (características, núcleo de definición, información histórica, información sobre los procedimientos...) ya que serán más o menos importantes según la naturaleza del término (entidad, actividad, término genérico...). Por ende, el término debe observarse en el discurso e implica más módulos de información que las características distintivas y la posición dentro de la clasificación lógica.
- (b) *Las categorías tienen estructura prototípica*: estas se estructuran independientemente de la lengua y en gran medida a partir de prototipos. La observación de las categorías se hace a partir del texto y es parte del proceso de comprensión de un término. En cuanto a la relación intensional, pocos términos tienen características claramente delimitantes y los límites entre categorías suelen solaparse de manera natural. Con respecto a la



extensión de los términos, los ejemplares dentro de una categoría muestran parecidos dentro de la estructura de la familia y cada uno de los términos miembros tienen distintos grados de pertenencia a la categoría.

El tercer principio es la *representación de las plantillas de definición (template representation)*: el enfoque tradicional estima que la definición del concepto debe basarse en características que cumplan condiciones suficientes y necesarias. En cambio, el enfoque sociocognitivo considera que la definición dependerá del tipo de unidad de comprensión, del nivel y tipo de especialización del emisor y del receptor y, finalmente, de la relevancia de la información en cuanto a su definición. La teoría propone el diseño de patrones de comprensión compuestos por diferentes módulos de información que serán esenciales dependiendo de diversos factores como la perspectiva desde la que se entiende determinada unidad de comprensión. Estos tipos de información pueden ser, principalmente, información histórica, pasos a seguir en el proceso, o ser atributivas como objetivos, aplicación y resultados (Montero Martínez, 2002).

El cuarto principio, es la *funcionalidad de la sinonimia y la polisemia (functionality of synonymy and polysemy)*: mediante la observación de los textos se determina que existe poca arbitrariedad en la categorización y la lexicalización de manera que la sinonimia y la polisemia son frecuentes en el proceso de comprensión y referencian estructuraciones distintas del mismo término.

La polisemia existe principalmente por tres factores. El primero hace referencia a cambios en el mundo o en el desarrollo de la tecnología, por ejemplo, el término “impresión” hace referencia tanto a las técnicas inventadas por Gutenberg como a la de las modernas impresoras láser. El segundo hace cambios en la comprensión de la categoría desde un punto de vista cognitivo, por ejemplo, a partir de los avances científicos. Finalmente, el tercer factor tiene en cuenta a los elementos de cambio en el lenguaje que acepta, por un lado, que la estructura prototípica de las categorías aporta conocimiento adicional de la evolución del significado y que los elementos en la lengua se influyen entre sí (Temmerman, 2000).

Por último, el quinto principio, *los modelos cognitivos están en constante transición (cognitive models are constantly in transition)*: este rompe con la idea de que los conceptos están cristalizados en un determinado espacio de tiempo. El continuo desarrollo de las unidades de comprensión se puede comprender a partir de varios factores que ocurren de forma simultánea y son influyentes en el ordenamiento de los modelos.

Estos serían tres: (a) la mayor necesidad de una comprensión mayor y de mejor calidad; (b) el intercambio e interacción entre hablantes; (c) la estructura prototípica dentro de la comprensión de categorías que es, a la vez, causa y resultado de la evolución en el significado. La motivación de la lexicalización se comprende a partir de se explica a partir de los modelos cognitivos como los ICM que tienen un papel importante en el desarrollo de nuevas ideas.

La teoría sociocognitiva de Temmerman supone un giro semántico de carácter empírico a la concepción estructuralista de la terminología. El enfoque sociocognitivo pone en un primer plano a los términos en contexto a la hora de trabajar con sistemas de lenguajes especializados. Este sistema propone dejar atrás la preestructuración lógica, normalizada y cristalizada de la terminología para poder estudiar los términos como entes vivos que son modelados por el uso y el acuerdo de los especialistas y que varían a lo largo del tiempo a partir del desarrollo del conocimiento humano y las relaciones entre usuarios del lenguaje especializado.

El modelo evidencia la escasa correspondencia con el uso real del esquema jerárquico del modelo conceptual y describe la topología de la clasificación de la terminología de manera más parecida a una red distribuida de prototipos relacionados que se articulan de manera asimétrica conforme a los límites difusos de los términos (intensión) y se extiende de manera rizomática generando nuevos términos según la carga cognitiva y las necesidades de expresión. Por lo tanto, se produce un giro en el objeto el estudio, el concepto como ideal abstracto e invariable pasa a convertirse en *unidad de comprensión* que introduce el uso de prototipos y refleja el conocimiento y sus continuas reformulaciones observables a través de los textos.

Desde un punto de vista crítico, Faber (2009) cuestiona la viabilidad del uso de los prototipos para el estudio de la terminología, al considerar imposible definir la naturaleza exacta de la categoría prototípica, o poder medir objetivamente los grados de prototipicidad. Al igual que con Cabré, cuestiona la cuidadosa manera en la que esta teoría evita el tratamiento de la sintaxis, debido a que este análisis sintáctico de la terminología debe basarse en una teoría gramatical y pone sobre la mesa la escasez de estudios que trabajen sobre la relación entre ambas disciplinas.

#### 4.4 Los términos: denominaciones y formas lingüísticas

La observación de la terminología puede llevarse a cabo desde dos perspectivas: *in vitro*, esto es, fuera de su uso corriente, a partir de los distintos productos terminográficos (lexicones, bases de datos o tesauros); o *in vivo*, es decir, a través de la observación de los términos en uso mediante la observación de los textos. Las corrientes modernas de la terminología se benefician de las posibilidades que ofrece el análisis computacional de grandes cantidades de textos. Como se ha visto en el apartado anterior, se pueden generar lexicones con rapidez y precisión para posteriormente organizarlos en ontologías o glosarios. Esto ayuda a realizar un relato conciso de las técnicas de nombrado, clasificación por significado y representación del conocimiento. Los datos estadísticos que nos ofrece la extracción terminológica aplicada a corpus especializados se utilizan de manera generalizada tanto en el entorno académico como profesional y se ha convertido en una herramienta esencial a la hora de realizar recopilaciones y análisis terminológicos (Santamaría Pérez, 2006).

Mientras que las palabras tienen un origen no normalizado, el origen de los términos es fácilmente trazable y, en muchos casos, planificado. Por ejemplo, la terminología botánica usa nomenclaturas con un alto nivel de regularidad y permite analizar las reglas de nombrado que se aplican a una disciplina en concreto, conocer sus causas y establecer reglas con el fin de prever la designación de términos relacionados futuros. Una terminología organizada ontológicamente y formulada a partir de patrones predecibles sería de gran ayuda a la hora de automatizar los procesos de detección, clasificación y extracción terminológica.

Los términos, en sus distintas formas, son una parte frecuente en el lenguaje natural y en la estructura de una lengua. Su trasfondo funcional no parte de un sistema léxico diferente, sino que, entre sus propiedades léxicas, activan un significado especializado en determinados contextos. Aunque la naturaleza de la terminología es sobre el papel planificada, no es sorprendente reseñar que en múltiples ocasiones emergen términos de manera disruptiva que rompen con la regularidad de una terminología aséptica y organizada. Esto se debe a que en muchos casos se producen por necesidad de sus usuarios. Fenómenos de formación de palabras como la derivación a partir de afijos (“up-”, “down-”), la acronimia (“IBEX”, “NASDAQ”), el argot (“gnomes of Zurich”, “dead cat bounce”), la metáfora (“bull market”, “bear market”); o elementos no lingüísticos como símbolos (“€”, “¢”) o fórmulas (“H<sub>2</sub>O”) son ejemplos de los distintos

tipos de términos en la terminología financiera. Este es el caso de la dupla “Bull market” y “Bear market”. El primero hace referencia a un mercado donde generalmente se anticipa que los precios de los valores aumentarán, mientras que el segundo hace referencia a un mercado en el que los precios bajan drásticamente en un contexto de pesimismo generalizado. Aunque los orígenes de estos términos no están bien documentados, los términos oso (“bear”) y toro (“bull”) se consideran derivados de la forma en que cada animal ataca a sus oponentes. Es decir, un toro empujará sus cuernos hacia arriba en el aire, mientras que un oso empujará a su presa hacia abajo. Estas acciones se relacionaron metafóricamente con el movimiento de un mercado: si la tendencia era ascendente, se consideraba un mercado alcista; si la tendencia era descendente, era un mercado bajista (Sager, 1990).

A la hora de designar los términos es necesario observar los conceptos con los que se opera y las designaciones requeridas. Las denominaciones se utilizan para referenciar tanto objetos materiales (“coin”) como objetos abstractos (“currency”). La metáfora es uno de los fenómenos más recurrentes en la elaboración de términos y sirve de puente entre el mundo material y el mundo abstracto, es frecuente observar nombres en relación metafórica con sus análogos cotidianos como “green light” o “balloon”. En los dos ejemplos tenemos una representación mental abstracta vinculada al término a partir de una generalización basada en la experiencia. De esta manera, los términos “coin” y “currency” no hacen referencia a una moneda o a una divisa definida, sino que la experiencia cognitiva de la existencia de una gran serie de objetos que comparten una serie de características nos lleva a sintetizar estos conceptos. Por otro lado, los términos suelen agrupar objetos similares de manera genérica y diferenciadora, de manera que engloban a muchos similares, pero no representan todas y cada una de las clases en el global de una ontología, como es el caso de “book” en el dominio de la biblioteconomía, que, aunque es una subclase de documento, se utiliza como elemento diferenciador de cualquier otro documento. Finalmente, la elección de los términos es, generalmente, consistente en la elección de la técnica de denominación aplicada. Por ejemplo, en inglés se utilizan sustantivos deverbales para hacer referencia a los procesos y métodos, como es el caso de “banking”, “slicing”, “consumption”, etc.

Otro fenómeno es la utilización de términos compuestos, en los cuales la utilización de una, dos o tres palabras indica una jerarquía dentro de la ontología p. ej., los elementos mutuamente excluyentes “fixed interest rate” o “variable interest rate” serían subcategorías de una jerarquía encabezada por “rate”. Otros términos se forman a

partir de composición mediante la identificación de la función particular de su objeto, por ejemplo: “capital control” o “capital cushion”. Sin embargo, la posibilidad de predicción que ofrecen los patrones terminológicos, especialmente en el primer caso, suelen ser la excepción más que la regla y suelen utilizarse convenciones terminológicas de acuerdo con la costumbre del dominio específico (Sager, 1990).

Rossi (2015) considera la metáfora como un laboratorio privilegiado para observar las dinámicas lingüísticas y sociales dentro de la elaboración de nuevas designaciones dentro de la denominación de los conceptos. Propone tres tipos de metáforas dentro de la terminología:

- La *metáfora de invención*, que asocia áreas remotas para crear un nuevo paradigma epistemológico. Por ejemplo, la economía con la imagen de la *invisible hand* que propone una especie de Gaia ordenadora dentro de la economía de mercado que la autorregula.
- La *metáfora de filiación cognitiva*, que se manifiesta bajo la forma de una red conceptual que es estable al ser compartida por una comunidad discursiva y tiene un carácter particularmente productivo. Por ejemplo, la relación entre la economía y la física (“balance”, “elasticity”, “liquidity”, etc.)
- La *catacrosis denominacional*, que se basa en una analogía formal, y es muy común en los vocabularios técnicos, pero también en los lenguajes científicos, como los de la medicina o la biología. Es principalmente de origen antropomórfico o zoomorfo de carácter aislado, por lo general, y no da lugar a una red ontológica. Es el caso por ejemplo de los “PIGS”, que juega mediante la acronimia con “pig” y hace referencia a las economías engordadas del sur de Europa o los ejemplos de “bull” y “bear” explicados anteriormente.

Para concluir, realizaremos una breve clasificación de los distintos tipos de unidades terminológicas desde el punto de vista de la forma a partir de ejemplos del dominio económico (Santamaría Pérez, 2009):

- a) *Número de morfemas*: los términos pueden ser de carácter simple (“stimulus”) o complejo (“stimulatory”).
- b) *Tipo de morfemas*: los términos pueden ser derivados por afijación o compuestos.

prefijación	sufijación	composición
“HYPERinflation”	“interventionISM”	“takeover”

- c) *Sintagmas terminológicos*: términos formados por la combinación de palabras a partir de una estructura sintáctica determinada. N-N (“gold standard”, “interdealer broker”) o Adj N (“high street”, “hiking interest”, “junk bond”).
- d) Términos de origen complejo, principalmente:
- *Siglas*: combinación de la letra inicial de un término formado por varias palabras. P.e.: “ECB” (“European Central Bank”) o “PIGS” (despectivamente, el conjunto de los países del sur de la Zona Euro: Portugal, Italy, Greece, Spain).
  - *Acrónimos*: combinación de segmentos (generalmente dos) de un sintagma desarrollado. P.e.: “brexit” por “British EXIT (of the European Union)” o complejas como Euribor, combinación de “European” + el acrónimo lexicalizado “IBOR” (“Interbank Offered Rate”).
  - *Abreviaturas*: formas, normalmente fijadas ad hoc, que reproducen el segmento inicial de una palabra. P.e.: “Q1” (“First Quarter”), “pct” (“percent”), “avg” (“average”).
  - *Abreviaciones*: son unidades utilizadas por economía en el discurso. Ej.: “div net” por “dividend net”.

## 4.5 La gestión terminológica

### 4.5.1 Enfoque textual de la terminología

La *terminografía*, también conocida como *procesamiento terminológico* (*terminology processing*) (Wright y Budin, 1997) o *gestión terminológica* (*terminology management*) (Sager, 1990) es la disciplina que se encarga de estructurar sistemas conceptuales especializados y recopilar términos de manera estructurada basándose, al menos parcialmente, en los métodos ya existentes en la lexicografía. Entre ellos está la recolección de materiales que documenten las unidades descritas, la redacción de definiciones, la detección de equivalentes lingüísticos o los distintos niveles de uso (Montero Martínez, 2002).

Cabe aclarar, entonces, las diferencias entre lexicografía y terminografía. Cabré (1995) considera que estas se distinguen, en primer lugar, por el método de trabajo que usan, principalmente en dos vertientes: el proceso de trabajo y la orientación de este. El proceso de trabajo de la terminología busca, al menos parcialmente, ofrecer

denominaciones para los distintos elementos de la estructura conceptual. En cambio, la lexicología trabaja esencialmente desde la validación de las hipótesis a través del análisis de las muestras textuales. Al respecto de la orientación del proceso de trabajo, este es inverso entre las dos disciplinas, el proceso lexicográfico suele ser semasiológico y, en cambio, la terminología suele desarrollarse desde una perspectiva onomasiológica. Es decir, el punto de partida de la lexicografía para la redacción de diccionarios es una lista de lexemas que son definidos semánticamente, mientras que en terminología se establece primero a lista estructurada de conceptos de una materia para, subsiguientemente, asignar una denominación a cada uno para el uso efectivo del hablante especializado, de manera que la comunicación entre expertos sea precisa, actualizada y unívoca. La última diferencia es el tratamiento de la definición, mientras que en los diccionarios generales se suele recurrir a la polisemia, de manera que la presentación de los artículos es mucho más sistemática, la definición terminográfica es de un carácter descriptivo más amplio. Además, la organización de las entradas es generalmente de carácter sistemático en los diccionarios profesionales. Aquí cabe destacar al *Center for Leksikografi* en Dinamarca, cuyo desarrollo teórico ha sido conocido internacionalmente como Escuela de Aarhus, (Bergenholtz, Nielsen y Tarp, 2009). Sus estudios se enfocan en las funciones de los diccionarios a la hora de proporcionar información y conocimiento, es decir, funciones dependientes del texto relacionadas con la comunicación (p. ej., recepción de texto, producción de texto, revisión de texto, edición de texto y traducción) y funciones cognitivas (como obtener conocimiento en general o sobre un tema específico) con el objetivo de proponer soluciones a los problemas que se plantean en la lexicografía especializada, la lexicografía didáctica y las políticas lingüísticas.

La terminografía prescriptiva o normalizadora (organismos estatales como AENOR o internacionales, ISO) tiene un papel primordial en el mercado para denominar productos de una manera homogénea y unificadora. Esta regularización terminológica debe seguir los dictados del uso de los especialistas, que deben recoger el uso de manera descriptiva y exhaustiva procurando, al menos sobre el papel, la habilitación de canales de comunicación fluidos entre todos los actores: expertos, terminólogos y agencias normalizadoras. Ya que esta no es sino una propuesta idealizada, el trabajo de la terminología descriptivista no pasa por la normalización, sino por recoger la terminología viva, es decir, la documentación a partir del discurso de todos los términos usados para denominar el marco conceptual de una rama del conocimiento. Dentro de esta parte de la terminografía, además de los enfoques teóricos, se debe tener en cuenta también el trabajo



terminográfico práctico, tradicionalmente llevado a cabo manera sistemática por científicos o traductores, cuyo fruto son los diccionarios, glosarios o bases de datos terminológicas. El resultado final de un trabajo terminológico *in vivo* será una amplia recogida de unidades de conocimiento con distintos grados de lexicalización: unidades terminológicas, fraseológicas y contextuales. Estas desplegarán un margen de variación formal significativo y es pertinente a los usuarios finales la elección de las unidades según parámetros de precisión y adecuación al discurso especializado (Montero Martínez, 2002).

La práctica terminológica contemporánea se caracterizará, pues, por la clasificación, gestión y procesamiento de grandes cantidades de datos textuales en formatos electrónicos. La implementación de herramientas para ayudar a construir terminologías a partir de textos es una de las áreas más productivas del procesamiento del lenguaje natural. Debido a esta influencia, la terminografía ha sufrido un cambio radical y, como se ha visto anteriormente, supone un nuevo desafío para los fundamentos teóricos de la terminología ya que deben entenderse desde las nuevas necesidades de aplicación y desde el desarrollo de novedosos objetivos de estudio (Campo, 2012).

Uno de los ejes con mayor número de estudios es el relacionado con la construcción de terminologías a partir de corpus, dada por las necesidades comerciales, ya sea la gestión de enormes tomos de documentación industrial o el estudio de volúmenes de información procedentes de usuarios en la web. Como se expone en la introducción a este capítulo, los desarrollos técnicos impulsados tanto por la informatización de las distintas disciplinas como el acceso a grandes cantidades de información en Internet han aportado nuevas perspectivas al desarrollo de la teoría y la práctica, en forma de nuevas visiones que acompañan a productos avanzados que hacen uso de la terminología. Este sería el caso de nuevas disciplinas como la minería de datos, la extracción de la información o el análisis de sentimiento.

Los desarrollos en lingüística de corpus y lingüística computacional han permitido nuevas perspectivas en la gestión terminológica mediante trabajos desarrollados a partir de enfoques basados en corpus. Los corpus, que han crecido exponencialmente en esta década, permiten a los terminólogos recopilar grandes cantidades de material especializado. Tiene grandes ventajas ya que reduce, por un lado, el tedioso trabajo de procesado, mejora con creces las posibilidades de gestión terminológica y ofrece la posibilidad de adaptar sus aplicaciones a las necesidades específicas. Asimismo, el trabajo con textos permite observar la terminología en uso, permitiendo al observador inferir las



variaciones producidas por el uso en determinados contextos (Aussenac-Gilles y Condamines, 2007).

Dentro de esta nueva línea de posibilidades, Bourigault y Slodzian (1999, citado por Campo, 2012) proponen una terminología textual. La información esencial para la construcción de una terminología (tanto los términos como su esquema relacional) se puede obtener a partir del procesado de corpus textuales especializados que tengan en cuenta los criterios de selección y representatividad. Para la construcción de un corpus especializado, Pérez Hernández (2002) propone tener en cuenta varios criterios para la recopilación de los textos, como se verá más adelante. Este corpus constituirá la fuente de observación y la descripción de los términos de manera que es posible establecer distintas técnicas para extraer información formal y contextual: p. ej., polisemia, neología, distintos patrones de sintaxis o aspectos semánticos. Estos criterios serían los siguientes:

1. *criterios externos*: son eminentemente extralingüísticos y determinan la modalidad, género, finalidad y origen de los textos con vistas a cubrir de manera representativa una gama de contextos situacionales.
2. *criterios internos*: identifican tipos diferentes de textos, de acuerdo con una serie de categorías lingüísticas. En el caso de los corpus de propósito general suelen centrarse en la distribución de pronombres, proposiciones o tiempos verbales.

Los corpus especializados deben tener una serie de criterios externos que tengan en cuenta las situaciones comunicativas entre los especialistas, de manera que se amolden al contexto lingüístico que se quiere estudiar. Pérez Hernández (2002) expone cuatro categorías que diferencian distintos niveles de complejidad en la situación textual comunicativa:

1. *Comunicación entre especialistas*: con una organización textual y vocabulario altamente especializado y preciso.
2. *Comunicación especialista-iniciado*: incluye un texto de más explicitado y un léxico especializado.
3. *Comunicación entre (semi-)especialista y lego*: de índole más accesible y un léxico parcialmente menos especializado.
4. *Comunicación profesor-alumno*: que tendrá un contenido y léxico especializado y un carácter pedagógico, con ilustraciones y ejemplos de los distintos conceptos y términos utilizados.

Una vez diseñado y recopilado el corpus, Bourigault y Slodzian (1999) realizan una propuesta de trabajo que permite la recopilación de terminologías en uso para, de esta manera, poder crear los distintos productos terminológicos: (a) exploración de corpus para adquisición terminológica, (b) exploración de corpus para encontrar relaciones entre unidades terminológicas, (c) con la asistencia de un experto en el tema, validación de términos candidatos y sus relaciones y, dadas las necesidades, (d) exploración de corpus (paralelos o no) para detectar equivalencias multilingües en los idiomas de destino y validación de estas propuestas por los expertos.

Es importante reseñar la labor cualitativa del tercer paso, pues aunque actualmente las distintas aplicaciones informáticas han simplificado la tarea de selección y clasificación terminológica y su posterior análisis, es necesario realizar una tarea precisa de verificación manual. Determinar la condición de término de una palabra es una labor compleja, ya que la condición de término es una cuestión gradual y no todos los términos están al mismo nivel (Alonso Campo y Renau Araque, 2013). Por esto, el siguiente paso cuenta con un grado variable de intervención humana, en el que idealmente colaboran el terminólogo y el experto en el campo. Tras el procesado de los textos, la lista de elementos resultantes debe ser filtrada, clasificada y ordenada empleando distintas técnicas supervisadas por el equipo de expertos (Slodzian, 2000).

Una vez obtenida una lista concisa de términos de acuerdo con los criterios preestablecidos, estos deben estructurarse y clasificarse para generar un producto de presentación final. Entre las distintas presentaciones terminográficas se pueden encontrar principalmente las *bases de conocimiento terminológicas*, los *tesauros* y las *ontologías*. Estas herramientas estarán más o menos especializadas y adaptadas al usuario final: enseñanza de idiomas con fines específicos, estudios de traducción, redacción de manuales o herramientas derivadas como es el caso de los lexicones utilizados en muchas aplicaciones de procesamiento de lenguaje natural.

#### **4.5.2 Las bases de conocimiento terminológicas**

El examen de los textos aproxima la terminología a la lingüística (especialmente a la lingüística de corpus y al análisis del discurso), en beneficio de la investigación en estas disciplinas. Los primeros bancos de terminología fueron el EURODICTION en Luxemburgo con orígenes en los años 1960 y el NORMATERM presentado en Francia en 1973. Sin embargo, su uso no se popularizó hasta la aparición de la informática

personal y el texto electrónico en la década de 1990. La informatización de los textos y al desarrollo de herramientas para estudiarlos, junto a la demanda comercial, pusieron a colaborar a expertos en informática y a lingüistas para rediseñar los modos de estudiar los textos en la construcción de terminologías. De aquí surgieron las *bases de conocimiento terminológicas* (BCT), en inglés *terminological knowledge base* (TKB) una estructura entonces novedosa de representación que relaciona en red a conceptos, términos y textos. Estas son una forma híbrida a caballo entre una base de datos terminológica y una base de conocimientos, esto es, contiene toda la información de un banco de datos terminológicos a la que se añade una gran cantidad de información conceptual altamente estructurada de manera que clasifica los conceptos de un dominio de manera explícita y estructurada (Aussenac-Gilles y Condamines, 2007).

Un ejemplo ilustrativo es *EcoLexicon*, desarrollado por la Universidad de Granada, una BCT sobre el medio ambiente mejorada a partir de técnicas lingüísticas y de representación del conocimiento. Está diseñada para la traducción de términos medioambientales entre seis idiomas europeos diferentes. Además, representa el conocimiento sobre el medioambiente mediante recursos multimodales: definiciones, concordancias, contextos, imágenes, animaciones, etc. (López Rodríguez, Buendía Castro, y García Aragón, 2010).

#### 4.5.3 Los tesauros

El término *tesauro* se ha utilizado para hacer referencia a distintas recopilaciones conceptuales similares dentro de la terminología o de la ciencia de la información de utilidad para distintos fines técnicos definida por Kilgariff y Yallop (2000) como “a resource in which words with similar meanings are grouped together” (p.1). Los tesauros reivindican enteramente la relación de los textos con el conocimiento manifiesto, por lo que estos juegan un papel esencial como clasificadores de la información en múltiples contextos. La historia de los tesauros para el lenguaje especializado y las aplicaciones computacionales de la lingüística se remontan a los albores de la disciplina a finales de la década de 1950 en la que tuvo lugar el diseño de la máquina de desambiguación, *Roget's Thesaurus*, que contenía un tesauro sobre tarjetas perforadas (Masterman, 1957, ap. Kilgariff y Yallop 2000). Desde entonces, el papel de la clasificación del conocimiento y sus términos ha sido esencial para el acceso a la información. Los autores destacan aquí cuatro tipos principales de tesauro:

- *Tesauros de estilo Roget*: (Bernard, 1986), publicados en formato impreso para ayudar a la selección de palabras durante la escritura a partir de las relaciones conceptuales y jerárquicas de cada término. Al contrario que los diccionarios monolingües, organizados alfabéticamente para la definición o la ortografía, este está estructurado según categoría y ofrece un amplio rango de posibilidades por cada entrada. No incluye definiciones y el usuario debe inferir el sentido apropiado de cada definición.

1. *WordNet* (Fellbaum, 1998): es una base de datos léxica desarrollada bajo principios psicolingüísticos de libre acceso. En la actualidad sigue siendo extensamente utilizada en investigación del procesamiento del lenguaje natural.

Originalmente desarrollada en inglés, a partir del proyecto *EuroWordNet* este sistema se ha convertido en una herramienta multilingüe. El significado de cada palabra se encuentra de acuerdo con un *synset* (*conjunto de sinónimos*) y estos tienen relaciones jerárquicas y de otro tipo para cada uno otro. Está organizado según relaciones de sinonimia y la hiponimia para sustantivos y verbos y mediante antonimia para adjetivos. Para ilustrar las posibilidades de la herramienta, en *WordNet 1.5* existen 25 clasificaciones relacionales de alto nivel para los sustantivos organizadas de manera jerárquica y, para los verbos, la clasificación de nivel superior tiene 15 tipos.

2. *Tesauros manuales para recuperación de información (IR-manual)*: son realizados parcialmente de manera manual y se utilizan para la extracción automatizada de información. Al igual que *WordNet*, se organiza mediante sinonimia y la taxonomía, lo que hace posible para ampliar o reducir las búsquedas, y comparar con documentos que usan sinónimos de los términos de búsqueda. Se destaca el *Unified Medical Language System* (UMLS) que incorpora una gran cantidad de conocimiento del dominio de la medicina (Sanfilippo et al., 1999).
3. *Tesauros automáticos*: se recopilan a partir del procesado de corpus, a partir de la similitud entre las palabras medidas directa o indirectamente a través de vectores de coocurrencia estadística. Parte de este trabajo tiene lugar a partir del procesamiento del lenguaje natural y otros bajo el título de recuperación de información.

El tesoro se relaciona en dos momentos clave de la interacción con la información especializada: *a priori* durante su indexación y *a posteriori* durante su investigación. A diferencia de otros productos de la terminología, desde el principio la relación con los textos como conocimiento manifiesto y su existencia está fundamentada por la necesidad de crear un vínculo entre los documentos y el conocimiento. Los tesauros fueron la primera evidencia práctica de que la ciencia de la información y la terminología. Como disciplinas, coinciden en el modo de representación de la información: conceptos conectados por relaciones que estructuran un sistema. Además, ambas tienen las mismas tensiones a la hora de establecer criterios para la normalización de las clasificaciones, por lo que es necesario encontrar puntos de equilibrio entre ideología y representatividad. Finalmente, el acercamiento de la terminología a la ciencia de la información permite realizar una mejor automatización de la tarea clasificación de los textos y, a cambio, ofrecer una nueva perspectiva desde el uso de la lengua (Aussenac-Giles y Condamines, 2005; Laine-Cruzel, 2001; Pédauque, 2005;).

#### 4.5.4 Las ontologías

Las ontologías se desarrollan dentro del campo de la ciencia de la información para la reutilización y el intercambio simplificado de conocimiento de dominio entre los sistemas que las utilizan. Para ello, los sistemas deben tener representaciones compatibles y consistentes del mundo, representados en el mismo lenguaje estándar. Las ontologías reflejan así la búsqueda de invariantes en un dominio, una descripción genérica, al menos consensuada, de conocimiento (Aussenac-Giles y Condamines, 2005). En cuanto a su definición, no existe un acuerdo claro sobre los roles de la ontología, hecho notablemente paradójico al ser una disciplina que se encarga de organizar, clarificar y consensuar el conocimiento (Moreno-Ortiz, 2008). La definición clásica parte de que la conceptualización es la representación formal de un área de interés que incluye sus objetos, conceptos, entidades, etc. La explicitación formal de una conceptualización, entonces, sería la ontología (Gruber, 1995), pues esta existe a través de su sistema relacional. El desarrollo de las ontologías es el punto de partida o campo de trabajo de la inteligencia artificial, la web semántica, las distintas ramas de la informática o la archivística las cuales hacen uso de ellas para simplificar y organizar la información. Esta forma de organización de conocimientos, según la definición clásica de Noy y McGuinness (2001), consta de los siguientes elementos:

- *Clase*: serie de objetos del que consta una entidad (físicos, tareas, funciones, etc.).
- *Relaciones*: enlazan los distintos conceptos ontología para representar su localización, jerarquía e interacciones.
- *Propiedades*: cada objeto es descrito a partir de un juego de atributos que acumularán distintas clases de valores. Entre ellas, además existen especificaciones, rangos y restricciones.
- *Instancias*: son objetos, miembros de una clase, que no pueden ser divididos sin perder su estructura y características funcionales. Pueden ser agrupados en clases.
- *Axioma*: en relación con las entidades, son los elementos que permiten la modelación de una serie de órdenes lógicas que se cumplen en la realidad. Existen tres tipos de axiomas: relacionales, no relacionales y generales.

Inspirada primero por la ingeniería de software, los métodos de construcción ontológica comenzaron a cuestionar la definición y organización de los conceptos alrededor de 1996 y con el cambio de siglo se comenzó a utilizar de manera sistemática a los textos como fuentes de información (Aussenac-Giles y Condamines, 2005; Fernández-López et al., 1999). El ciclo propuesto aquí es de carácter múltiple: el procesamiento automático del lenguaje permite la automatización parcial de la extracción de información que es posteriormente tratada mediante aprendizaje automático de manera más eficiente que la clasificación manual de la información. Con los datos obtenidos se cierra el círculo epistemológico tras la renovación de hipótesis sobre el estado de los conceptos y su relación con los términos.

El desarrollo de la web como red mundial masiva de intercambio de información ha supuesto la necesidad de resolver los problemas de eficiencia y extracción de información a partir de nuevas arquitecturas surgidas de la web semántica. Además de la información sin clasificar, esta hace uso extensivo de los metadatos para la anotación y clasificación de documentos. Moreno-Ortiz (2008) considera que la terminología, desde su propia naturaleza clasificadora, ha sido partícipe de este desarrollo de las ontologías de manera sistemática. Desde las taxonomías que dieron lugar a la nomenclatura biológica hace varios siglos hasta los modernos sistemas léxicos como *WordNet*, la clasificación de la información conceptual siempre ha sido un campo de trabajo fundamental para esta disciplina y, desde su concepción, ha aportado formas de hacer que han trascendido a la mera comunicación técnica. Las ontologías solucionan muchos problemas a la hora de definir términos, pues al clasificarse de manera jerárquica, al menos sobre el papel, el

usuario (ya sea humano o digital) puede localizar de manera lógica la ubicación del significado dentro del mapa de conocimiento. De esta forma, el problema de la definición queda mejor planteado, ya que esta puede estandarizarse con mayor claridad y evitar ambigüedades y circularidad, pues determina un vocabulario especializado más preciso y con la cobertura más ajustada.

Un ejemplo sería *BabelNet* (Navigli y Ponzetto, 2012), una ontología lexicalizada desarrollada de forma automática a través de la integración de *Wikipedia* y *WordNet*. *BabelNet* es un diccionario enciclopédico multilingüe que enlaza 380 millones de conceptos y entidades mediante amplias relaciones semánticas en conjuntos de sinónimos similares a los *synsets* de *Wordnet*. El uso de ontologías como *BabelNet* tiene múltiples aplicaciones en el procesamiento del lenguaje natural, desde desambiguadores a traductores automáticos a mejoras de la eficiencia en buscadores.

#### 4.5.5 Extracción terminológica

El diseño de soluciones para la extracción y gestión de terminologías es una de las áreas más exitosas de la lingüística de corpus y el procesamiento del lenguaje natural. Los extractores terminológicos se definen como un conjunto de aplicaciones computacionales que reconocen y extraen las unidades terminológicas que aparecen en los textos. El reconocimiento automatizado de las unidades de dominio específico a partir de textos se conoce como extracción terminológica. La extracción terminológica es en la actualidad uno de los pasos principales esenciales del análisis terminológico y su aplicación es a partir de herramientas computacionales que facilitan la identificación de términos candidatos (o candidatos a término). Sus fuentes de datos son los corpus textuales y bases de datos léxicas, así como los textos etiquetados a mano para entrenar herramientas de aprendizaje automático. Dichos candidatos serán palabras o grupos de palabras consideradas etiquetas en un cuadro conceptual más amplio.

A menudo hay una combinación de este trabajo con el estudio de las relaciones jerárquicas, mediante herramientas que ayuden a identificar las relaciones conceptuales. En conjunto con la extracción, el diseño de aplicaciones de estructuración de los términos extraídos tiene un papel altamente relevante para la construcción de ontologías, indexación de textos o traducción automatizada. Según el tipo de técnica aplicada, esta tarea dependerá de la clasificación previa de al menos una serie de términos. De hecho, la identificación de las relaciones semánticas es uno de los métodos utilizados con más



frecuencia para la identificación de conceptos y para la identificación de las distintas instancias para la asociación de los términos con sus respectivas clases a través del uso de principios de aprendizaje y extracción de información. Las tecnologías como el *aprendizaje de ontologías* o las relacionadas con la *Web semántica* dependen de la información extraída automáticamente desde corpus textuales mediante el uso de herramientas capaces de detectar terminología y relaciones terminológicas. (Cabré, Estopà, 2000; Estopà, y Vivaldi, 2001; Pazienza et al., 2005;).

La dificultad y las diferencias teóricas a la hora de aportar definiciones a la palabra *término*, como se ha podido observar en este capítulo, puede generar contrariedades a la hora de convertir de manera operativa las propiedades de los lenguajes especializados en un sistema funcional. El desafío de la extracción terminológica se afronta principalmente desde tres enfoques diferentes: *estadístico*, *lingüístico* e *híbrido* (Pazienza et al., 2005). Según Estopà (2002), las características de la extracción terminológica implican seis parámetros diferentes:

1. Información de partida
2. Estrategias de reconocimiento
3. Estrategias de filtrado
4. Estrategias de adquisición del conocimiento
5. Interacción del sistema con el usuario
6. Resultados, grado de satisfacción.

Una vez obtenidos los resultados, la selección de los candidatos relevantes depende del terminólogo. Estas herramientas se basan en varios principios diferentes, aunque no incompatibles, que dejan ver las distintas concepciones del funcionamiento del lenguaje. Ya hace más de cuarto de siglo, Bourigault (1992) propuso el diseño de extractores basados en la sintaxis a través de *parsing*, ya que la forma gramatical de las unidades terminológicas es relativamente predecible. Distintos *enfoques lingüísticos* extraen la terminología mediante la captura de sus propiedades sintácticas y determinaron que los términos tienen con frecuencia composiciones sintácticas claras, por ejemplo, una amplia mayoría de los términos son sintagmas nominales. Las herramientas de este primer tipo consideran cualquier texto como la implementación de un sistema muy estable en el que los términos respetan los patrones gramaticales recurrentes (p. ej., “Adj N”, “N of N”) a partir de una gramática local similar a la propuesta por Hunston y Francis (2000). Estos patrones se identificarían mediante observación e introspección. Por ejemplo, se



conocería la hiponimia mediante patrones como el ejemplo propuesto por Condamines (2005) para el francés. En él, N1 sería el hiperónimo de N2:

Paul aime toutes les FLEURS sauf les ROSES.

Tous les N1 sauf le/la/les N2.

En resumen, un sistema extractor de terminología a partir de un corpus que siga un enfoque lingüístico debe contener dos módulos básicos funcionando en cascada. Primero, un módulo de procesamiento sintáctico que garantice la identificación de sustantivos, verbos y adjetivos, y; segundo, un módulo identificador de términos que descarte formas irrelevantes.

El segundo tipo de herramienta considera que no todos los fenómenos lingüísticos que pueden aparecer en los textos son predecibles. Este enfoque sigue dos dimensiones principales: *calidad de unidad (unithood)* y *calidad de término (termhood)*. La *calidad de unidad* se encarga básicamente de las colocaciones, mientras que la *calidad de término* determina el grado de pertenencia de la unidad dentro de un dominio específico. Ambas medidas se clasifican de acuerdo con la metodología y las presunciones que tomamos a partir de ellas. Este enfoque espera descubrir los elementos que no son susceptibles de ser hallados a través de la introspección a partir de medidas estadísticas que utilizan cálculos de frecuencia para determinar el grado de *calidad de término* que tienen las unidades candidatas a término. Se parte de una tradición distribucional que considera la recurrencia de contextos y distribuciones. En él se detectarían las unidades léxicas cuya ocurrencia en el dominio específico es más alta que los resultados obtenidos desde un corpus de dominio general. Sin embargo, los autores indican que los enfoques puramente estadísticos raramente aportan resultados totalmente satisfactorios, *p. ej.*, términos poco frecuentes son más difíciles de extraer si no se filtran después mediante una perspectiva lingüística ya que pueden incluir elementos no deseables (Cabré, et al., 2001; Pazienza, et al., 2005). Se han resumido estas técnicas en la Tabla 6: el grado de asociación, la significancia de las medidas de asociación y las medidas heurísticas. A pesar de que los resultados obtenidos de cada esquema son altamente precisos, su funcionalidad difiere a los resultados obtenidos en los sistemas con enfoque lingüístico ya que según Cabré et al. (2001) extraen resultados que son, en cierto modo, diferentes.

Las medidas *heurísticas* se basan en suposiciones intuitivas y estadísticas. Definen el *grado de asociación (Mutual Information y Dice factor)* y la *significancia de asociación (Z-Score, T-Score, X y Log-Likelihood Ratio)*, los cuales están basados en datos

fuertemente estadísticos. Las fórmulas que se usan para encontrar los resultados numéricos fueron recopiladas por Velardi, Fabriani, y Misikoff (2001). El objetivo de las medidas relacionadas con el *grado de asociación* es determinar la posibilidad de ver si dos o más palabras adyacentes son colocaciones (Tabla 6). Por otro lado, las medidas de *significancia de asociación* evitan errores de estimación (Pazienza et al., 2005).

El enfoque *híbrido*, a medio camino entre ambos, usa ambas visiones y toma en cuenta pistas de tipo lingüístico y estadístico de manera automatizada. Lo que ambos enfoques tienen en común es que analizan un corpus de textos especializados y extraen listas de palabras que posteriormente deben ser confirmadas a mano. Por ejemplo, los resultados obtenidos desde un determinado enfoque lingüístico se refinan mediante el uso de medidas estadísticas. Como consecuencia, los resultados finales se consideran más fiables, ya que los resultados extraídos desde un enfoque lingüístico se ordenan y clasifican a posteriori según las cifras estadísticas de *termhood* y *unithood* (Cabré, et al., 2001; Condamines, 2005; Pazienza et al., 2005).

		<i>Dimensión estadística</i>		
		Grado de asociación	Significancia de la asociación	Heurística
<i>Dimensión lingüística</i>	Calidad de unidad			MI <sup>2</sup> MI <sup>3</sup>
		Mutual Information (MI)	Z-score T-score	
	Calidad de Término		X <sup>2</sup> Ratio LogLikelihood	Frequency C-value Coocurrencia
		Dice Factor		

**Tabla 6** Clasificación de medidas estadísticas en las dimensiones estadística y lingüística (Pazienza, Pennacchiotti, & Zanzotto, 2005).

Existe una serie de desafíos para la extracción terminológica automática, siendo el *silencio* y el *ruido* los problemas más comunes. Estos conceptos están relacionados estrechamente con las métricas de precisión y exhaustividad que se presentarán en el Capítulo 6. En términos generales, un análisis hecho a mano por un experto no coincide con los resultados de un sistema automatizado. Son el *silencio* y el *ruido* los retos más evidentes que deben ser resueltos por los lingüistas computacionales. De acuerdo con Estopà (2000), hay dos tipos de *silencio*: *silencio intrínseco* y *silencio extrínseco*.



El *silencio intrínseco* se refiere al conjunto de elementos no detectados que deben ser encontrados por el sistema. Según los autores, asciende al 5%-10% de las *unidades terminológicas multipalabra*, y se produce por errores de ambigüedad, superposición de términos y elementos discursivos ocultos. Entre estos errores destacan:

- Los *errores de ambigüedad* son un tipo de unidades silenciadas debido a errores de etiquetado que tenemos a la hora de identificar el tipo de palabra. Por lo general, no responden a la estructura gramatical predicha por el sistema. Por ejemplo, cuando la palabra “attorney” se identifica como adjetivo en la expresión multipalabra atípica “attorney general”.
- Los *términos superpuestos* son términos complejos que contienen otras unidades simples o complejas. Por ejemplo, “high tech entrepreneur” se compone de dos términos independientes combinados: “high tech” y “entrepreneur”. El problema surge cuando es necesario detectar todas las TUs que son parte de otras TUs complejas, ya que algunas unidades constituyentes pueden ser también términos de manera aislada.

Relacionado con el *silencio intrínseco*, el problema más complejo que debe superarse es la *extracción de anáforas*: unidades cuyo núcleo o complemento haya sido omitido por razones discursivas. La mayoría de las anáforas usadas en el discurso se utilizan para hacer la lectura más sencilla, generalmente para no repetir elementos mencionados previamente. Por ejemplo, podemos encontrar:

- (a) Unidades escondidas en oraciones que hacen uso de cópulas (p. ej., “economic and political costs”).
- (b) Unidades ocultas en sintagmas de tipo especificativo (p. ej., “New Perspectives: Social, Political and Economic impacts of emerging Web 2.0 platforms”).
- (c) Unidades ocultas en oraciones comparativas (p. ej., “implications in the politics, as economic level”).
- (d) Unidades ocultas en oraciones condicionales (“in macro or even microeconomics”).
- (e) Unidades ocultas en oraciones predicativas (p. ej., “the currency is pegged”).

Por otro lado, el *silencio extrínseco* sería el conjunto de segmentos terminológicos ignorados por el sistema que no son parte de los objetivos de extracción. Según la autora,

esto se causa por los límites propios de los sistemas de extracción terminológica. Los extractores, generalmente, detectan *Unidades Terminológicas Poliléxicas* (UTP), sin embargo, no son el único tipo de unidad terminológica. Hay más unidades especializadas, como las unidades terminológicas monoléxicas, los acrónimos (p. ej., “IMF”), las unidades fraseológicas (p. ej., “under demand”), latinismos, etc. Para resolver esto, la autora propone varios puntos para facilitar la extracción de unidades no multipalabra que facilitan, por tanto, el proceso de extracción terminológica. Entre ellos, se promueve el uso de nomenclaturas a la hora de escribir sobre un dominio específico para facilitar el procesado automatizado en un mundo informatizado. El uso de terminología estándar facilita la tarea de comprender lo que está escrito. Además, propone el uso de una heurística de morfología léxica. Por ejemplo, la detección de adverbios que derivan de términos adjetivales que generalmente significan "from a [ADJ] point of view".

Para concluir, el segundo desafío de los extractores terminológicos es el *ruido* o la cantidad de términos candidatos que son inválidos. Los principales factores de ruido se originan desde el concepto de unidad terminológica con los que trabajan los extractores. Para los sistemas de extracción terminológica, los términos son una forma exclusiva que depende de un determinado dominio y la mayoría de los sistemas solo tienen en cuenta las estructuras sintácticas. Sin embargo, Estopà (2000) recalca que una unidad terminológica no es sino la relación entre el contenido y la forma. La causa del ruido depende de la imposibilidad de diferenciar las unidades terminológicas poliléxicas de las estructuras morfosintácticas generales. Las unidades que comparten su misma estructura son las UTP, las unidades fraseológicas especializadas, las combinaciones recurrentes, las unidades de discurso y las unidades multipalabra que no comparten el mismo dominio.

#### **4.6 Características del lenguaje periodístico en lengua inglesa**

Si bien los textos periodísticos no alcanzan el nivel de especialización que tienen los textos de índole más técnica como por ejemplo el inglés médico o el inglés económico, el lenguaje periodístico tiene características específicas diferenciadoras. El lenguaje periodístico es un uso lingüístico especializado que en gran medida realiza la labor divulgativa de interpretar textos especializados (políticos, económicos, etc.) al gran público, ya que la selección de los elementos léxicos y sus construcciones tienen altos niveles de sofisticación, reflejando de manera consciente la ideología de sus direcciones editoriales. Dada la naturaleza de este trabajo, es necesario describir las particularidades

del lenguaje periodístico y es interesante comenzar hablando sobre la naturaleza de la producción de la noticia, es decir, el cómo se produce una noticia desde que se produce un evento hasta que queda plasmado en un medio de comunicación para la lectura del público. Posteriormente, se revisarán aquí las características léxicosintácticas, las construcciones de evidencialidad y, finalmente, la estructura de los textos de las noticias.

Bednarek y Caple (2011) hablan de un *contexto comunicativo* en el discurso periodístico que además de los distintos actores engloba un contexto socioeconómico determinado y un histórico de circunstancias previas que retroceden a mayor o menor distancia en el tiempo. La relación dialógica de la noticia abarca tres elementos principales, que se relacionan de manera triangular y que son fundamentales para que se produzca el hecho noticiable: (a) discurso periodístico, (b) productor/es y (c) audiencia/s. La construcción de este lenguaje especializado depende de la compleja asimetría en la relación discursiva entre el productor y la audiencia.

Aunque la naturaleza de este estudio no abarca el estudio en profundidad de estos factores, es interesante resumir el trasfondo en la construcción discursiva de una noticia. Tras ella existe una larga cadena de trabajo y *cocina* periodística, es decir, la negociación discursiva de jefes de sección, redactores, editores, etc. que se ocupan del texto periodístico: las temáticas cubiertas, sus enfoques particulares y la selección de las líneas de lenguaje con el que se elaborarán los titulares. El público recibirá el producto final: los cuerpos de la noticia que se adaptarán a los medios en los que aparezcan (redes sociales, noticieros informativos, prensa escrita, etc.). Sin embargo, Bell (1991) discute que la elaboración de la postura u opinión de un medio no reside en el equipo editorial que redacta la noticia, sino en el medio editorial como institución, es decir, propietarios, gerentes y sus relaciones de poder con el entorno sociopolítico. Este constructo de relaciones de poder hará que la noticia refleje unas líneas generales ideológicas y unos valores que respondan, en gran medida, a la visión ideológica fomentada de estas instancias. Un ejemplo paradigmático de este modelo sería el conglomerado multinacional *News Corporation* del magnate Rupert Murdoch con cabeceras tan importantes a nivel global como *The Wall Street Journal* o *New York Post* en Estados Unidos o *The Times* o *The Sun* en Gran Bretaña que han sido históricamente muy influyentes en los círculos de poder del mundo anglosajón.

Finalmente, en cuanto al lenguaje especializado, es esencial tener en cuenta el complejo rol de la audiencia. Se parte históricamente de una selección tradicional de audiencia muy estática basada en factores sociales como la ideología, edad, clase y

género, además del nivel de especialización del texto. La selección es previa al advenimiento de los medios en línea y, obviando secciones muy limitadas como las cartas al director, tienen un carácter unidireccional e impersonal en la que no existía posibilidad de respuesta de parte del público. Las grandes cabeceras siguen siendo consideradas voces de autoridad y referentes creíbles para las informaciones que se publican en las redes sociales. Sin embargo, desde el siglo XXI, se ha producido una transformación en la presentación de los artículos, y en gran medida del lenguaje utilizado. En la actualidad, la audiencia ejerce un rol más activo, actuando de modulador al ofrecer comentarios, críticas o reelaboración de las noticias presentadas en línea (*mash-ups*) y una mayor contribución de los lectores a los medios, p. ej., como ciudadanos que ejercen de colaboradores y aportan imágenes o declaraciones exclusivas de última hora. Por ello, la selección del lenguaje y de los factores textuales es más delicada que nunca. Teniendo estos factores en cuenta, es necesario destacar que, a partir de la digitalización de los medios, se está produciendo una mayor fragmentación y especialización del público, de manera que el lenguaje se enfoca discursivamente cada vez más a un público determinado (Bell, 1991; Bednarek y Caple, 2012).

Entrando en la cuestión específica del periodismo económico, este tiene como objetivo la divulgación de hechos y cuestiones relacionadas con los distintos niveles de la economía pública y privada, así como el desarrollo de los eventos en el sector financiero, conectando con la continuidad de los eventos económicos en el mundo y las relaciones de causa y efecto entre todos estos eventos. El periodismo económico tiene la particularidad de no solo informar al público sobre hechos relevantes a través de un lenguaje que, si bien tiene una abundancia de elementos terminológicos, este es más plano y asequible para acceder a la mayor cantidad posible de público. Así, su objetivo será narrar e interpretar la constante evolución de la realidad económica a través de distintos indicadores interrelacionados que suelen ser de carácter muy abstracto (p. ej., tasa de crecimiento, desempleo, inflación, etc.) y que están valorados en la inmensa mayoría de los casos desde una perspectiva capitalista, matizada y valorada a través de sus distintas corrientes (Paiva Jacobini, 2008).

El resultado final de toda esta red de factores socioeconómicos, de poder, de noticiabilidad y de organización del trabajo se materializa en textos que comparten una serie de funciones lingüísticas y extralingüísticas, características morfosintácticas y de organización textual de manera que se cumplan los objetivos requeridos por el formato

de noticia, el estilo periodístico y la concisión, objetividad, claridad y precisión en los textos que reportan el hecho noticioso.

Las características léxicas y sintácticas revisadas por Bednarek (2008) y Bednarek y Caple (2012) se relacionan estrechamente con el contexto comunicativo del discurso periodístico. Entre ellas se encuentran los sustantivos y frases nominales, los verbos y frases verbales y los adverbios y circunstanciales. Trabajos como los de Bell (1991) o Jucker (1992) sugieren que los sustantivos son muy frecuentes<sup>23</sup> en la redacción periodística en inglés y que, según Bednarek y Caple (2012), actúan como modificadores de otros sustantivos (p. ej., “shareholder revolt” o “investor anxiety”). Otros fenómenos son las nominalizaciones, adjetivos atributivos que actúan como modificadores de los sustantivos (p. ej., “strategic growth”) en detrimento de adjetivos predicativos, así como frases preposicionales a la derecha de un sustantivo (p. ej., “no experience of the clothing and fashion businesses”). Estas pre y posmodificaciones permiten abreviar los textos y agregar varios tipos distintos de información en un corto espacio, hecho especialmente destacable en los titulares con vocabulario muy específico elegido no solo por un menor número de caracteres, sino por su sonoridad (p. ej., “bigwigs por executives”).

Estos sustantivos hacen referencia a ciudades, países o nombres propios, en contraste con la escasa frecuencia de pronombres personales<sup>24</sup>. En cuanto a sus usos principales Bednarek y Caple (2012) encuentran los siguientes:

- Expresión de tiempo y lugar (“Today’s agenda” o “Syntagma Square”).
- Etiquetado de actores y fuentes en las noticias (“the co-leader of the Democratic Council”, “no investors”, “protesters claim that...”).
- Intensificación o evaluación (“took the market by STORM”, “efforts to help eurozone STRUGGLERS such as Greece”).

El segundo conjunto lexicosintáctico destacable son los verbos léxicos y la escasez de uso de verbos modales. Bednarek y Caple (2012) mencionan una ratio de 16 a 1, con la excepción de “will/would”.

---

<sup>23</sup> En el caso del corpus de trabajo, el *Esmeraldas Corpus of Business English*, de carácter periodístico, cuenta con una frecuencia de 71,89 ppM en combinaciones de sustantivos, mientras que en un corpus canónico como el BNC, la frecuencia baja a 45,066 ppM.

<sup>24</sup> Se encuentran diferencias de frecuencia de 35 ppM en el corpus ECBE contra 57ppM en el BNC.



Sobre los tiempos, los verbos presentes se usan para hacer referencia a procesos en marcha o, en otros casos, verdades intemporales de manera que se enfatiza en la novedad del evento y su valor como noticia. También los tiempos presentes tienen una ligera preponderancia sobre los tiempos en pasado. Asimismo, se omiten con frecuencia los verbos auxiliares y perífrasis por cuestiones de espacio y dinamismo, p. ej., sacrificando el uso de *going-to future* en favor de tiempos simples “Royal family quits” en lugar de “The Royal family is going to quit”, o la omisión de la perífrasis *is going to* en “Mexico's new government to cut VAT in region bordering United States”.

Tanto la voz activa como la voz pasiva son frecuentes. La voz pasiva se utiliza para aportar información de fondo y suele omitirse el agente de una manera consciente, para ocultar la agencia o por irrelevancia de este. Las interrogaciones son un fenómeno poco frecuente en este lenguaje y se suelen utilizar de manera retórica p. ej., “Does Income Inequality Lead to Consumption Inequality?”. La mayoría de las entrevistas están escritas en estilo indirecto (p. ej., “When asked about”) ya que el interés reside en la información que ofrece la persona entrevistada.

Los sintagmas adverbiales y sus construcciones afines son frecuentes en algunas de sus categorías. Principalmente, se habla aquí de los circunstanciales de tiempo y de lugar, mientras que los adverbios utilizados para estructurar el discurso (“nevertheless”, “however”) no son frecuentes. Los primeros (“yesterday”, “tomorrow”, “this quarter”) sitúan las historias en un eje cronológico o se producen hipótesis o anuncios sobre el futuro. Los adverbiales de lugar ofrecen información relevante sobre el trasfondo de la historia situándola en un espacio físico.

Especialmente en la prensa económica, es abundante el uso de cantidades y cifras que permiten ilustrar a partir de datos estadísticos que aporten un nivel mayor de credibilidad y evidencialidad a la noticia dando un aspecto más factual a la información aportada. Para facilitar la lectura, se prefiere redondear las cifras sacrificando en muchas ocasiones la precisión a favor de una mejor legibilidad. De esta manera, son frecuentes construcciones utilizadas para este fin, como “up to” (“The company has earmarked up to £400m this year for acquisitions.”) o about (“About 85 percent of its patients are in the Medicare trial”).

En este tipo de discurso la voz de la enunciación contiene significados específicos. Las citas directas en este tipo de discurso suelen utilizarse, además de lo expuesto anteriormente, para aportar intensidad al texto y distanciar la opinión de la persona citada de la postura de la cabecera. El uso de la voz pasiva, más frecuente, da un aspecto de



mayor neutralidad y alineación con la narración del texto. No obstante, Bednarek y Caple (2012) aportan un listado de los distintos tipos de verbos utilizados para el estilo indirecto:

- Neutral: cuando no se aporta otra información que lo expresado, p. ej., “say”, “tell”.
- Illocucionario: cuando da información sobre el propósito del hablante, p. ej., “demand”, “promise”.
- Declarativo: hace referencia a actos institucionales, p. ej., “acquit”, “plead guilty”.
- Señalador de discurso: hace de conector intratextual, p. ej., “conclude”, “add”.
- Paralingüístico: aporta información extra sobre la calidad del discurso, p. ej., “whisper”, “scream”.

La estructura textual de la noticia se ha dividido tradicionalmente en titular, entradilla y cuerpo. El titular se publica en letras de mayor tamaño y encuadra en pocas palabras y de manera atractiva la naturaleza del evento descrito. Las entradillas típicas son cortas y ofrecen datos claves en una, dos o tres oraciones de manera que aporten brevemente los elementos noticiables de la historia y sirvan como comienzo de la redacción principal de la noticia. Bell (1991) resume sus principales características: (a) mencionan el evento principal y el lugar (el clásico *who, what where*), (b) construyen noticiabilidad, comienzan desvelando el actor, (c) cubre más de un evento, (d) aporta trasfondo a la historia.

En cuanto a la estructura del cuerpo de la noticia, suele seguir un patrón de pirámide invertida en el que la información parte de lo más general a lo más específico, de manera que recicla la información, añadiendo un detalle nuevo en cada paso, sin seguir un orden cronológico establecido. De esta manera, se genera una tensión en la lectura, manteniendo los detalles más noticiables sobre los detalles menos destacados.

Uno de los hechos más característicos del lenguaje periodístico en inglés es el llamado, no sin sorna, *headlines*<sup>25</sup>, palabra que representa la forma tan específica de escribir titulares en lengua inglesa. Los titulares cumplen con la función de resumir brevemente la historia y con la metafunción interpersonal de atraer al lector y enfocan o muestran la postura o perspectiva del resto de la historia a través de estrategias visuales (uso de tipografías específicas y otros efectos visuales) y, especialmente, lingüísticas:

---

<sup>25</sup> Traduciendo libremente a Bednarek y Caple (2012), *titularés* combinando las palabras “titular” y p. ej., “japonés” (“headline” + “Japanese”).

- Uso de un lenguaje intenso o altamente evaluativo (p. ej., “Blow for House of Fraser as Chinese firm drops plan to invest”).
- Figuras retóricas (p. ej., “Donald Trump gets a helping hand from Brussels bean counters”).
- Omisión de palabras gramaticales (p. ej., “Insulin maker Sanofi stockpiles medicines ahead of Brexit”).
- Si hay verbo, estará generalmente en tiempo presente y sin auxiliares (p.e “FCA proves to be a paper tiger in case of RBS mistreatment”).
- Frases oracionales con sustantivos premodificadores (p. ej., “flight cancellation crisis”).
- Escasez de usos de marcadores de tiempo o atribución.

La *evidencialidad* es la información que ofrece el texto sobre el origen de la certeza a la hora de enunciar el discurso puede o no hacer referencia a las fuentes de información. Así una noticia puede permanecer sin atribución al no ser desvelada la forma en la que obtuvo la información, ofrecerla de una manera vaga a través de formulaciones como p. ej., “allegedly” o “is said to” (p. ej., “One Tullett broker allegedly told Verrier: Your company's shit. Why should we do this?”), las cuales se utilizan principalmente por motivos legales. La atribución a fuentes específicas es la más prestigiosa de las tres muestras de evidencialidad de manera que se indica la fuente de una manera parcial o total, mencionando explícitamente a la persona que ofrece la información (“Mr Cullum has stated that he does not want to stay in charge”) o aportando la información relevante sin citar a la persona expresa que provee dicha información (“Senior Fed officials have stated repeatedly that they had to take unusual steps in 2008”).

Los patrones de lenguaje evaluativo, complementando lo revisado en el Capítulo 3, son transversales al discurso periodístico. Bednarek y Caple (2012) revisan aquí elementos frecuentes en el lenguaje periodístico utilizados con fines de noticiabilidad de manera que el sentimiento relacionado con los distintos elementos pueda ser identificado fácilmente de manera positiva o negativa. Por ejemplo, los eventos (“historic” o “fiasco”) o los actores (“top” o “wannabe”). En términos similares a los expuestos en el sistema de gradación relativo a la teoría de la valoración que se revisa aquí en el Capítulo 3 de Martin y White (2005), existen cuantificadores muy frecuentes en el lenguaje periodístico que se utilizan para maximizar la evaluación de los textos, ya sean intensificadores de cantidad (“millions of”) o tamaño (“huge”, “peanut-size”) o fuerza (“blaze”, “ravage”), etc. Se

destaca aquí también la presencia de léxico negativo en el lenguaje periodístico, tanto en el lenguaje evaluativo (p. ej., el uso figurativo de “nightmare” es muy frecuente) o las referencias a las emociones (“petrified”, “devastating”). En conjunto, estos fenómenos se utilizan para generar un efecto de superlatividad, de manera que se maximiza para los lectores la exposición de los eventos y el impacto, y se explota mediante recursos lingüísticos el número de las visitas a las noticias.

#### 4.7 Características del lenguaje económico en lengua inglesa

El conjunto de lenguas de especialidad del inglés de la economía, las finanzas y los negocios es quizá el dominio especializado con estudios de mayor recorrido histórico, dado su carácter como lengua franca y en la que mayor material escrito se genera (Robinson, 1980). Dentro del estudio de lenguas de especialidad, Swales (1985) denominó a esta especialidad como uno de los referentes didácticos primordiales para el desarrollo del inglés de la ciencia y la tecnología y, como consecuencia, del estudio del resto de lenguas de especialidad. Fuera del ámbito de la didáctica, comenta Spillner (2007) que existe un déficit dentro de la investigación lingüística en textos de comunicación económica, mencionando a los textos ya añejos en lengua inglesa como Gallais-Hammono (1980; 1982) o Gläser, (1988).

Desde una perspectiva que relaciona el lenguaje con la filosofía de la economía, Mongin (2006) recuerda que aún existen economistas que dan por sentado el principio positivista de que solo los juicios matemáticos de carácter factual y no los juicios de valor son susceptibles de discusión racional. De hecho, el debate filosófico de la economía anglosajona se remonta a los tiempos de Keynes, en la tricotomía que distingue entre *ciencia positiva*, *ciencia normativa* y *arte de la economía*. Desde hace cuatro décadas, autores como Sen (1970) discutieron esta corriente *dura* de pensamiento, y la reemplazaron con una corriente epistemológicamente más *suave* que propone la existencia de diferencias marcadas a la hora de evaluar las afirmaciones de carácter económico. En efecto, la visión de la economía como ciencia no puede ser la misma si se la interpreta como una disciplina preponderadamente imparcial o, como se considera aquí, dependiente de los valores y el pensamiento económico.

Según Mongin, cuando alguien expresa: “X is a good economic policy”, se deben considerar cuatro factores. La *oración* en sí, como entidad lingüística; la *declaración* o significado tras la oración; la *expresión*, que dibuja el acto material realizado por el

hablante; y la *opinión* o *valoración* que se está emitiendo. La novedad en este dominio reside a la hora de la expresión de este último factor. Un hablante experto en el dominio económico realizará un juicio de valor si la declaración es evaluativa, en los términos expresados a lo largo el Capítulo 3; y haría un juicio *factual* en el que considerarán los hechos positivos o negativos a partir de información económica si la declaración es lógicamente objetiva y lo afirma sinceramente. Por lo tanto, si se tiene en cuenta que una decisión de evaluativa “good” significa que un predicado evaluativo de “X is a good economic policy”, en el caso de un juicio factual se preguntará si el economista juzga que X es una buena política económica. Por ello, una declaración de carácter económico que se salga de las férreas líneas positivistas está incluida en uno de estos cuatro supuestos: (a) afirmaciones que no son directamente prescriptivas ni lógicamente evaluativas; (b) declaraciones que son prescriptivas o lógicamente evaluativas, y son afirmadas sinceramente por el economista, (c) afirmaciones más controvertidas que no cuentan como prescriptivas o lógicamente evaluativas y se deben “coger con pinzas”; o (d) declaraciones para las cuales no hay evidencia de que el economista las haya afirmado con sinceridad. De estas cuatro, solo las dos primeras son relevantes para este estudio, ya estas dos últimas se escapan del foco de estudio del análisis de sentimiento.

A pesar de ser considerado un ente de alto nivel de abstracción, la esfera económica es un mundo en el que priman las relaciones comunicativas. Los tan renombrados “mercados” a los que muchos políticos hacen referencia como entes indeterminados han sido descritos como sistemas eminentemente comunicativos totalmente compatibles con la teoría de actos de habla austinianas (Lusztig y Schwag, 2014). En relación con esto, Palmieri y Palmieri (2012) comparan las relaciones económicas y financieras, como el caso de la compra de bonos, con la carga dialógica y la relación interpersonal de confianza que hay tras el hecho de hacer una promesa. Ya que el objetivo principal de las finanzas es la generación de plusvalía mediante inversiones, los especialistas en economía afrontan la incertidumbre no solo mediante estudios cuantitativos, sino con información verbal. Estas comunicaciones se realizan, además de a través de la prensa, mediante otros textos cargados de datos y opinión como las *divulgaciones obligatorias*, que son prescritas por los entes reguladores del poder del mercado, o las *divulgaciones corporativas*, realizadas de manera estratégica y convenientes para las firmas comerciales. Además de esto, los autores también tienen en cuenta la información cuantitativa (p. ej., operaciones financieras, adquisiciones de acciones, etc.) las cuales son consideradas signos, y por tanto, entes comunicativos.

No es posible hablar del inglés económico o financiero sino a partir de la suma de varias disciplinas reunidas dentro de una sola familia de lenguas de especialidad que comparten, en gran parte, campo de trabajo. Son múltiples los campos de estudio y especialidades, destacando los siguientes: economía, contabilidad, finanzas, banca y comercio. Esto se justifica en que muchos de los términos y expresiones utilizadas en este dominio no son exclusivas de uno solo de estos subdominios, sino que son transversales y utilizados por varios de ellos a la vez. Verbos como *amortize* son utilizados indistintamente en contabilidad y finanzas, en relación con el eje horizontal de especialidad de Hofmann (1998), o sustantivos como “turnover”, que es utilizado en cuatro campos distintos: contabilidad, comercio y finanzas. Asimismo, existen diferencias diatópicas, en términos de Coseriu, en la selección de la terminología, contrastando especialmente entre las modalidades británica y estadounidense. De esta manera, por ejemplo, la tradición británica utiliza “debtors” para referirse a la cantidad de dinero que se debe a una empresa, mientras que los estadounidenses tendrán preferencia por el uso de “accounts receivable” (Nguyen Phuoc, 2012).

Antes de explicitar las distintas características del léxico es interesante describir de la fluida relación entre el lenguaje de la prensa y el lenguaje de la economía como discurso de divulgación. El lenguaje especializado de la economía no presenta ninguna homogeneidad, pues su grado de elaboración se nivela a partir del lenguaje teórico puro hacia los textos de ciencias económicas y posteriormente al periodismo económico, de ahí a la prensa cotidiana, pasando además por la lengua científica del momento entre colegas de especialidad que hacen uso de un argot. El objetivo principal del uso del lenguaje económico en prensa es divulgar los asuntos económicos que, aunque tienen un carácter altamente abstracto, afectan directamente a toda la población. La lengua económica juega un papel importante en la prensa escrita, así como en revistas y boletines de economía especializada, donde se encuentran textos publicitarios o informaciones expositivas financieras y bursátiles con un carácter altamente técnico que en gran medida debe transformarse de manera práctica para poder ser comprendido por sus destinatarios (Spillner, 2007). Además de plasmar las informaciones corporativas e institucionales que se ofrecen a través de ruedas de prensa y declaraciones, ofrecen informes e interpretaciones cargados de lenguaje evaluativo, de manera que frecuentemente se encuentran varios puntos de para tratar de nivelar la asimetría de información entre los emisores y el público. El poder de la prensa es notorio, ya que en gran medida el periodista

económico modela la información ofrecida por de los *datos crudos* e influencia al público con su punto de vista (Palmieri y Palmieri, 2012; Bushee et al., 2010).

El uso de la lengua se desliza a partir del eje vertical de Hofmann, que permite hacer un uso riguroso del lenguaje especializado y a su vez asequible para un público con niveles variables de especialización. Se siguen distintas estrategias que reformulan el discurso mediante estrategias lingüísticas de forma que no solo es posible hablar a través de la terminología, sino a través de todo el universo lingüístico que gira a su alrededor (oraciones, declaraciones, expresiones y evaluaciones). En relación con la economía, los periódicos en muchas ocasiones ofrecen distintas estrategias mediadoras y divulgativas que sirven para comprender el trasfondo de los textos, no solo a través de glosarios sino mediante parafraseo, conjuntos pregunta-respuesta o explicaciones teóricas (Janot, 2012).

El dominio económico-financiero de la lengua inglesa se caracteriza por emplear definiciones claramente delimitadas que convierten en términos a una gran serie de elementos léxicos no poco frecuentes en el dominio general. Así, es fácil encontrar términos económicos que tienen un significado especializado y en muchos casos polaridad distinta a la de la lengua general. Este es el caso de términos como “asset”, “balance”, “capital”, “interest”, “facilities”, o “honour” (todos estos positivos en el dominio general). Asimismo, el lenguaje económico es especialmente productivo en cuanto a colocaciones. Así, un sustantivo como “debt” hace uso específico de múltiples verbos que aportan un significado preciso, por ejemplo, “to service a debt”, “to forgive a debt”, “to retire a debt”, así como conjuntos de sustantivos como “debt forgiveness” o “debt restructuring”. Este tipo de construcciones son las más frecuentes en la terminología económica, especialmente a la hora de extender la terminología. Así, el significado del término *account* se precisa dentro de su estructura jerárquica a través de términos compuestos: “deposit account”, “current account”, “demand account”, “deposit account”, “dormant account”, “instant access account”, etc. Finalmente, quedan fuera de la clasificación otros compuestos frecuentes en los textos económicos que proceden de la metaforización de elementos culturales aplicados a determinados factores económicos como es el caso de: p. ej., “sleeping beauty”, “white knight” y “black knight” o “cash cow”, entre tantos.

En cuanto a las realizaciones morfológicas del lenguaje económico, se pueden encontrar los siguientes casos (Nguyen Phouc, 2011):

- Palabras, que pueden ser especializadas como “asset”, “capital”, “debit”, “earnings”, “fraud”, etc., así como palabras procedentes del léxico general que en el dominio específico cambia su significado, por ejemplo “friendly” es un adjetivo que en finanzas hace referencia sobre las intenciones de comprar o controlar una empresa.
- Expresiones multipalabra: además de las revisadas anteriormente, es necesario destacar las nominalizaciones de conjuntos verbo + preposición como “buy-back”, “buy-in”, “take-out”, “take-over”, “turn-over”, “write-back”, “write-down”, “write-off”, que fosilizan a los verbos frasales. De una naturaleza similar son las conversiones frasales de adverbios + sintagma nominal como “above-par”, “below-the-line”, etc. Otros patrones de colocación poco frecuentes en otros ámbitos y que destacan aquí son las construcciones sustantivo + adjetivo como “accounts payable”, “bills receivable”, “total repayable”. Finalmente, el dominio económico se caracteriza por incluir frases nominales premodificadas como “public sector borrowing requirement”, y posmodificadas, como “law of diminishing returns”, “return on capital employed”, “balance brought down”.
- Abreviaciones: muchas unidades multipalabra son utilizadas en el discurso frecuentemente a través de sus iniciales como p. ej., “A.G.M.” (“annual general meeting”), asimismo, muchas de ellas han pasado a lexicalizarse sin que muchos usuarios conozcan su procedencia p. ej., “Swift” que abrevia a “Society for Worldwide Interbank Financial Telecommunications”.

Para concluir, es interesante observar las características evaluativas del lenguaje económico. Existen escasas referencias bibliográficas sobre el estudio de la evaluación en este lenguaje especializado y se destaca Samson (2006). En este trabajo se revisan las propiedades evaluativas del discurso económico a nivel textual, principalmente la interacción con los oyentes para subrayar los puntos cruciales del texto y promover que lo afirmado es la interpretación correcta sobre un asunto determinado. Al contrario de lo que pueda parecer al público, el discurso económico contiene una gran cantidad de adjetivos que forman parte de construcciones de modo que se evidencia una organización compleja que se estructura claramente a medio camino entre la evaluación y el del metadiscurso. Aunque está difundida a lo largo del texto, los segmentos evaluativos suelen concentrarse a mitad y en las conclusiones. Asimismo, la autora destaca que en su discurso los economistas no tienden a remarcar sus posiciones de una manera drástica,



sino gradual a lo largo del desarrollo del texto. Dichos adjetivos están restringidos por el contexto, ya que se colocan con indicadores (p. ej., “sales”, “returns”, “rates”, etc.) directamente relacionados con modelos de economía discutidos y evaluados teniendo presente la postura de los oyentes.

## 4.8 Conclusiones

En este capítulo se ha tratado de aportar una panorámica de los lenguajes especializados enfocada al tratamiento sistematizado del léxico con el objetivo de conocer los factores principales implicados en la construcción de un lexicón de sentimiento de dominio específico. Tras una revisión superficial de la literatura en el ámbito de los lenguajes de especialidad, se ha realizado un recorrido por la historia de los desarrollos teóricos sobre terminología y su relación con la lingüística, la gestión terminológica. Finalmente se ha ofrecido una descripción del lenguaje de los dominios relevantes para este trabajo: el periodismo y la economía.

En el primer apartado se revisaron las fronteras, muchas veces difusas, entre el dominio general y las lenguas de especialidad. Si bien existen consideraciones que imponen mayores distancias entre dominios, desde hace dos décadas existe un consenso en que la diferencia está, principalmente, en los usos comunicativos en el campo de trabajo cada disciplina. La terminología sigue siendo un factor muy importante en la discusión, pero esta debe tenerse en cuenta como un factor más del estudio, junto con factores textuales y contextuales.

El modelo terminológico propuesto por la *teoría general de la terminología* de la Escuela de Viena propone un punto de vista conceptual monolítico en el que las definiciones son altamente normativas, unívocas y sincrónicas. El significado evaluativo o cualquier relación con el sentimiento está fuera de cuestión. Es decir, solo existirá una denominación para cada concepto y su definición será eminentemente neutral, sin ninguna consideración a su evolución a lo largo de los tiempos. Los nuevos trabajos surgidos en los albores del siglo XXI, destacando aquí la teoría comunicativa de Cabré y la socioterminología de Temmerman suponen un giro copernicano al tratamiento de lenguajes especializados y, además de las posibilidades que ofrecen los sistemas de transmisión del conocimiento, se anticipan a las distintas aplicaciones en el terreno del procesamiento del lenguaje natural, como la web semántica o el análisis de sentimiento. A partir de entonces, se consideraría que las unidades terminológicas compartirían



propiedades lingüísticas con el léxico general, diferenciándose en el modo de significación y en sus aspectos pragmáticos, sociales y cognitivos dependientes enteramente del discurso. Aquí reside el primer interés de esta tesis: se considera que los términos pueden contener más significados que la denominación (p.e. sentimiento) y que este no es sincrónico, pues en muchos casos el significado tendrá relación con el momento histórico y con el uso de sus hablantes.

La siguiente cuestión de la tesis es eminentemente técnica: la construcción de un lexicón de sentimiento especializado a partir de textos de la prensa. Con el fin de aportar una infraestructura teórica sólida, se han revisado los puntos más importantes de la literatura sobre la terminografía, que se encarga de estructurar sistemas conceptuales especializados y recopilar términos de manera estructurada. El desarrollo del tesoro, las ontologías y sus técnicas relacionadas han trascendido el dominio lingüístico en cuanto a la capacidad de estructuración compleja de la lengua y el conocimiento. Esto no sería posible sin el salto de gigante que ha supuesto la universalización de la informática en el tratamiento terminológico a partir de los años 90 y la revolución que supone en la actualidad la aplicación de las tecnologías relacionadas con el aprendizaje automático

Para terminar, los dos últimos apartados resumen la literatura sobre la descripción de los factores subyacentes correspondientes al léxico, organización textual y contextual de los dominios periodístico y económico. Su conocimiento aporta una perspectiva más amplia que permitirá llevar a cabo el trabajo de elaboración de un lexicón de sentimiento del dominio económico a partir de la riqueza textual encontrada en un amplio compendio de artículos periodísticos de actualidad económica. En el siguiente capítulo se estudiarán los pormenores del análisis de sentimiento como disciplina en boga dentro del procesamiento del lenguaje natural. Se trata de un viaje de ida y vuelta entre dos disciplinas pues, en relación con este apartado, se observarán las ventajas y desventajas de la aplicación del enfoque léxico, eminentemente terminográfico, con vistas a cumplir el objetivo de analizar la evolución temporal de la orientación semántica de los términos especializados a partir del procesamiento de la opinión de una serie de oraciones.

## 5. Análisis de sentimiento

« But man's capacities have never been measured ; nor are we to judge of what he can do by any precedents, so little have been tried. »  
Henry David Thoreau (1817-1862), *Walden* (1854).

### 5.1 Introducción y definición

El análisis de sentimiento (*sentiment analysis*), también llamado *minería de opinión* (*opinion mining*), es el campo del análisis textual que estudia las opiniones, sentimiento, valoraciones, actitudes, y emociones de individuos o de un público en relación con las distintas entidades y sus atributos que son expresadas en textos escritos (Liu, 2012). Son muchas las disciplinas que hacen uso de sus aplicaciones, entre otras: la lingüística, la psicología, las ciencias políticas o la economía. Se han utilizado, asimismo, muchas etiquetas para reflejar los distintos enfoques del análisis textual: lingüística computacional, procesamiento del lenguaje natural, recuperación de información, etc. Aunque la denominación original *sentiment analysis* apareciese tan solo al inicio del siglo XXI, ya existían estudios previos que proponían modelos computacionales relacionados con la opinión, los adjetivos de sentimiento o el análisis de subjetividad.

Dovring (1954) documenta el uso de técnicas de análisis textual en la Suecia del siglo XVIII, cuando se realizaban exámenes de censura mediante la búsqueda de selecciones léxicas en los himnos religiosos susceptibles de resultar amenazantes para la moral de la época. Durante la década de 1960, el grupo de investigación *General Inquirer* (GI) comenzó un proyecto para realizar una lista de palabras clasificadas de manera multicategoría para aplicarse en estudios de la rama de la psicología. Esta lista aún hoy sigue siendo utilizada en múltiples aplicaciones de procesamiento del lenguaje natural de todo tipo, especialmente de análisis de sentimiento. Con el advenimiento de la era de Internet y la producción y facilidad de acceso a grandes cantidades de texto de todo tipo, el término en inglés *sentiment analysis* comenzó a adquirir entidad propia en distintos estudios estadísticos y existe consenso en la literatura de que este fue utilizado por primera vez en el artículo “*Sentiment Analysis: Capturing Favorability Using Natural*

*Language Processing*” de Nasukawa y Yi (2003). Este fue, muy probablemente, el primer documento de trabajo relacionado con el PLN que intentó atribuir orientación semántica a temas, o *aspectos*, específicos en el texto. Paralelamente apareció la denominación *opinion mining*, término que se utiliza de manera casi equivalente, en el artículo “*Mining the peanut gallery: Opinion extraction and semantic classification of product reviews*” de Dave, Lawrence y Pennock (2003) que trataba sobre la identificación de las propiedades únicas de varios textos extraídos de la web para la detección automática de reseñas positivas y negativas. Otras denominaciones para la disciplina utilizadas han sido *extracción de valoraciones* o *minería de reseñas* (*review mining*) (Pang y Lee, 2008).

El estudio de las entidades con relevancia global y de la opinión del público siempre ha sido una cuestión destacada, especialmente a partir de los estudios de popularidad y de mercado llevados a cabo a partir de la segunda mitad del siglo XX. Mientras el número y los tipos de recursos informativos crecen exponencialmente debido al auge de Internet en el siglo XXI, los seres humanos afrontan hoy el reto de transformar la riqueza de esa información para desarrollar las disciplinas desde nuevas visiones y perspectivas. Aún actualmente, la demanda de opiniones en el Banco Central Europeo, o la valoración de los líderes políticos se lleva a cabo mediante sondeos tradicionales, a través de encuestas financiadas por organizaciones que se realizan de manera rutinaria o por encargo. De estas, solo un pequeño porcentaje trascendía finalmente al público en general y en muchos casos sus métodos de análisis de datos no son del todo transparentes. De hecho, en muchas ocasiones no se aportan facilidades de acceso a los datos y se obstaculiza el contraste de la información para la realización de interpretaciones alternativas. Entre los ejemplos más notorios están el *Barómetro del CIS* que se realiza mensualmente en España o el *Monthly Michigan Index of Consumer Sentiment* en los Estados Unidos (Batra y Rao, 2010; Yu, Duan y Cao, 2013).

En cuanto al uso de los términos *sentiment* y *opinion* (y sus equivalentes en lengua castellana), existen diferencias sutiles ya que coinciden semánticamente en muchos aspectos. Mientras que una opinión hace referencia a la visión concreta y racional de una persona, el sentimiento entra dentro del mundo de las sensaciones. Por ejemplo, Liu (2015) ilustra esta diferencia de una manera sencilla: “Estoy preocupado por la situación económica del país” en contraposición con “Pienso que la economía del país va fatal”. En la primera podríamos responder “Comparto tu sentir” mientras que en la segunda normalmente responderíamos “Estoy de acuerdo / en desacuerdo contigo”. Los segmentos que expresan sentimiento u opinión incluyen segmentos de carácter subjetivo,

en contraposición con los segmentos textuales objetivos que incluyen datos o hechos consumados. Asimismo, dichas oraciones objetivas también pueden incluir sentimientos positivos o negativos, al indicar hechos deseables o indeseables. Como disciplina, el análisis de sentimiento se ha centrado en detectar las opiniones e identificar su polaridad, de forma que se puedan cuantificar segmentos de sentimiento positivo o negativo.

El análisis de sentimiento, desde sus principios, se centra en extraer la información sobre las actitudes emocionales a partir del análisis de grandes corpus de documentos, especialmente desde las redes sociales. Comparado con los campos que le preceden (procesamiento del lenguaje natural y *affective computing*) los problemas de análisis se enfocan de una manera más directa y efectiva, ya que no se proponen métodos de comprensión completa y automatizada de los textos. Estos se tratan como conjuntos léxicos que incluyen características semánticas, p. ej., en este caso una polaridad positiva o negativa, dentro de un nivel de mayor o menor detalle, basadas en la presencia o ausencia de algunos determinados elementos léxicos y patrones gramaticales. Aunque algunos sistemas incluyen dentro de este conjunto al análisis de emociones (p. ej., Toret, 2015), clasificando las enunciaciones en un conjunto limitado de emociones (felicidad, ira, tristeza), los sistemas típicos de análisis de sentimiento no suelen abarcarse dentro de este ámbito. El resultado de un análisis de sentimiento suele ser una clasificación: las opiniones son etiquetadas, tras su análisis a través de distintos métodos, con una puntuación positiva, negativa o neutra (Waloszek y Waloszcek, 2017). Aún dentro del procesamiento del lenguaje natural, como disciplina de la lingüística computacional, se han realizado estudios en múltiples disciplinas a partir de las aplicaciones de análisis de sentimiento. La atención levantada por el análisis de sentimiento justifica el rápido paso de disciplina meramente científica, hacia una enorme variedad de aplicaciones en una suma dispar de estudios y aplicaciones comerciales de todo tipo. La mayoría de las compañías dedicadas a la inteligencia de mercado (por ejemplo, Microsoft) ofrecen sus propias soluciones comerciales para la minería de opinión y muchas otras (Amazon, Facebook, etc.) cuentan con sus propios sistemas internos de análisis.

En relación con la lingüística, en términos de comprensión del lenguaje natural, el análisis de sentimiento se puede reconocer como una subárea del análisis semántico ya que el interés recae en observar el significado que ofrece la orientación semántica de los textos. Existe una amplia gama de aplicaciones de esta disciplina: análisis de volumen de ventas, destinos turísticos (Moreno-Ortiz et. al, 2010) o evaluar las actitudes de los inversores (Bollen et al., 2011). Asimismo, los análisis de sentimiento se han aplicado en

una gran cantidad de ocasiones en el estudio del debate político (Wang et al., 2012) o como factor para predecir los resultados de elecciones a partir de los comentarios realizados por los usuarios en la red (Tumitan y Becker, 2014). Otros campos de aplicación han sido estudios del discurso político (Gallardo Paúls, 2017), análisis periodísticos (Godbole et al., 2007), jurídicos (Garofalo, 2018), estudios en psicología (Tao et al., 2016) o la educación (Fernández-Cruz, 2017).

## 5.2 Propiedades del análisis de sentimiento

El sentimiento es la sensación subyacente, actitud, evaluación o emoción asociada con una opinión. Dentro de la quintupla del sentimiento (*e, a, s, h, t*) ya revisada en el Capítulo 3 (Liu, 2015), el sentimiento es el factor central e indispensable para cualquier tipo de análisis de orientación semántica. Mientras, el resto de los factores se amoldará a las necesidades del estudio y a las posibilidades que ofrecen los distintos sistemas o plataformas de análisis. Un esquema de la clasificación se puede consultar en la Figura 8. Los distintos tipos de estudios sobre análisis de sentimiento pueden resumirse de la siguiente manera:

- Enfoques de trabajo: los estudios de análisis de sentimiento cubren dos amplios campos. El primero es de carácter teórico y experimental que abarca dos disciplinas principales: la lingüística y la psicología. En cuanto a la lingüística, hay necesidad de expandir el conocimiento sobre las propiedades globales del lenguaje humano en relación con la expresión de la valoración, mientras los estudios en psicología abarcan observaciones relacionadas con la subjetividad y las emociones en los textos. El segundo es de carácter más sociológico, y en muchos casos, con vocación comercial, con vistas a conocer más sobre las características y propiedades de productos y marcas. De hecho, existen multitud de investigaciones relacionadas con la actitud del ciudadano y el consumidor ante los productos.
- Granularidad del análisis: los sistemas de análisis de sentimiento tienen varios niveles de detalle en a la hora de analizar los textos: documento, oración y aspecto.
- Métodos de análisis. Concurren dos enfoques principales: aprendizaje automático (*machine learning*), relacionado con la construcción de modelos analíticos, y otro basado en extensos lexicones (*lexicon-based*) que incluyen unidades léxicas de todo tipo previamente etiquetadas de manera manual o semiautomática. Muchos

sistemas, además, desempeñan propuestas *híbridas* que utilizan técnicas características de ambos enfoques. Se puede ver un esquema en la Figura 8.

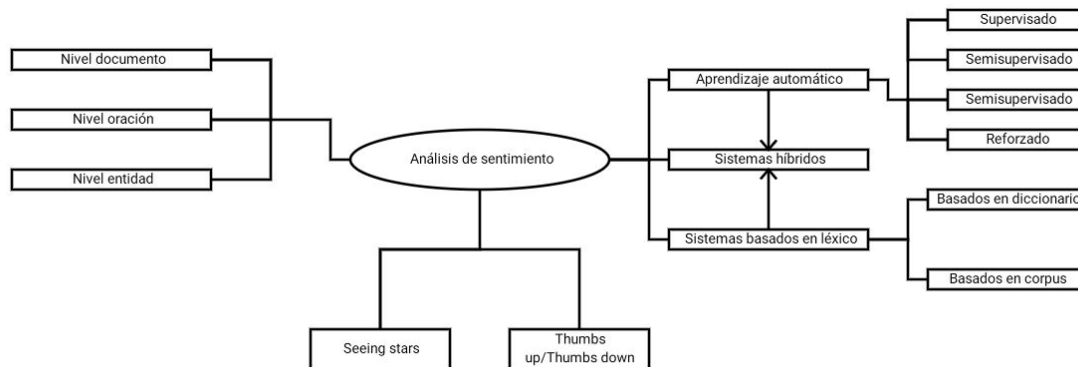
- La orientación semántica: puede ser positiva, negativa, y en muchos casos, neutra. En la literatura, se le suele referenciar con matices como sentimiento, valencia, polaridad u orientación semántica.
- Nivel de detalle o intensidad de sentimiento: además de un enfoque que asigne etiquetas categóricas a los resultados del análisis textual (enfoque *Thumbs up/Thumbs down*), los sentimientos de cualquier tipo pueden tener distintos niveles de intensidad (enfoque *Seeing stars*). Este último enfoque asigna una puntuación al sentimiento de un texto analizado dentro de una escala (Pang y Lee, 2005). Para hacer referencia a estas puntuaciones se utilizará el término polaridad. Por ejemplo, “good” es más leve que “excellent” y “terrible” tiene mayor intensidad que “bad”.

Para aplicar los enfoques de manera efectiva, Liu (2015) recoge que la clave para la clasificación del sentimiento es la ingeniería de características efectivas (*engineering of effective features*). Aunque estas características no son exclusivas solo de este enfoque son las siguientes:

- Términos y frecuencia: se trata de palabras individuales (*unigrams*) y sus *n-grams* asociadas a recuentos de frecuencias, algo también conocido como *Bag-of-Words*. Son las características más usadas en la clasificación textual tradicional que se basa en distintas temáticas.
- Parte de la oración (*Part of Speech*, POS): los elementos se etiquetan según la parte de la oración a la que correspondan a través de distintos algoritmos (p. ej., *PennTree-Tagger*). Los adjetivos principalmente, y también, adverbios, sustantivos y verbos, son los indicadores de sentimiento más importantes.
- Palabras y sintagmas con carga de sentimiento: además de las palabras individuales, existen modismos y otro tipo de expresiones multipalabra, como por ejemplo “cost somebody an arm and a leg” que en conjunto tendría una carga negativa.
- Reglas de opinión o reglas contextuales y modificadores de sentimiento (*sentiment shifters*): además de palabras y frases con carga semántica, hay otros constructos o composiciones del lenguaje que se pueden utilizar para expresar

sentimiento u opinión. Además, existen expresiones que se usan para cambiar la orientación del sentimiento, por ejemplo, de positivo a negativo o viceversa.

- Dependencia semántica: como se pudo observar en el apartado de gramáticas locales, determinadas estructuras gramaticales están basadas en la dependencia de la clasificación de las palabras y están estructuradas a través de árboles de dependencia.



**Figura 8** Resumen de las principales características del análisis de sentimiento

## 5.3 Descripción de los principales niveles de clasificación

### 5.3.1 Clasificación de sentimiento a nivel de documento

La clasificación de sentimiento a nivel documento se considera la tarea más simple de análisis de sentimiento. Sencillamente se clasifica un documento como positivo o negativo o lo sitúa en una escala de mayor granularidad. En este tipo de análisis no es necesario detectar entidades o aspectos.

Existe una amplia gama de algoritmos clasificadores susceptibles de realizar, entre otras, tareas de análisis de sentimiento. Un ejemplo de esto sería la clasificación a través de distintos idiomas y/o distintos dominios del lenguaje. Este tipo de clasificadores se han aplicado generalmente a partir de textos preclasificados utilizados como datos de entrenamiento, como reseñas comerciales de un solo producto, ya que se evaluaría tan solo a una única entidad.

Los sistemas que utilizan este enfoque no son recomendables para aplicar a p. ej., discusiones en foros de internet, entrevistas o noticias pues en ellas suelen entremezclarse múltiples entidades. Incluso si, dada una clasificación previa, se tiene la certeza de que un documento evalúa exclusivamente una única entidad, este tipo de sistemas aportará una única puntuación generalizada de sentimiento. Por tanto, a partir de los resultados



obtenidos será difícil determinar con precisión la opinión sobre los distintos aspectos que componen dicha entidad.

Una traducción de esta problemática a términos formales, dentro del formato descrito en el apartado 3.8.2 sería la siguiente: la clasificación de sentimiento a nivel documento asume que un documento de opinión  $d$  expresa el sentimiento  $s$  sobre una entidad general  $e$  y que solo contiene opiniones de un único emisor de opinión  $h$ . Según Liu, (2015), aquí, al sentimiento  $s$  se le asignaría a un aspecto “GENERAL” dentro de la quintupla de opinión. En cuanto a la entidad, el emisor de opinión y el tiempo son irrelevantes para el estudio, o directamente desconocidos.

### 5.3.2 Clasificación de sentimiento a nivel oración

Como se ha revisado en el apartado anterior, la clasificación de sentimiento a nivel de documento ofrece resultados con bajo nivel de detalle en relación con los requisitos de muchas aplicaciones prácticas. Por ejemplo, cuando en un documento se hace referencia a distintos objetos de opinión (*targets*) es necesario trascender el nivel documento y obtener mayores niveles de detalle. Como es el caso de esta investigación, es factible tomar el análisis de sentimiento a nivel oración como un análisis a nivel de documento a partir de documentos mono-oracionales. Sin embargo, Liu (2012) considera que esta tarea es más compleja ya que la información contenida en una oración es mucho más corta que la contenida en un documento típico que contiene múltiples oraciones. Muchos de los sistemas que realizan su clasificación a este nivel, por lo general, tienen dificultades para clasificar de manera precisa el sentimiento en tres clases básicas (positivo, negativo y neutro) y con frecuencia obvian la clasificación de oraciones neutras, tendiendo a *positivizar* o *negativizar* la puntuación. Sin embargo, esto no puede ser ignorado porque un documento puede contener muchas oraciones que no expresen sentimiento ya que un resultado neutro suele, aunque no exclusivamente, significar factualidad o ausencia sentimiento.

La clasificación a nivel oración es un paso intermedio en la tarea general de análisis de sentimiento y no considera varios factores de la quintupla de análisis de sentimiento: el objeto de opinión (*opinion target*), detección de la entidad y aspecto, emisor de opinión o el tiempo. De esta forma el problema de la clasificación a nivel oración queda clasificado formalmente “given a sentence  $x$ , determine whether  $x$  expresses a positive, negative or neutral (or no) opinion” (Liu, 2015 p. 70). A efectos



prácticos, la mayoría de las oraciones expresan un único sentimiento por lo que en la mayoría de las tareas de análisis de sentimiento no necesitan determinar el aspecto de las entidades, ni el emisor de opinión y, además, estos elementos son fáciles de detectar mediante técnicas de clasificación textual previa.

Sentimiento y subjetividad no son términos equivalentes a pesar de ser atravesados por una línea secante. La subjetividad es un término clave que se ha utilizado ampliamente en el análisis de sentimiento a nivel oracional y su uso suele ser objeto de confusión. En muchas publicaciones considerar una oración como subjetiva y suponer que contiene una carga de sentimiento eran consideradas conceptualmente equivalentes. Sin embargo, estos conceptos no son equivalentes. Por eso es necesario reiterar que la definición de oración como subjetiva se dará cuando esta proceda de una perspectiva mental de un individuo en lugar de una experiencia contrastable y procedente de la experiencia en el mundo físico. Por lo tanto, mientras que una oración objetiva ofrece la información de manera factual, una oración subjetiva expresará sentimientos personales, juicios o creencias. La tarea de determinar si una oración es objetiva o subjetiva se llama *clasificación de subjetividad* (Riloff y Wiebe, 2003; Wiebe, 1999).

Se debe tener en cuenta que las expresiones subjetivas relacionadas con el sentimiento expresan apreciación, evaluación, deseo, creencia, sospecha, especulación, etc. (Wiebe, 2000). Como contrapunto, se ha de tener en cuenta que existen oraciones subjetivas que no contienen opinión ni sentimiento. Además, las oraciones objetivas pueden implicar opiniones o sentimientos causados por hechos deseables o indeseables. Por ejemplo:

Floor collapses at Jakarta stock exchange, dozens injured.<sup>26</sup>

Por ello, Liu (2015) considera obviar esta problemática y plantea diferenciar preferiblemente entre *opinionated* (si implica un sentimiento positivo o negativo), que se traducirá aquí como ‘con carga de opinión’ o si lo expresado no incluye un sentimiento positivo o negativo, esto es, *non-opinionated* ‘sin carga de opinión’.

---

<sup>26</sup> Ejemplo obtenido de CNN: <https://edition.cnn.com/2018/01/15/asia/jakarta-stock-exchange-intl/index.html>

Considerando estos factores, la clasificación a nivel oración se lleva a cabo mediante la resolución de dos problemas: el primero es la clasificación de subjetividad y se trata de catalogar si una oración expresa una opinión o si por el contrario se trata de una información objetiva o factual. El segundo problema sería ya la clasificación de sentimiento y trataría de clasificar esas opiniones en clases positivas o negativas. Entre las implementaciones prácticas, existen varios modelos basados en aprendizaje automático o en lexicón, destacando los propuestos por Wiebe et. al. (1999) o Riloff y Wiebe (2003) basados en patrones sintácticos presumibles ofrecer perspectivas subjetivas.

Un sistema de análisis de sentimiento a nivel oración presume que una oración contiene una única expresión de sentimiento. Suponer esto también impone varias restricciones que complejizan la aplicación de este análisis:

- (a) Esta presunción funciona bien en oraciones simples (sujeto-verbo-objeto) con un único sentimiento y no es apropiada para aquellas que incluyen varias opiniones, ni oraciones complejas. En el caso de que una oración tenga varios sentimientos, puede asignársele una clase comodín llamada *MIXED* ‘MIXTA’ (Liu, 2015).
- (b) Se debe detectar un tono general positivo o negativo, pero puede obviar detalles. Por ejemplo.

HSBC keeps increasing dividends despite regulatory uncertainties.

El sentimiento positivo tiene mayor peso que el negativo. Existen sistemas como Moreno-Ortiz (2017) que ofrecen soluciones para ponderar los elementos positivos y negativos en un mismo segmento.

La clasificación a nivel oración solo se puede aplicar a las oraciones que expresan opiniones de carácter común pero no puede aplicarse a oraciones que expresen opiniones comparativas. La oración del ejemplo a continuación expresa sentimiento de una manera clara, sin embargo, no es posible clasificarla eficazmente a través de una única puntuación positiva, negativa o neutra.

A reloadable debit card works better than a bank account.

Para concluir, la clasificación de sentimiento a nivel oración se aproxima a solucionar el problema de relacionar los objetivos de opinión con su sentimiento de una manera más eficiente y detallada que la clasificación a nivel documento. Sin embargo, su aplicación práctica tiene una serie de limitaciones. El problema es inverso al análisis a

nivel de aspecto que se verá a continuación. Debido al escaso nivel de detalle, no es posible conocer más características de la opinión, como por ejemplo qué entidades y aspectos son considerados positivos o negativos. Aunque se puede realizar una generalización al disponer del sentimiento de los targets de opinión, en oraciones que incluyan dos objetos de sentimiento dispares es muy difícil otorgar una única puntuación de polaridad. Finalmente, este enfoque no puede aplicarse en oraciones comparativas u oraciones complejas que incluyan varias opiniones. Para superar algunas de estas limitaciones, se puede profundizar en el análisis a través del nivel aspecto que se analizará a continuación.

### 5.3.3 Clasificación de sentimiento a nivel aspecto

El análisis de sentimiento aspecto supera el nivel de detalle de la clasificación de sentimiento a nivel documento u oración. La razón subyacente de este enfoque parte de que una puntuación negativa de un documento no implica necesariamente que el mensaje que ofrece el documento sea totalmente pesimista con respecto a todos los aspectos encontrados en el documento. Este enfoque se centra particularmente en sintagmas en lugar de palabras, ya que aportar los fragmentos más significativos del texto. Cuando se presenta un documento *d* determinado, existen distintos asuntos o entidades dentro del texto que son evaluados y que, dentro de su enunciación, pueden portar opinión. Se entiende como entidad a cada uno de los temas, personas, fenómenos o eventos sobre los que queremos conocer su sentimiento de manera automatizada (Liu, 2015). En distintos dominios de aplicación, el análisis de sentimiento basado en aspecto se le conoce con otros nombres, al no existir una terminología unificada al respecto. La literatura propone términos como *análisis de sentimiento basado en temas (topic based sentiment analysis)* o *análisis de sentimiento basado en entidades (entity-based sentiment analysis)*.

La categoría de una entidad (*entity category*) representa a una entidad con un nombre exclusivo para uso del clasificador de sentimiento, por ejemplo, la entidad RESTAURANTE. Esta categoría puede ser referida a través de distintas expresiones o menciones de entidad (*entity expression* o *entity mention*) realizadas mediante palabras o frases sinónimas; por ejemplo, dentro de la categoría de entidad RESTAURANTE entrarían las expresiones de entidad *hamburguesería*, *fonda* o *cevichería*. El proceso de agrupar las expresiones de entidad en categorías de entidad se llama *agrupación o resolución de entidades (entity resolution* o *entity grouping)*. Para la extracción de

aspectos de cada entidad, el problema es básicamente el mismo que para las entidades y utiliza los conceptos categoría de aspecto (*aspect category*) y expresión de aspecto (*aspect expression*). Por ejemplo, *color*, *frame rate* y *resolution* se refieren al mismo aspecto (VIDEO) para las cámaras de video. Asimismo, el proceso de agrupar expresiones de aspecto en categorías de aspecto se llama resolución de aspecto (*aspect resolution* o *aspect grouping*) (Henríquez Miranda y Guzmán Luna, 2016; Liu, 2015).

Las expresiones de aspecto suelen ser sustantivos o sintagmas nominales, pero también sintagmas verbales, adjetivos, adverbios y otras construcciones. Las expresiones de aspecto explícitas (*explicit aspect expressions*) sería las expresiones de aspecto que aparecen abiertamente en el texto como sustantivo o sintagma nominal. Muchos de las expresiones de aspecto implícitas son adjetivos o adverbios que califican o describen categorías de aspectos específicos. Por ejemplo, el aspecto “Calidad de imagen” de la entidad CÁMARA: “La calidad de imagen de esta cámara es genial”. Las *expresiones de aspecto implícitas* (*implicit aspect impression*) serán las que no son de tipo sintagma nominal o sustantivo. Su detección suele ser un desafío a la hora de desarrollar una implementación e incluso de extraer de manera manual.

Con la base de la discusión previa, es posible definir un modelo de entidad y un modelo de opinión según lo establecido por Liu (2011) y resumir las principales tareas de la disciplina del análisis de sentimiento basado en aspecto:

- Un *modelo de entidad* (*model of entity*) es una entidad que se representa en términos absolutos y contiene un conjunto finito de aspectos  $A=\{a_1, a_2, a_3, \dots\}$  de manera que se puede expresar de manera textual a través de un conjunto finito de entidades de expresión  $\{ae_1, ae_2, \dots, ae_r\}$ . Cada aspecto se puede expresar con cualesquiera de los elementos de un conjunto finito de sus expresiones de aspecto.
- Un *documento de opinión* (*model of opinion document*): se define como un documento  $d$  que contiene opiniones de un conjunto finito de entidades  $\{e_1, e_2, e_3, \dots, e_r\}$  y cada entidad incluye un subconjunto de aspectos. Cada una de las opiniones formará parte de un conjunto finito de portadores de opinión  $\{h_1, h_2, h_3, \dots\}$  que tienen lugar en un tiempo determinado  $t$ .

Teniendo en cuenta las definiciones anteriores, un sistema de análisis de sentimiento realizará al menos una serie de seis tareas principales:

1. Extracción de entidad y su resolución: extraer todas las expresiones de entidad en un documento  $d$ , así como reconocer y agrupar bajo una sola entidad a las expresiones de entidad sinónimas.
2. Extracción de aspecto y su resolución: extraer todas las expresiones de aspecto de cada una de las entidades y agrupar las expresiones de aspecto sinónimas en grupos.
3. Extracción del emisor de opinión y su resolución: resolver dentro la expresión quién es emisor de opinión de cada una de las opiniones del texto los datos estructurados y agruparlos de manera análoga a las tareas 1 y 2.
4. Extracción del tiempo: extraer el tiempo de publicación de cada publicación y homogeneizar los distintos formatos de tiempo de acuerdo con los estándares, p. ej.: *Sun, 15 May 2011 19:00:30 GMT*.
5. Clasificación de sentimiento: Al clasificar determinamos si una opinión sobre un determinado aspecto  $a$  perteneciente a una entidad  $e$  es positiva, negativa o neutra en el caso de una clasificación *Thumbs up/Thumbs down* (Pang et al., 2002). En el caso de una clasificación tipo *Seeing stars* (Pang et al., 2005) se le asignará una valencia numérica en una escala (p. ej., de 0 a 5, siendo 5 la puntuación máxima).
6. Generación de la quintupla de opinión: producir todas las quintuplas de opinión  $(e, a, s, h, t)$  expresadas en  $d$  basadas en los resultados de las tareas 1-5.

Llevar a cabo esta serie de tareas es una tarea de gran complejidad que requiere de clasificadores de sentimiento que sean diseñados *ad hoc* y combinen las distintas técnicas y herramientas aquí expuestas. En un análisis más avanzado, se contemplan dos tareas adicionales análogas a la tarea 2 (Liu, 2015):

7. *Extracción del motivo de la opinión*: extraer las expresiones de motivo para cada una de las opiniones y agruparlas en clústeres.
8. *Calificador de opinión*: extraer las expresiones de calificación para cada opinión y agrupar todas estas calificaciones en clústers.

Jindal y Liu (2006) destacan la problemática de la clasificación de los tipos de opinión a la hora de generar un sistema de clasificación. El tipo más simple y frecuente de opinión expresa un sentimiento sobre una entidad determinada o un aspecto concreto de una entidad, por ejemplo, “Economy is doing well”. Por otro lado, una opinión

comparativa es difícil de clasificar con una opinión final, pues se realiza una labor de contraste entre varias entidades que comparten aspectos. Por ejemplo, en el caso de “Debt management in Greece has been worse than in Argentina”, el aspecto es debt management y compara dos entidades distintas que comparten aspecto: Greece y Argentina. En ellas, la puntuación final del aspecto *management* de la entidad *debt* no sería concluyente para un clasificador automático.

Incluso si se tuviese la certeza de que en un proyecto de análisis de sentimiento cada documento evaluase exclusivamente a una entidad, esto no necesariamente se traduciría en que en una evaluación de sentimiento positivo el emisor de opinión considere como tales todos y cada uno de los aspectos de las entidades. Esta es la razón de ser del análisis de sentimiento a nivel aspecto: conocer con detalle el sentimiento atribuido a cada parte de una entidad. De hecho, de la quintupla de sentimiento, la tarea más compleja aquí es la extracción de aspecto.

Es pertinente, entonces, conocer algunas aplicaciones prácticas de este modelo. La clasificación de sentimiento a nivel aspecto también múltiples formas de implementación (p. ej., a través de aprendizaje automático, sintaxis u ontologías). La tarea de reconocer las distintas entidades en el texto parte de la tarea de procesamiento del lenguaje natural relacionada con la recuperación de información (*information retrieval*) llamada extracción o reconocimiento de entidades nombradas (NER, *named entity recognition / extraction*) (p. ej., Batra y Rao, 2009). El proyecto más conocido y ampliamente utilizado en la actualidad es el *Stanford NER* (Finkel, Grenager y Manning, 2005) que etiqueta las secuencias de palabras de un texto que referencia a los nombres de las cosas, desde nombres de personas y empresas a términos de biología, química y genética. Viene con extractores de funciones bien diseñados para el reconocimiento de entidades nombradas, y muchas opciones para definir extractores de características. Tiene tres categorías principales (PERSONA, ORGANIZACIÓN, UBICACIÓN), y además ofrecen varios modelos para diferentes idiomas y circunstancias.

Conseguir extraer el objeto de opinión es un desafío ya que no existen muchos modelos capaces de relacionar de manera efectiva entidades y/o aspectos con el objeto de opinión. Liu (2015) propone aproximarse a esto de la siguiente forma:

1. Generar un conjunto de características dependientes jerárquicamente de la entidad de estudio.
2. Determinar el alcance de cada expresión de sentimiento para determinar si esta cubre la entidad destino o el aspecto de la oración. Por ejemplo, en “Apple is doing

very well in this bad economy”, el alcance de aplicación de la palabra de sentimiento “bad” cubre solo a “economy”, pero no a “Apple”.

La propuesta de Boiy y Moens (2009) computa el peso de cada una de las características mediante el cálculo de la distancia entre la palabra que representa la entidad o el aspecto y el objeto de opinión relacionado. Diferencia entre dos distancias (1): diferencia de profundidad: el peso de las características es inversamente proporcional a la diferencia en profundidad en la palabra que representa la característica y la entidad observada en el árbol sintáctico. (2) Distancia de ruta: si el árbol sintáctico se observa como un grafo, el paso de una palabra que representa a la característica es inversamente proporcional a la longitud de la ruta entre la característica y la entidad de interés usando una búsqueda en anchura.

Otra solución para este problema se puede llevar a cabo a través de ontologías (Cadilhac et al., 2010; Freitas y Vieira, 2013; Lau et al., 2014). La ontología es necesaria en el análisis de sentimiento *basado en aspecto* porque se usa para definir las entidades y aspectos relevantes, darles estructura y definir restricciones de aplicación. Es posible simplificar la estructura ontológica para facilitar la tarea de análisis. Por ejemplo, cuando la opinión se expresa sobre una característica específica del objeto, entonces se entiende claramente que ese sentimiento es específico al objeto y no a la totalidad. Esto se puede ilustrar con la siguiente oración: “Greece’s economic figures are terrible”, donde la entidad “Greece” tendría el atributo “economic figures”. El sentimiento negativo no va dirigido a toda Grecia, sino tan solo a su economía. En el mundo real, sin embargo, ocurre que los hablantes a menudo están habituados a la existencia de la característica y la considerarán como un elemento independiente de la entidad. Debido a esto, en muchos casos no la observarán directamente a través de una estructura en árbol completa, sino a través de la simplificación de esta jerarquía dividiendo la entidad inicial en múltiples entidades (Pawar et al., 2016).

Se puede concluir que el análisis de sentimiento a nivel aspecto es el nivel más refinado, ya que incluye en su análisis todas las variables relativas a la opinión, así como su división en aspectos y subaspectos propone el desarrollo estructural completo de la disciplina. Resolver cada uno de estos factores a partir de lenguaje natural no estructurado implica un gran desafío a nivel de desarrollo computacional y de teoría lingüística. En la práctica, la mayoría de los problemas no necesitan identificar todos y cada uno (o no en todas las ocasiones) de los elementos de la quintupla para realizar un análisis de



sentimiento eficiente ya que algunos elementos pueden ser desconocidos, estar implícitos pragmáticamente o ser directamente irrelevantes.

## 5.4 Enfoques de análisis de sentimiento

Se procede entonces a desgranar las características de los grandes bloques clasificatorios en el análisis de sentimiento: el enfoque de *aprendizaje automático*, el enfoque *basado en léxico* y los *sistemas híbridos* que combinan características de ambos. El enfoque de aprendizaje automático aplica conocidos algoritmos relacionados con la inteligencia artificial en combinación con características lingüísticas, entre las técnicas estadísticas más frecuentes están *naïve Bayes* y *Support Vector Machines* que son paradigma de los métodos de aprendizaje supervisados y no supervisados.

El *enfoque basado en léxico* se fundamenta en diccionarios de sentimiento, es decir, extensas colecciones relacionales que recopilan entradas y valores de sentimiento. Existen dos ramas según la forma de adquirir el léxico: un enfoque basado en diccionarios y un enfoque basado en corpus. El *enfoque basado en el diccionario* que depende de buscar palabras clave de opinión, y luego busca en el diccionario sus sinónimos y antónimos. Y finalmente el *enfoque basado en corpus*, el cual comienza con una lista inicial de palabras de opinión, y luego encuentra otras palabras de opinión en un corpus grande para ayudar a encontrar palabras de opinión con orientaciones específicas del contexto. Esto podría hacerse mediante el uso de métodos estadísticos o semánticos. Hay una breve explicación de los algoritmos de ambos enfoques y artículos relacionados en las siguientes secciones (Medhat, 2014).

Finalmente, la tercera clasificación correspondería a *enfoques híbridos* que combinan las potencias de ambos enfoques, al ser muy común que los léxicos de sentimiento jueguen un papel clave en la mayoría de los métodos.

### 5.4.1 Aprendizaje automático (*Machine learning*)

Los métodos de *aprendizaje automático* (*machine learning*) están relacionados con la informática y la inteligencia artificial y su objetivo es desarrollar técnicas que permitan que los sistemas “aprendan” de datos con el fin de identificar patrones y clasificar información con mínima intervención humana a través de algoritmos y modelos estadísticos. Un modelo de aprendizaje automático dependerá de la existencia de



conjuntos de datos etiquetados y una selección de algoritmos de clasificación automática (Liu, 2015). Un proceso de aprendizaje automático típico consta cuatro pasos:

1. Preprocesado: en el que se normalizan los datos a los formatos requeridos por los sistemas y posteriormente se fragmentarán los datos de forma aleatoria.
2. Aprendizaje: se selecciona un algoritmo de aprendizaje y se genera un modelo a partir de los datos de entrenamiento en el sistema.
3. Evaluación: mediante distintas técnicas se evalúan y comparan los resultados de los distintos algoritmos de aprendizaje para determinar cuál es el mejor método de análisis.
4. Predicción: una vez establecido cuál es el modelo con mejor rendimiento, se proceden a la clasificación automática de los datos.

Los conjuntos de datos necesarios para diseñar un modelo son al menos dos: un *conjunto de datos de entrenamiento* etiquetados previamente (*training set* o *training data*) y el *conjunto de datos de prueba* (*test set* o *test data*) que se utiliza para proporcionar una evaluación imparcial de un ajuste final del modelo en el conjunto de datos de entrenamiento. Si los datos en el conjunto de datos de prueba nunca se han utilizado en el entrenamiento del modelo, el conjunto de datos de prueba también se denomina *conjunto de datos de reserva* (*holdout dataset*).

Un *conjunto de datos* determinado tendrá un número de filas (*observación*) compuesta de una o varias columnas o *características* (*features*). Los *conjuntos de datos de entrenamiento* incluirán además una columna extra: la *etiqueta* u *objetivo* (*label*) que será el valor que se pretende predecir. El modelo se *entrenará*, es decir, “observará” y “aprenderá” de estos datos. Para ilustrar esto, un *conjunto de datos de entrenamiento* típico será un fichero de texto plano con una cantidad considerable de *observaciones* (filas) y cada una de ellas incluirá al menos dos columnas: la primera incluirá el texto que debe ser analizado y la segunda una asignación generalmente manual de una *etiqueta* de valencia para dicha oración (p.e. “*pos, neg, neu*”).

Existen algunas variaciones sobre la clasificación de los tipos de algoritmos de aprendizaje automático, pero frecuentemente se pueden dividir en categorías según su propósito. Estas son las siguientes:

- *Aprendizaje supervisado* (*supervised learning*).
- *Aprendizaje no supervisado* (*unsupervised learning*).

- *Aprendizaje semisupervisado (semi-supervised learning)*.
- *Aprendizaje reforzado (reinforcement learning)*.

En un proceso de *aprendizaje supervisado (supervised learning)* los modelos aprenden de los conjuntos de *datos de entrenamiento* tras realizar ajustes en sus parámetros que mejor se adapten a los datos de entrada. Los algoritmos de aprendizaje supervisado tratan de modelar relaciones y dependencias entre las características de entrada y las etiquetas, de modo que se puedan predecir los valores de salida para los nuevos datos. Una vez que el modelo se entrena adecuadamente a través de un *modelo de evaluación*, el modelo podrá etiquetar nuevos conjuntos de datos no catalogados (Medhat et al., 2014).

Los principales modelos de aprendizaje supervisado incluyen algoritmos de *clasificación* (valores discretos, p.e. etiquetar una oración como “pos” / “neg”) y *regresión* (valores continuos, p.e. asignar una valencia a una oración en una escala de 0-100). Un modelo de *clasificación* encontrará las relaciones entre las *características* de cada observación y su etiqueta, y establecerá la línea divisoria entre las distintas clases. Así, al ser alimentado con nuevos datos, el modelo será capaz de determinar la clase a la que pertenecen, de acuerdo con sus características. Un modelo de *regresión* tomará una serie de características predictoras (texto) y una característica de respuesta numérica (puntuación de sentimiento) como resultado. El algoritmo tratará de encontrar una relación entre dichas variables que provea un resultado consistente y continuo. Algunos algoritmos utilizados frecuentemente son: *K-Nearest Neighbor*, *Naïve Bayes*, *Regresión Linear*, etc.

En los procesos de *aprendizaje no supervisado*, el sistema está entrenado con datos no etiquetados. Esta familia de algoritmos se utiliza principalmente en la detección de patrones y modelos descriptivos sin la referencia de variable de etiquetado conocidas. Sin embargo, aquí no hay etiquetas de salida en las cuales el algoritmo puede tratar de modelar relaciones. Los algoritmos aplicarán distintas técnicas en los datos de entrada para buscar reglas, detectar patrones y agrupar los puntos de datos que ayudan a obtener información significativa y describir mejor los datos al analista. Algunos algoritmos más frecuentes son *K-means clustering* o *Association Rules*.

El *aprendizaje semisupervisado* es un punto intermedio entre los dos modelos anteriores ya que también utiliza datos sin etiquetar para su entrenamiento. En muchos casos, el costo de etiquetar textos es bastante alto, ya que requiere de labor humana

especializada. Estos métodos aprovechan la idea de que estos datos contienen información importante sobre los parámetros del grupo, aunque se desconoce la pertenencia a grupos de los datos sin etiquetar. Por lo tanto, este tipo de modelos es idóneo cuando se cuenta con un conjunto de datos etiquetado parcialmente.

El *aprendizaje reforzado* tiene como objetivo utilizar las observaciones recopiladas de manera conductista a través de la interacción con el entorno para tomar medidas que maximicen la recompensa o minimicen el riesgo. El algoritmo de aprendizaje de refuerzo (*agente*) aprende iterativamente de sus experiencias en el análisis textual hasta que explora toda la gama de estados posibles. Así, el sistema es capaz de determinar automáticamente la categorización ideal dentro de un contexto específico, con el fin de optimizar su rendimiento y mejorar el uso contextual del análisis de sentimiento. Este modelo se ha utilizado en análisis multimodales de sentimiento que combinan el análisis del léxico utilizado en videoconferencias con la detección de expresión facial de la emoción. Considerando la palabra “crazy”, la unidad léxica puede tener un sentimiento positivo si está acompañada por una sonrisa o puede tener un sentimiento negativo si está acompañada por un ceño fruncido. Al mismo tiempo, “great” acompañada de un ceño fruncido implicaría un discurso sarcástico y, por tanto, negativo (Chen *et al*, 2017).

Los modelos de clasificación de aprendizaje automático se pueden evaluar mediante diversas técnicas, destacando el método *Holdout*, que divide el conjunto de datos etiquetado en un conjunto de entrenamiento y otro de prueba (por convención, 2/3 conjunto de datos de entrenamiento y 1/3 designación de conjunto de datos de prueba). De esta forma construye un modelo a través de los datos de entrenamiento y evalúa el rendimiento de este a través del conjunto de prueba.

El *gold standard* se refiere a pruebas de diagnóstico (en este caso, oraciones etiquetadas) que tienen la máxima fiabilidad a la hora de detectar el sentimiento de forma que se determina que tiene una fiabilidad máxima según unos criterios. Frecuentemente, este útil tiene como objetivo la correcta clasificación de la oración en su conjunto en las tres categorías (pos/neg/neu) de manera independiente.

Posteriormente, la evaluación de los resultados suele cuantificarse a través de las métricas de *precisión* y *exhaustividad* las cuales se emplean para cuantificar el rendimiento de los sistemas de búsqueda y recuperación de información (Figura 9). Se denomina *precisión* a la fracción de entidades detectadas que son relevantes para el estudio sobre el total, es decir la ratio de aciertos. Mientras *exhaustividad* es la fracción de entidades relevantes recuperadas sobre el total de entidades relevantes. Además del

análisis de sentimiento, se utiliza para evaluar tareas como el reconocimiento de imágenes o el rendimiento de motores de búsqueda, entre muchas. En los estudios de este tipo, se pretende obtener una precisión y exhaustividad alta denominada utilidad teórica (esto es, con valores próximos a 1). Con el objeto de ponderar y ver la distancia de ambas medidas a la utilidad teórica, suelen combinarse los valores de ambas métricas en una media armónica denominada *F-Score* o *valor-F*.

$$\text{Precisión} = \frac{VP}{VP + FP} \quad \text{Exhaustividad} = \frac{VP}{VP + FN} \quad \text{Valor F} = \frac{2VP}{2VP + FP + FN}$$

**Figura 9** Fórmulas para calcular la *precisión*, la *exhaustividad* y el *Valor-F*.

(VP= verdadero positivo, FP = Falso positivo, FN = falso negativo)

Una vez evaluado el modelo, se procederá a realizar predicciones con datos que se quieren clasificar. Los algoritmos distan de ser totalmente imparciales pues se basan suposiciones y generalizaciones subjetivas realizadas por humanos. Estas parten de preconcepciones sobre la forma en que deberían aprender los modelos y, obviamente, introducen un sesgo con una visión específica del mundo. Los expertos humanos actúan como “maestros” que enseñan a los modelos los resultados considerados deseables (p.e. a través del etiquetado de los datos de entrenamiento) que contienen los datos de entrada y, a partir de estos, el modelo aprende los patrones y los reproduce con los datos que se desean analizar.

#### 5.4.2 Enfoque basado en léxico

La segunda familia de sistemas de análisis de sentimiento son los sistemas *basados en léxico*. A diferencia de los sistemas principalmente estadísticos entrenados a partir de datos etiquetados manualmente, estos sistemas se basan en la disponibilidad de diccionarios de sentimiento. Estas recopilaciones, por lo general, toman como punto de partida elementos extraídos de corpus que posteriormente son codificados a un formato determinado, ampliados y anotados mediante una gran variedad de técnicas con el objetivo de aportar una base léxica sólida a los sistemas de análisis de sentimiento.

La fuente de conocimiento primordial de los lexicones para el análisis de sentimiento son palabras cuyo significado contenga una polaridad positiva o negativa. De esta manera, las palabras se convierten en la materia prima para los distintos sistemas de análisis de sentimiento. Un lexicón prototípico incluirá, al menos, una serie de entradas relacionadas con su polaridad (p. ej., “{bad: NEG}”). Además de utilizar fórmulas que

agreguen el sentimiento asignado a cada una de las entidades de un texto, se debe tener en cuenta la distancia entre las distintas palabras que incluyan en el texto y, en muchos sistemas, las unidades consideradas neutras. Además de las palabras, muchos lexicones incluirán también con frecuencia frases y expresiones idiomáticas que contienen expresiones de polaridad (Liu, 2012).

Para llevar a cabo la recopilación de la lista de léxico de opinión, existen tres enfoques principales. El *enfoque manual* consiste en asignar valencias a listados completos de léxico. Aunque esta técnica ofrece resultados con un alto nivel de refinamiento, consume mucho tiempo de elaboración y no se suele aplicar en exclusiva. Por ello, esta tarea se combina generalmente con los otros dos enfoques automatizados: recopilación mediante diccionarios o a través de corpus. Asimismo, los lexicones suelen requerir de trabajo cualitativo (encarecidamente, la revisión manual por parte de expertos) para realizar una verificación final con la intención de evitar errores típicos e inevitables en el procesamiento léxico automatizado. De esta forma, se aumentan los niveles de precisión y exhaustividad.

Un ejemplo prototípico sería el enfoque de Ding, Liu y Yu (2008) que propone un modelo generalista de análisis a partir de lexicones:

- Usa un lexicón de expresiones de sentimiento que incluye palabras, expresiones multipalabra, locuciones, y reglas de composición.
- Un conjunto de reglas para manejar los distintos constructos idiomáticos (*sentiment shifters*, y cláusulas adversativas).
- Una función de que agrega las puntuaciones de sentimiento a partir de las relaciones con sentimiento-objeto extraída del análisis sintáctico. Así, puede determinarse la orientación de sentimiento para cada uno de los objetos de opinión que haya en una oración determinada.

Según Liu (2015), los sistemas basados en léxico con clasificación *a nivel documento* obtienen las puntuaciones de orientación semántica mediante algoritmos que baremen estadísticamente todas las expresiones de sentimiento en un texto. La clasificación de sentimiento será positiva o negativa dependiendo del resultado final, pudiendo en muchos casos encontrarse con sistemas que asignen un sentimiento neutro a un texto si el sumatorio final de partes positivas o negativas es igual o similar a 0. Este tipo de enfoque tiene múltiples variantes que pueden clasificarse en las diferencias en el

valor que se le asigna a cada expresión de sentimiento, en la forma en la que se manejan las negaciones y en qué casos se asigna la información contextual.

En cuanto a los sistemas basados en léxico con clasificación a *nivel de aspecto*, se tiene en cuenta la distancia o la relación sintáctica entre las unidades léxicas de sentimiento. En este caso, el sistema debe hallar fácilmente el ámbito de aplicación de cada expresión de sentimiento y determinar si este debe cubrir el aspecto o entidad objetivo en la oración. Así, al ser el ámbito de estudio supraoracional, la explotación del léxico de sentimiento debe llevarse a cabo mediante el análisis de las relaciones sintácticas entre las expresiones de sentimiento y los objetos de opinión a través de distintas estrategias (Ding et al., 2008). Utilizando la clasificación de Liu (2015), aquí se agrupan las dependencias en tres categorías:

1. Dependencias sintácticas tipo sustantivo-adjetivo y verbo-adverbio. Por ejemplo, “This camera takes great pictures”<sup>27</sup>, puede usar la relación de dependencia sintáctica entre “great” y “picture” para identificar el objeto “picture” y la valencia asignada a “great”. En cuanto a la segunda relación, p. ej., “It's not just that many houses today don't convert easily to rental property.” relaciona al objeto “convert” con el objeto “easily”.
2. La palabra de sentimiento en muchos casos actúa como un indicador de aspecto. Por ejemplo, muchos adjetivos, p. ej., “ugly”, describen al aspecto *BELLEZA* en el dominio general y mientras que “easily” describiría al aspecto *DIFICULTAD*.
3. Las relaciones semánticas: esta es la tarea más compleja de esta lista, ya que a menudo es muy difícil reconocer el significado en patrones de uso de distintas unidades léxicas. Queda pendiente la implementación de modelos que lleven a cabo tareas tan complejas como la resolución de anáforas o determinar si es necesario llevar a cabo un modelo que procese de manera sencilla y eficiente no solo las estructuras sintácticas, sino también las estructuras semánticas. P.e. “Edgar admires Jacqueline”, en el que Edgar es el emisor de opinión y el objeto de la palabra de sentimiento positivo “admires” tiene claramente como objeto de sentimiento a Jacqueline. Si en cambio se utilizase la oración a “Jacqueline loathes Edgar”, el objeto de sentimiento negativo sería en este caso Edgar.

---

<sup>27</sup> Ejemplo (Liu, 2015, p. 95)

Es posible, entonces, ultimar que las distintas vertientes del análisis de sentimiento basado en léxico ofrecen una respuesta muy robusta y versátil al problema del análisis automatizado de la opinión. El rendimiento de los resultados, además, es mucho más transparente y fácil de mejorar a partir de técnicas de recopilación de información lexicográficas o terminográficas. Los distintos desafíos que ofrece su estudio (ironía, idiolectos, etc.) pueden afrontarse a partir de la profundización en los estudios del lenguaje evaluativo en uso; por ejemplo, desde la extracción de distintos patrones de la gramática local de un determinado contexto, o la adaptación del lexicón a una situación determinada. Como nota crítica, no obstante, fuera del estudio pormenorizado de los factores contextuales, pocos trabajos realizan estudios que incluyan niveles superiores de análisis lingüístico. Sin profundizar sobre ello, dada la cuestión de espacio, sería interesante abordar la complejidad que supone la manera que se tenga en cuenta la aplicación de reglas sintácticas o de detección de estructuras temáticas semánticas.

### 5.4.3 Enfoques híbridos

Las distintas metodologías descritas pueden combinarse entre sí y también hacer uso de técnicas de clasificación no supervisadas que se pueden combinar entre sí. Por ejemplo, Martín-Valdivia y Martínez-Cámara (2013) propusieron el uso de meta clasificadores para desarrollar un sistema de clasificación de polaridad. Trabajaron en un corpus español de reseñas de películas junto con un corpus paralelo equivalente traducido al inglés. En primer lugar, generaron dos modelos individuales tras utilizar estos dos corpus como fuente de datos y luego aplicaron algoritmos de aprendizaje automático (SVM, NB y otros). En segundo lugar, integraron la ontología de sentimiento *SentiWordNet* en el corpus en lengua inglesa de forma que se generase un nuevo modelo no supervisado que utilice un enfoque de orientación semántica. Finalmente, combinaron los tres sistemas utilizando un meta clasificador. Sus resultados superaron los resultados del uso de corpus individuales y demostraron que su enfoque aportaba una buena estrategia para la clasificación de polaridad cuando hay corpus paralelos disponibles (Medhat et al., 2014).

Qiu et al. (2010) proponen un enfoque integrador de métodos basados en léxico y autoaprendizaje (*self-learning*) que consiste en dos fases. En la primera, se hace uso de un método iterativo en el que algunas reseñas de productos son clasificadas a través de un lexicón de sentimiento. Posteriormente se clasifican otras reseñas mediante un proceso iterativo basado en un control de ratio positivo-negativo. En una segunda fase, un



clasificador supervisado entrenado toma algunas reseñas preclasificadas en la primera fase como modelo de entrenamiento y las aplica a otras reseñas sin clasificar. Este es, en esencia, un método sin supervisión que hace uso de una técnica supervisada que no requiere de reseñas clasificadas a mano en cada uno de los dominios específicos.

### 5.4.3 Ventajas, desventajas y desafíos de los distintos enfoques

La mayoría de los sistemas de análisis de sentimiento basados en léxico requieren una profunda comprensión del lenguaje evaluativo y trabajo lexicográfico a mayor o menor escala, de manera que se utilicen unidades léxicas y reglas contextuales para desempeñar los distintos métodos estadísticos o de aprendizaje automático. Mientras, los basados en *aprendizaje automático* extraen los términos positivos y negativos de forma automática a partir de los datos de entrenamiento. A pesar de la diversidad de enfoques de trabajo sobre la minería de opinión, no existe un tipo de sistema o técnica que sean preponderante a la hora de obtener mejores resultados pues, dependerá en gran medida de una resolución eminentemente práctica adaptada a las necesidades propias de cada investigación.

Los sistemas de *aprendizaje automático* son muy flexibles y se adaptan a todo tipo de dominios con facilidad. Sin embargo, estos sistemas precisan una gran cantidad de datos de entrenamiento etiquetados a mano y son muy dependientes del contexto, por lo que la portabilidad entre dominios es muy reducida. Para ilustrar esto, en algunas ocasiones se necesita etiquetar hasta 10.000 oraciones para comenzar a obtener buenos niveles de precisión y exhaustividad (Moreno-Ortiz et al., 2015).

Por otro lado, el enfoque basado en léxico es capaz de gestionar parcialmente los textos generalistas, ya que al tener una mayor independencia entre dominios no necesita que el usuario etiquete una gran cantidad de oraciones que incluyan contenido general para cada situación particular o dominio específico. En el caso de requerir una mayor precisión en la clasificación de un determinado contexto, este tipo de sistemas más flexible que el basado en aprendizaje automático. Los lexicones se pueden mejorar y ampliar de manera muy sencilla y adaptarlos a las necesidades particulares para afinar los resultados. En cuanto a desventajas, la construcción de un lexicon inicial con alto nivel de detalle necesita de una gran inversión en tiempo y conocimientos de gramática general y local para extraer el léxico general y/o específico, sintetizar el sentimiento y definir reglas contextuales (Moreno-Ortiz et al., 2015; Liu, 2015).



## 5.5 Modificadores contextuales

Además de las palabras y los sintagmas, una serie de construcciones gramaticales modulan el sentimiento cuando se ponen en relación con otros recursos léxicos. De esta manera, el cálculo de la polaridad de un texto u oración está dependiente no solo de las palabras con polaridad presentes, sino también de otros elementos como pueden ser partículas de negación o el uso de intensificadores y atenuadores. Con vistas a implementar un sistema que afronte este problema, Polanyi y Zaenen (2006) propusieron a nivel teórico lo que denominaron *sentiment shifters* o *contextual valence shifters*, una serie de reglas que modifican el valor de la expresión mediante inversión de la polaridad, su intensificación o su inversión. En el sistema de Hun y Liu (2004) y Kim y Hovy (2004), a cada una de las expresiones de sentimiento positiva se le asigna el valor de +1 y a cada expresión de sentimiento negativa se le asigna el valor -1. En este modelo, las palabras de negación como *not* y *never* invierten el valor de orientación semántica. Otros de los pioneros en su implementación técnica fueron Kennedy e Inkpen (2006) y en español, el trabajo pionero fue el sistema de Moreno-Ortiz et al., (2013), cuya descripción de *Context Rules* fue implementada primero en *Sentitext* (Moreno-Ortiz, 2010) y posteriormente en *Lingmotif* (Moreno-Ortiz, 2016) será la base de la descripción del concepto. Su modelo pone en marcha un clasificador de valores contextuales en el texto de tres tipos: negación, intensificadores y atenuadores.

Estas reglas provocarán variaciones en la puntuación de sentimiento inicial de manera que se hace posible el aumento, disminución, anulación total o inversión de la polaridad y su gradación. Un ejemplo de aumentadores de intensidad será la presencia del adverbio “highly” que aumentará el valor positivo de la palabra “profitable”. En el segundo ejemplo en cambio, disminuirá, al ser “vulnerable” una palabra con sentimiento negativo. Por ejemplo:

Governments have also nationalized HIGHLY PROFITABLE industries.

Vermont's maple sugaring is HIGHLY VULNERABLE to climate change.

El conjunto “not very” que se muestra en el siguiente ejemplo, disminuirá la intensidad de la carga semántica de *good*.

Investors are NOT VERY GOOD at market timing.

El modelo paradigmático en lengua inglesa es el *Sentiment Orientation Calculator* (SO-CAL) de Taboada et al. (2011) que extendió el modelo anterior y propusieron un

enfoque más refinado que merece ser observado con más detenimiento. En este sistema, la puntuación de sentimiento de cada elemento se mueve en un rango [-5, +5].

En este sistema, la intensificación o atenuación de las unidades con carga semántica funciona de la siguiente manera (a) cuando un adjetivo con una etiqueta de sentimiento positiva o negativa se coloca junto a un adverbio modificador (p. ej., “very good”) o (b) cuando un sustantivo con una etiqueta de sentimiento es modificado por un adjetivo anotado como intensificador o atenuador (p. ej., “total failure”).

El efecto de los intensificadores y atenuadores funciona a partir de un sistema lineal de adición o substracción. Esta serie de intensificadores o atenuadores están anotados con determinado peso porcentual. Por ejemplo, la partícula “slightly” tiene asignado un -50% con relación a la puntuación del sustantivo al que precede, “somewhat” un -30%, “pretty” un -10%, “really” suma un 15%, “very” +25% y “extremely” +90%, etc. En el caso de que estos elementos estén encadenados, estos se aplican de manera recursiva tomando como punto de partida la unidad léxica más próxima a la palabra de sentimiento (p. ej., “most excellent”). De esta manera, el sentimiento de “really very good” se modelaría de una forma más eficiente a partir de multiplicaciones, en lugar de a través de adiciones, de la siguiente forma  $[3 * [100\% + 25\%]]$  (Benamara et al., 2017).

La inversión de la orientación de sentimiento es una tarea más problemática. En ella existe una gran dificultad a la hora aplicar las reglas de negación a una palabra fuertemente positiva y en muchos casos, simplemente atenúa el sentimiento. La negación se lleva a cabo a partir de morfemas (“un-”, “a-”, “-less”) o a través de elementos de polaridad negativa (*Negative Polarity Items*). Estos son de distinta naturaleza (sustantivos, adverbios o verbos) y aparecen solo en las oraciones de carácter negativo, y no tiene correspondiente en las oraciones de carácter afirmativo (Israel, 2011). Existen varios enfoques para tratar la negatividad. Mientras que Taboada et al. (2011) consideran que debe asignar un sentimiento neutro (p. ej., “This is not good management”), otros autores como Benamara et al. (2012) observan que las acumulaciones de elementos de polaridad negativa refuerzan los sentimientos dentro de su ámbito (p. ej., “There is no lack of liquidity in the US economy” sería una expresión positiva). Algunos estudios consideran que la negación en la oración es mucho más marcada que la afirmación desde su propia esencia psicológica (Osgood y Richards, 1973), a pesar de que las expresiones negativas suelen presentarse acolchadas por recursos retóricos de carácter positivo, por lo que Taboada et al. (2011) propusieron que las expresiones de sentimiento negativas pesaran más a la hora de calcular el sentimiento de manera automática.

Una solución propuesta por Moreno-Ortiz et al. (2010) es el caso de la unidad “excellent” con una puntuación de +5: su negación “not excellent”, está muy lejos del sentimiento de una palabra en el otro lado de la clina de sentimiento como puede ser “atrocious” (-5), sino que además su tonalidad será mucha más positiva que la negación “not atrocious”. Por esto, se propone que las partículas negativas tengan un desplazamiento fijo en el valor de sentimiento (p. ej., 4), en lugar de invertir el signo como proponen algunos modelos (p. ej., Wilson et al., 2005). De esta manera, un adjetivo con valencia +2 restaría 4 puntos y el clúster “not atrocious” pasaría a tener una puntuación final de -2.

Otros factores contextuales que deben tenerse en cuenta son el modo y la modalidad, cuya relación con la expresión de la evaluación ya se revisó en el Capítulo 3. Estos aspectos extraproposicionales del significado aportan información dentro de la oración que trasciende el contenido proposicional de la misma. Por ejemplo, la modalidad, el modo irreal y la serie de marcadores gramaticales ofrecen información de carácter no factual y pueden ofrecer resultados poco confiables. Especialmente, hacemos referencia a sintagmas adverbiales que indican evidencialidad (“perhaps”, “maybe”), condicionales, modales (“could”, “may”), verbos intensionales (“think”, “consider”) y las expresiones multipalabra que actúan como modales (“a possibility is...”, “we consider that...”). En los casos *irrealis*, implementar una solución que determine cuándo estas oraciones tienen sentimiento se vuelve una tarea realmente compleja. Para ello, Brooke et al. (2009) proponen ignorar este tipo de proposiciones, estrategia a la que llamaron *irrealis blocking*. Benamara et al. (2017) consideran que existen algunos casos que deben ser estudiados con detenimiento, ya que es posible estructurar reglas que calculen el sentimiento de la acumulación de modalidad (*cumulative modality*) a partir de la gramática local. Sin embargo, no existen muchos estudios que describan la relación de estos patrones de modalidad, ni la modificación contextual de los valores de sentimiento. Para ilustrar esto proponen la siguiente oración, la cual es un ejemplo claro de acumulación de modalidad que implica claramente un sentimiento positivo:

(3) You definitely must see the movie.

Además de las citadas anteriormente, una implementación práctica se encuentra en las *reglas de opinión* de Liu (2010), llamadas posteriormente *reglas de composición de sentimiento* (2015). Estas son diez reglas que aquí se resumirán en seis. Se utilizan principalmente para modificar la puntuación de sentimiento en un sistema basado en

léxico, aunque también pueden utilizarse como características efectivas, muchas de ellas etiquetados manualmente, en el aprendizaje supervisado. Liu cita a Montague (1974), quien expuso que el significado de una expresión compuesta es una función del significado de sus componentes, y se estructura a partir de la sintaxis. Sin embargo, los esquemas de representación existentes distan de la representación de muchas de las reglas de sentimiento que se expondrán a continuación. El pseudolenguaje de anotación que se utiliza está basado en el *formato Backus-Naur* (BNF). A continuación, se presenta un compendio con las líneas principales de las reglas de composición:

- (a) Reglas de sentimiento general (*general sentiment rules*): son el nivel superior de la mayoría de las reglas generales y deciden la polaridad final de cada expresión. Esta puntuación final estará determinada a partir de la aplicación de las distintas reglas a partir de la jerarquía expuesta. Asigna reglas simples al léxico como PO o NE para valoraciones positivas o negativas, dobles negaciones (NEGATION NE), o expresiones adversativas (NE BUT). Por ejemplo, el algoritmo de Hatzivassiloglou y McKeown (1997) propone examinar oraciones como las siguientes. En el primer ejemplo, “expensive” y “complicated” deben tener la misma polaridad ya que se utilizan para expresar la misma opinión sobre el mismo tema. En el otro caso, “regrettable” debería tener la polaridad opuesta a “necessary”.

- a. It's expensive and complicated, especially for a nonprofit.
- b. These steps are regrettable but necessary to ensure a sustainable business.

- (b) Disminuir o incrementar la “cantidad” en un ítem de sentimiento: este set de reglas detecta elementos que disminuyen o incrementan la cantidad de un sintagma nominal y modifica la valencia de sentimiento. En la oración siguiente “fear” es palabra negativa de sentimiento y su reducción indica un efecto positivo o deseable.

The fear gradually subsided as CEBS lifted the veil on its methodology.

- (c) Disminuir o incrementar la cantidad de un ítem potencial positivo (*positive potential item*, PPI) / negativo (*negative potential item*, NPI): para algunos elementos, el incrementar o disminuir sus unidades o cantidades es positivo (o

negativo). Estas oraciones aportan datos y son generalmente oraciones factuales p. ej.:

Eurotunnel's first quarter revenue rose 9% this year.

House prices are rising

Aquí “revenue”, es un PPI porque su incremento es deseable. Sin embargo, para otros ítems, incrementar (o disminuir) su valor es negativo, por ejemplo, el caso de (b). Aquí “price” es un NPI porque su aumento no es deseable (para la perspectiva de la ciudadanía).

Los PPI y NPI no contienen sentimiento *per se*, y su descubrimiento y clasificación es esencial en muchos dominios específicos. Por ejemplo, en economía, “growth” y “budget” son PPI, pues su aumento es positivo. En el dominio de las hipotecas, “interest rate” and “down payment” son NPI desde el punto de vista de los receptores del préstamo (Wen y Wu, 2011).

- (d) Hechos deseables o indeseables: existe información deseable o indeseable tras hechos presentados como objetivos y/o factuales que implican sentimiento positivo o negativo. Estas deben ser descubiertas y etiquetadas a mano o de manera supervisada para cada entidad, por ejemplo:

After winding down, Adelphia will cease to do business.

- (e) Sentimiento parcial y total: las oraciones que cumplen con este conjunto de reglas son por lo general oraciones compuestas con diferentes sentimientos en distintas partes de la oración. La segunda parte frecuentemente sirve para resumir el sentimiento general sobre una entidad. Liu propone centrarse en el sentimiento general de la entidad y darle menos peso al sentimiento relacionado con el aspecto. Al ser estas oraciones generalmente complejas, la entidad suele encontrarse en la cláusula dominante:

Despite grim results on the high street, M & S still expects to hit its target of \$500m Internet sales in three years.

## 5.6 Compilación de lexicones de sentimiento

Un lexicón de sentimiento incluye un compendio de unidades léxicas a los que se les ha asignado sentimiento. Dado que la recopilación manual de unidades léxicas es un trabajo laborioso, que necesita de ingentes cantidades de tiempo, para la recopilación de

lexicones existen dos enfoques principales: el enfoque basado en diccionario y el enfoque basado en corpus (Deng, 2014; Liu, 2015).

### 5.6.1 Enfoque basado en diccionario

El uso de diccionarios a la hora de recopilar palabras de sentimiento es una aproximación acertada dado que muchas herramientas lexicográficas (p. ej., *Wordnet* o, incluso, el mismo corrector ortográfico de *LibreOffice*) contienen redes léxicas de sinónimos y antónimos que pueden ser aprovechadas.

En el enfoque basado en el diccionario, las nuevas palabras de sentimiento se identifican por su relación con un pequeño conjunto de palabras de opinión anotadas a mano, conocido como el *léxico semilla*, organizadas y relacionadas de forma que se recopilen en *lexicones semilla* y permitan tomarlos como base para extraer nuevas entradas aplicando técnicas de *bootstrapping* a partir de la estructura de sinónimos y antónimos de un diccionario. Deng, (2014) recoge un método simple de etiquetado de palabras que funciona de esta forma:

- (a) Se recopila un pequeño lexicón semilla a mano con orientaciones semánticas conocidas.
- (b) El algoritmo hace crecer este conjunto a través de la búsqueda de términos similares/opuestos en un diccionario de sinónimos y antónimos. En algunos proyectos (p. ej., Godbole, Srinivasaiah y Skiena, 2007; Kim y Hovy, 2004), se propone evaluar una a una las palabras de sentimiento iniciales mediante *WordNet* de manera que se generan listas de palabras negativas y positivas a partir de la expansión de listas de semillas hasta un determinado número de pasos. Las palabras que sobrepasen este umbral en la red de sinonimia serán consideradas neutras. En esta línea, Kamps et al. (2004) crearon una base de datos de caminos más cortos en redes de sinonimia.
- (c) Si es necesario, estas palabras se añaden a la lista de semillas y el algoritmo se vuelve a lanzar de manera recursiva.

Las relaciones entre las entradas suelen tener altos niveles de precisión (Deng, 2014). La efectividad de este método dependerá en gran medida de la cantidad y calidad de relaciones en las entradas de sinónimos y antónimos del diccionario utilizado. Otro desafío es la portabilidad entre dominios, de manera que se el análisis de sentimiento se

ajuste a la naturaleza del dominio de los textos. Sin embargo, aunque se han publicado muchos trabajos relacionados con este enfoque, no hay mucha literatura sobre evaluaciones independientes de estos métodos (Liu, 2015).

Aunque estos métodos suelen tener errores, es sencillo realizar una revisión manual para corregirlos, aunque esto suele requerir para un hablante nativo al menos un esfuerzo de varios días. Por otro lado, como podemos ver en la Figura 9, se presenta el problema de que la coherencia en la lista de sinónimos se debilita conforme se profundiza en la distancia de sinonimia en *WordNet*.

### 5.6.2 Enfoque basado en corpus

El segundo enfoque para la generación de lexicones de sentimiento es la serie de técnicas que componen el enfoque basado en corpus. En él, las palabras de opinión se reconocen en función de las relaciones semánticas que se han aplicado a dos escenarios principales.

El primero propone que a partir de una lista de palabras de sentimiento semilla (frecuentemente de propósito general), se descubran otras palabras de sentimiento relacionadas a través de la búsqueda de elementos con distinta polaridad en dominios específicos a través de un corpus especializado. Una forma de afrontar de manera semiautomatizada esta cuestión, sería a través de la conversión de un lexicón de propósito general en uno de dominio específico. Una estrategia para identificar palabras con sentimiento en un corpus propondría explotar reglas de conjunción para identificar simultáneamente algunas palabras con sentimiento y determinar su orientación semántica en un corpus determinado. Hatzivassiloglou y McKeown (1997) identificaron palabras con sentimiento utilizando patrones que hacían uso de conjunciones en un corpus de 21 millones de palabras del *Wall Street Journal* y propusieron generar un lexicón a partir de una lista de adjetivos semilla. Consideraron que dos adjetivos separados por una conjunción copulativa (ADJ + “and” + ADJ) deben tener la misma polaridad, por lo tanto, debe añadirse este adjetivo a la lista semilla. Inversamente, los adjetivos separados por una conjunción adversativa (ADJ + “but” + ADJ) deben ser asignados con el valor opuesto de sentimiento.

El segundo escenario propone utilizar palabras iniciales para identificar nuevas palabras de opinión. Turney (2002) combinó una serie de técnicas de recuperación de información y la información mutua puntual (PMI) para identificar la polaridad de frases. Aunque este método no está diseñado para desarrollar un léxico de sentimiento, permite



la identificación de la polaridad potencial de las palabras de sentimiento en los documentos que se clasifican. La PMI es una técnica estadística usada para medir la coincidencia entre dos palabras / frases. De esta manera, la polaridad de una palabra será la diferencia de su PMI con una palabra positiva y con una palabra negativa. Las dos palabras clave elegidas por el autor son “excellent” y “poor”, las cuales se consideran las más utilizadas en las reseñas en línea.

### 5.6.3 Lexicones de sentimiento en la literatura

A lo largo de estas dos décadas de estudio de análisis de sentimiento se han elaborado una gran cantidad de lexicones. A continuación, se destacan aquí los diez trabajos lexicográficos principales en lengua inglesa junto con sus principales características (Moreno-Ortiz y Pérez-Hernández, 2018):

- *The General Inquirer*<sup>28</sup> (Stone y Hunt, 1963). Este es uno de los lexicones de sentimiento más veteranos y populares que parte de un estudio de psicología cognitiva y análisis de contenido. Al estar fácilmente accesible en la red se utiliza en una gran cantidad de trabajos. Ofrece información sintáctica, semántica y pragmática además de marcadores de discurso en una clasificación de polaridad. La clasificación de la tonalidad está basada en 182 categorías, principalmente: positivo, negativo, fuerte, débil, activo, placer o dolor. Incluye 1915 unidades positivas y 2291 negativas.
- *Opinion Lexicon*<sup>29</sup> (Hu y Liu, 2004). Recopila aproximadamente 6.800 palabras extraídas de reseñas de productos, etiquetados originalmente mediante *bootstrapping*. Parte de los conjuntos adjetivales de sinónimos y antónimos de WordNet. Contiene 2.006 unidades positivas y 4.783 negativas.
- *Multiperspective Question Answering Subjectivity Lexicon*<sup>30</sup> (MPQA) (Wilson *et al.*, 2005). Este lexicón incluye contiene 2.718 unidades léxicas positivas y 4.912 negativas. Fueron extraídas de una serie de fuentes combinadas combinación, incluyendo el léxico de *General Inquirer*, el sistema de Hatzivassiloglou y

<sup>28</sup> Disponible en <http://www.wjh.harvard.edu/inquirer/>.

<sup>29</sup> Disponible en <http://www.cs.uic.edu/liub/FBS/opinionlexicon-English.rar>

<sup>30</sup> Disponible en [http://mpqa.cs.pitt.edu/lexicons/subj\\_lexicon/](http://mpqa.cs.pitt.edu/lexicons/subj_lexicon/).



McKeown (1997) y la lista de pistas subjetivas obtenida aplicando técnicas de *bootstrapping* (Riloff y Wiebe, 2003). Este lexicón fue etiquetado a mano e incluye cuatro categorías de polaridad (positivo, negativo, neutro y ambos) y una etiqueta de nivel de subjetividad. La mayoría de las palabras se marcan como positivas (33,1%) o negativas (59,7%), un pequeño número de palabras (0,3%) están anotadas con ambas polaridades y un 6,9% del léxico está etiquetado como neutro.

- *SentiWordNet 3.0*<sup>31</sup> (Baccianella, Esuli y Sebastiani, 2010). No es mero lexicón, sino una estructura ontológica de sentimiento. Está diseñada a partir de la anotación estadística de la estructura jerárquica de *WordNet 3.0* a cuyas entradas se le asignaron valores positivos, negativos o neutros. En su versión actual, el algoritmo de anotación ha refinado sus puntajes aplicando estrategias estadísticas masivas *random walk* junto al aprendizaje semisupervisado.
- *Macquarie Semantic Orientation Lexicon* (MSOL) (Mohammad, Dunne y Dorr, 2009). Este lexicón incluye unos 76.400 elementos léxicos (51.208 palabras y 25.192 expresiones multipalabra) etiquetados como positivos o negativos de manera semiautomática. Está recopilado a partir de un lexicón semilla y un algoritmo que etiquetó las entradas de un tesoro de tipo Roget. Incluye 30.458 unidades positivas y 45.942 negativas.
- *NRC Emotion lexicon/EmoLex* (Mohammad y Turney, 2010). Este es un lexicón que además del sentimiento (positivo y negativo), etiqueta las palabras según ocho emociones básicas (ira, miedo, anticipación, confianza, sorpresa, tristeza, gozo, y asco). Sus 14.182 entradas fueron etiquetadas manualmente a partir de *crowdsourcing* a través de la plataforma *Amazon Mechanical Turk*.
- *Sentiment140 Lexicon*<sup>32</sup> (Mohammad, Kiritchenko y Zhu, 2013). Este lexicón fue creado a partir de una colección de 1,6 millones de tweets. Incluye 62.468 unigramas, 677.698 bigramas y 480,010 pares etiquetados como positivo o negativo.

---

<sup>31</sup> Disponible en <http://sentiwordnet.isti.cnr.it/>.

<sup>32</sup> Disponible en <http://saifmohammad.com/WebPages/lexicons.html>

- *LIWC2015*<sup>33</sup> (Pennebaker et al., 2015). Este lexicón se incluye además en una *suite* de análisis de sentimiento. Incluye 6.400 entradas (palabras, raíces morfémicas, etc) y, además, clasifica según las distintas características sintácticas y psicológicas.
- *Valence Aware Dictionary and sEntiment Reasoner* (VADER) (Hutto y Gilbert, 2014). Se trata de un lexicón que contiene 7.500 entradas (léxico, emoticones y argot de Internet) y está enfocado al estudio del sentimiento en las redes sociales a partir de una revisión manual realizada a través de *Amazon Mechanical Turk*.
- *Lingmotif-Lex* (Moreno-Ortiz y Pérez-Hernández, 2018). Es la base de *Lingmotif* y tiene varios componentes principales: el lexicón que incluye unidades léxicas simples, expresiones multipalabra y *emojis*, y un juego de reglas contextuales. Está etiquetado según sentimiento (pos, neg, neu) e intensidad (1-3).

<i>Nombre</i>	<i>Autores</i>	<i>Descripción</i>
<i>General Inquirer</i>	Stone y Hunt, (1963)	1.915 positivas y 2.291 negativas
<i>Opinion Lexicon</i>	Hu y Liu, (2004)	2.006 positivas y 4.783 negativas
<i>MPQA</i>	Wiebe, Wilson y Cardie, (2005)	4.912 negativas y 2718 positivas
<i>Sentiwordnet</i>	Baccianella et al., (2010)	11.7000 <i>synsets</i>
<i>MSOL</i>	Mohammad, Dunne y Dorr, (2009)	30.458 positivas y 45.942 negativas
<i>NRC Emotion Lexicon</i>	Muhammad y Turney, (2010)	14.182
<i>Liwc</i>	Pennebaker, et al., (2015)	6.400
<i>Vader</i>	Hutto y Gilbert, (2014)	7.500
<i>Lingmotif-Lex</i>	Moreno-Ortiz y Pérez-Hernández (2018)	67.400

**Tabla 7** Resumen de los principales lexicones de sentimiento

## 5.7 Clasificación de sentimiento en dominios específicos

En una discusión temprana sobre búsqueda semántica, Hampp y Lang (2005) debatieron sobre la factibilidad de adaptar el léxico a industrias específicas, como manufactura, servicios o gobierno, con el fin de aprovechar al máximo la capacidad de detectar particularidades y entender el sentimiento asociado con esas palabras. Sugieren que, de esta manera, el conocimiento específico del dominio puede aprovecharse para facilitar la desambiguación de los términos y comprender mejor el contexto. Como se reitera a lo largo de este trabajo, la clasificación de sentimiento es muy sensible al dominio específico

<sup>33</sup>Web comercial: [www.liwc.net/](http://www.liwc.net/)

(como la economía o la medicina), al que pertenece el lexicón o de los datos de entrenamiento de un sistema.

A la hora de estudiar el sentimiento de los lenguajes especializados en el análisis de sentimiento, cabe hacerse dos preguntas: ¿Qué factores tienen mayor peso a la hora de determinar el tono de un texto? ¿Hay palabras positivas o negativas relativas a un sector determinado que no lo son en el dominio general o en otros dominios? Los clasificadores entrenados con datasets específicos de dominio suelen obtener una precisión mayor porque identifican las palabras de sentimiento automáticamente. Además, algunas unidades léxicas pueden ser positivas en un dominio y negativas en otro. Es por este motivo que en los sistemas basados en léxico una adaptación de dominio se hace necesaria. De hecho, los sistemas con gran índice de precisión en el dominio general o en un dominio específico determinado, no suelen tener altos índices de precisión en un contexto distinto. Por ello, es necesario hacer un estudio en profundidad del dominio para establecer estrategias pertinentes a la hora de estudiar de manera automatizada el sentimiento (Malo et al., 2014; Moreno-Ortiz y Fernández-Cruz, 2015, 2018; Liu, 2015).

Uno de los primeros estudios en afrontar este problema fue el realizado por Aue y Gamon (2005) quienes propusieron una transferencia de dominios debido a que no existían de datos etiquetados en un nuevo dominio. De esta manera, experimentaron con cuatro estrategias:

1. Entrenar una mezcla de reseñas etiquetadas para otros dominios y probarlas en el dominio objetivo.
2. Entrenar un clasificador como en la primera estrategia, pero limitándose a las características observadas en el dominio objetivo.
3. Usar conjuntos de clasificadores de dominios con datos que hayan sido previamente etiquetados y probarlos en el dominio objetivo.
4. Combinar pequeñas cantidades de datos etiquetados con grandes cantidades de datos sin etiquetar del dominio objetivo (este es el tradicional ajuste de aprendizaje semisupervisado).

Entre los problemas más frecuentes es destacable que incluso dentro de un mismo dominio muchas palabras pueden expresar diferentes orientaciones semánticas dependiendo de la entidad analizada. Según Liu (2012), estas situaciones suelen ocurrir con adjetivos de cantidad (p.e “The battery life is long” vs “It takes a long time to

focus”<sup>34</sup>). Es por ello por lo que encontrar palabras de sentimiento dependientes del dominio y sus orientaciones semánticas es, en limitadas ocasiones, insuficiente. Algunas aplicaciones que trabajan a nivel entidad han utilizado relaciones entre aspecto y sentimiento de las palabras y proponen el uso de duplas p. ej., (“content\_sentiment\_word”, “aspect”) de manera que se representen los contextos de opinión, p. ej., (“long”, “battery life”) (Deng et al., 2008). Algoritmos como el de Popescu y Etzioni (2007) proponen partir de una etiqueta global que vaya refinando el sentimiento de los elementos del texto a través de barridos graduales: primero a partir de un corpus específico y, finalmente, a nivel entidad.

Una opción destacada es la aplicación de sistemas basados en léxico y, por tanto, recopilar un lexicón de palabras de sentimiento para el dominio específico de manera que se puedan analizar textos especializados. Existen varios problemas fundamentales:

- Cómo identificar y gestionar las palabras y sintagmas que dependen del dominio específico o del contexto. Con frecuencia, las palabras de sentimiento cambian su orientación semántica en un dominio específico distinto. Por lo tanto, si las palabras de sentimiento dependientes del dominio no son identificadas adecuadamente, pueden afectar significativamente la precisión del clasificador de sentimiento. Por ejemplo, las entradas en un diccionario de sentimientos de dominio general como el *General Inquirer* tienen una polaridad diferente en el contexto de la economía o la medicina. Por ejemplo, “growth” en inglés hace referencia a crecimiento económico en el este dominio (positivo), mientras que hace referencia al crecimiento en altura (neutro) de p. ej., un niño o a tumores (negativo) en el caso de la oncología. Es por esto por lo que un sistema basado en léxico requiere de la información sobre la orientación semántica que es específica para el dominio objetivo.
- Cómo detectar, y en muchas ocasiones descartar, las palabras y sintagmas factuales que implican sentimiento en un dominio determinado. El factor del experto en el manejo de terminología, modismos y jerga es esencial, al existir una gran cantidad de técnicas de extracción de información (Liu, 2015).

---

<sup>34</sup> Ejemplos de Liu (2012 p. 84)

*Lingmotif* (Moreno-Ortiz et al., 2015) propone un método simple y novedoso para extraer elementos léxicos de dominio específico (palabras y expresiones multipalabra) a partir de textos especializados que divergían en la orientación semántica de elementos léxicos de dominio general. En lugar de utilizar solo un léxico específico, combina el uso de dos lexicones: el general y uno específico al dominio de análisis. Como se verá en el Capítulo 6 de manera más detallada, una de sus principales ventajas es un mecanismo que permite importar lexicones proporcionados por el usuario mediante el uso de archivos CSV. Estos archivos de texto contienen una lista de palabras de opinión dependientes del dominio que anularán solamente los valores de opinión en su lexicón central. De esta manera, se aprovechan las fuentes léxicas preexistentes y solo se modifican los elementos propios del dominio que en su uso especializado tengan un sentimiento diferente.

### 5.7.1 Lexicones de dominio económico-financiero

El análisis de sentimiento en el dominio de la economía y las finanzas es considerado un problema destacado en la literatura. El desarrollo del análisis de sentimiento como disciplina se ha desarrollado durante una época de incertidumbre financiera, por lo que no es casualidad que muchos estudios en esta disciplina estén destinados a analizar los sucesos económicos para establecer formas de medir diversos parámetros (informes financieros, análisis de discursos de las figuras relevantes, análisis de la prensa económica, etc.) de manera fiable y precisa. Los resultados de estos estudios suponen una ventaja ante los métodos tradicionalmente (análisis manual de informes o índices oficiales como el *Indicador de Sentimiento Económico de la Zona Euro*), de tal forma que es posible tomarles el pulso a unos indicadores a través del análisis automatizado de opiniones relevantes para determinado índice. La economía es un lenguaje especializado, y la noción de dependencia del dominio es particularmente relevante en este contexto, pues requiere del juicio de expertos para definir qué tipo de polaridad tiene un texto determinado (Krishnamoorthy, 2017; Loughran y McDonald, 2015; 2016; Malo, et al., 2014).

La idea de aplicar el análisis textual a los mercados financieros no es del todo novedosa y la idea del impacto del sentimiento en los mercados está bien establecida (cf. Klein y Prestbo, 1974; Ederington y Lee, 1993). Uno de los pioneros fue Niederhoffer (1971), veterano académico y gestor de *hedge funds*, que analizó manualmente dos décadas de titulares económicos del *New York Times*. Estos fueron asignados a

diecinueve categorías semánticas en una escala *bueno-malo* determinando que los mercados tienden a exagerar la reacción ante noticias negativas. De una manera profética propuso el desarrollo de técnicas de análisis computacional de los textos para obtener mayor objetividad en los resultados (Devitt y Ahmad, 2007ab). Los estudios de análisis de sentimiento en el dominio económico y financiero surgieron a mediados de la década de los 2000 y estudiaron su relación, principalmente, con las predicciones bursátiles (p. ej., Melvin y Yin, 2000), el impacto de informaciones exógenas, especialmente, los titulares de prensa en los resultados financieros (Chan, 2001; Poon y Granger, 2005), o el sentimiento de expertos hacia distintas firmas y sus acciones (O'Hare et al., 2009), etc. El trabajo de Tetlock (2007) es un estudio pionero que analiza una estrategia de inversión abiertamente basada en análisis de sentimiento y discute los inconvenientes de usar métodos probabilísticos basados en ensayos difíciles de replicar. Aunque los estudios expresados en el apartado anterior han propuesto varios métodos de generación de léxico, la mayoría de ellos no tuvieron en cuenta la necesidad de aportar información específica de dominio en los lexicones. Loughran y McDonald (2011) desarrollaron el primer lexicon de sentimiento extensivo relativo al dominio de las finanzas extraído a partir de informes corporativos.

Hace medio siglo, la escuela denominada *finanzas modernas* investigó el impacto de las noticias en los mercados financieros. Los expertos del campo financiero tienen distintas visiones sobre la forma en la que los mercados responden ante tal información y cómo el estudio de los artículos de prensa puede incorporarse a las estrategias financieras. Entre las más destacables está la *hipótesis del mercado eficiente* (*Efficient Market Hypothesis*) que considera que la nueva información se difunde con rapidez y esta influye de manera racional en el flujo de inversiones y en los precios de manera inmediata. Así, los cambios de valores futuros reflejarán noticias futuras (Chan y Chong, 2017). Este enfoque empezó a ser desafiado a partir de la década de 1980 por las *finanzas conductuales*. En la actualidad, la mayoría de los enfoques consideran que el proceso económico capitalista asume que las emociones rigen, como un factor más, el comportamiento de los inversores.

Chan y Chong (2017) hacen una revisión del historial de los antecedentes principales del análisis textual a los mercados financieros. Una encuesta de Klein y Prestbo (1974) muestra cómo un informe de noticias financieras pesimista puede afectar a los mercados y apoya firmemente la sugerencia de que las noticias y los mercados se influyen mutuamente. Ederington y Lee (1993) concluyen que los textos financieros,

especialmente los comunicados de prensa, pueden arrojar luz sobre la volatilidad dentro del mercado. Engle y Ng (1993) propusieron el estudio de la curva de impacto de las noticias, que proporciona un mecanismo para explicar los rendimientos del mercado utilizando noticias. Blood y Philips (1995) analizaron la influencia causal de los titulares periodísticos con el indicador de sentimiento de los consumidores, la popularidad presidencial y varios indicadores macroeconómicos a través de una línea temporal entre 1989 y 1993. Cuando hacemos referencia a “indicador de sentimiento”, es necesario no confundir el término *sentimiento* como lo hemos venido usando con el indicador homónimo, que define el *sentimiento financiero* como “the expectations of market participants relative to a norm: a bullish (bearish) investor expects to be above (below) average, whatever “average” may be” (Fama, 1970; Brown y Cliff, 2004; citados en Van de Kauter et al., 2015; pp. 4999-5000).

Shiller (2000) discute el papel clave que juegan los medios en el escenario económico para los movimientos bursátiles que, en gran manera, son provocados por estos en una suerte de profecía autocumplida. Esto se debe a que los inversores siguen determinadas publicaciones consideradas influyentes y sugiere que la opinión del mercado está, en cierto modo, impulsada por el contenido de las noticias. En la línea de este trabajo, Tetlock (2007) muestra que el número de palabras negativas en la columna *Abreast of the market* del *Wall Street Journal* predice retornos de acciones en la frecuencia diaria de 1984 a 1999. A partir de estas investigaciones, García (2013) construye una medida longitudinal de sentimiento basada en noticias financieras del New York Times de 1905 a 2005, y estudia su relación con el desempeño financiero. La larga serie temporal muestra que la polaridad de noticias ayuda a predecir el rendimiento de las acciones en una frecuencia diaria, particularmente durante las recesiones.

Generalmente, los estudios que se han realizado sobre el análisis de sentimiento no se han dirigido específicamente a estudiar en profundidad los textos financieros, y se han centrado en múltiples estrategias de tipo estadístico basado en previsiones financieras. Es muy pertinente, entonces, profundizar en la noción de dependencia de dominio, hecho altamente relevante no solo en el terreno lingüístico, sino en el periodístico y, obviamente, en el contexto económico financiero. Como cualquier estudio lingüístico dedicado a las lenguas altamente especializadas, su estudio requiere de un alto nivel de experiencia para definir qué tipo de polaridad tienen las unidades léxicas de un texto, por lo que generalmente es necesario contar con un equipo de expertos para analizar en profundidad la naturaleza de este. Por esto, consideremos que para realizar un análisis



de sentimiento preciso de textos del dominio económico-financiero es esencial desarrollar un lexicón especializado con la amplitud suficiente como para obtener buenos niveles de exhaustividad y precisión.

### **5.7.2 Características del dominio económico-financiero en relación con el análisis de sentimiento**

Los clasificadores de sentimiento basados en léxico dependen de amplias fuentes léxicas, por lo que la construcción de un lexicón específico detallado es un prerequisite esencial para construir un analizador de sentimiento económico eficiente que tome en cuenta las características y particularidades del dominio específico. Sin embargo, Krishnamoorthy (2017) considera inadecuada la predicción de polaridad de textos financieros exclusivamente a través de lexicones de dominio específico. Al contrario que la mayoría de los enfoques de análisis de sentimiento, en este dominio los textos de noticias no siempre incluyen el sentimiento explícito del autor, sino que generalmente aportan opinión factual con carga semántica que refleja hechos deseables o indeseables (Van de Kauter et al., 2015) en un determinado contexto.

En economía, la variación es más importante que los valores absolutos. Por lo tanto, para evaluar la orientación de opinión dentro de un texto económico, se busca el estado futuro sugerido por los indicadores encontrados. Los términos que indican el crecimiento o disminución del indicador al que se refieren generalmente se denominan *modificadores* (p. ej., “rise” o “fall”). Estos también se dividen en positivos y negativos, con los modificadores positivos señalando un aumento cuantitativo y los negativos una disminución de su indicador asignado. Una gran cantidad de terminología económica es neutra por naturaleza (p. ej., “sales”), no obstante, incluye una fuerte orientación semántica cuando se combina con verbos y otros elementos léxicos direccionales. Los sistemas que hacen uso solo de términos económicos/financieros probablemente clasifiquen las oraciones como neutras, por lo que es necesario complementar el lexicón con una serie de reglas de direccionalidad que se verán a continuación (Musat y Trausan-Matu, 2010).

Esta dificultad añadida se debe a la existencia de palabras que son consideradas del dominio general (“rise”, “escalate”, “decrease”), pero que sin embargo tienen algún tipo de carga semántica cuando son relacionadas con determinado léxico del dominio. Mientras que en otros dominios (p. ej., las reseñas cinematográficas) los sentimientos se



expresan frecuentemente en combinaciones de adjetivos y adverbios, en el dominio económico se necesita además determinar si la dirección de los eventos es favorable (o no) desde un punto de vista económico (p. ej., “sales are rising”) a través de combinaciones de verbos y sustantivos (Malo, et al., 2014). Algunos ejemplos propuestos por Krishnamoorthy (2017) son los siguientes:

- a. *Turnover* ROSE to EUR 21mn from EUR 17mn.
- b. The *EPS* outlook was INCREASED by 5.6
- c. Unit *costs* for flight operations FELL by 6.4%
- d. The company intends to RAISE *production capacity* in 2006.
- e. *Demand* for fireplace products was LOWER than expected, especially in Germany.

Los autores de textos relacionados con el lenguaje económico son frecuentemente prudentes a la hora de expresar un sentimiento negativo. Los enfoques de análisis de sentimiento de tipo *bag-of-words* no suelen ser los más efectivos a la hora de analizar el lenguaje financiero, de manera que el clasificador contraste de manera exclusiva el recuento de palabras positivas y el número de palabras negativas. Por otro lado, es frecuente encontrar oraciones complejas como estrategia psicológica para expresar sentimiento negativo, de manera que este se diluya en la prosodia del texto (Chang y Chong, 2017). Al respecto, Loughran y McDonald (2011) determinaron que un 73,8% de las palabras negativas del *General Inquirer*, como “tax”, “cost”, “board”, o “vice” no lo son en el contexto económico-financiero.

Varios autores (Krishnamurthy, 2017; Malo *et. al*, 2014; Musat y Trausan-Matu, 2010) examinan los factores objetivos en el texto, tanto léxicos como estructurales, destacando tres categorías de palabras:

1. *Indicadores de rendimiento (performance indicators)*: términos indicadores de seguimiento o liderazgo (*lagging/leading indicators*) en el mercado. Los indicadores de seguimiento reflejan los resultados de la actividad de una firma o institución: “market share”, “operating profit”, “operating cost”, etc. En cambio, los indicadores de liderazgo hablan de posibles rendimientos futuros de la firma: “new stores”, “new customers”, “increase in production capacity”, etc. Estos términos no tienen sentimiento *per se* y son potencialmente positivos o potencialmente negativos, ya que su aumento o disminución hará que el resultado final sea positivo o negativo.

2. *Direccionalidad de los indicadores de seguimiento y liderazgo*, es la palabra o expresión multpalabra que describe la dirección de los eventos. Por ejemplo, “increase”, “rise”, “skyrocket”, “decrease”, “plummet”, “collapse” marcan tanto el sentimiento como la intensidad de los indicadores de rendimiento. Los indicadores económicos se dividen en dos subconjuntos: uno que contiene índices que aumentan cuando la economía está creciendo (p. ej., “revenue”, “exports”) y uno con índices que tienen una evolución positiva durante un período de declive económico (p. ej., “unemployment”, “risk”, “yield”). Al primer conjunto de indicadores se le denominará “positivo” y a sus opuestos, “negativo”.
3. *Palabras o expresiones multpalabra propias del dominio específico*, incluye terminología, jerga y términos utilizados más frecuentemente con una polaridad distinta que en el dominio general.

### 5.7.3 Sistemas de análisis de sentimiento económico-financiero

Los principales estudios llevados a cabo para determinar la polaridad en los textos sobre la economía o las finanzas han sido de una naturaleza muy variada y con objetivos eminentemente prácticos. Principalmente, se han llevado a cabo para predecir movimientos bursátiles a partir de noticias en la prensa, redes sociales o informes corporativos de las empresas. No existen, bajo nuestro conocimiento, aplicaciones interesadas en analizar otros fenómenos de carácter discursivo o sociológico relacionado con el impacto de la economía en los distintos ámbitos de la vida humana. Dada la naturaleza del estudio, se resumen aquí las principales aproximaciones del estudio del análisis de sentimiento basado en léxico:

- (1) *Lexicones de dominio general*. Al ser el dominio de la economía un lenguaje altamente especializado, la aplicación de lexicones generalistas de dominio suele clasificar de forma equívoca los textos especializados, lo cual impacta en el rendimiento de los clasificadores. También en este dominio, el *General Inquirer* ha sido uno de los lexicones más utilizados. Se destacan aquí los sistemas, descritos más arriba, de Tetlock, Saar-Tsechansky y MacSkassy, (2008) y Tetlock (2016). Engelberg (2008) discute que la lista de palabras positivas del GI falla en la clasificación para el dominio financiero. De hecho, el estudio del GI realizado

por Loughran y McDonald (2011) indica que el 75% de las palabras negativas no lo son en el dominio financiero.

- (2) *Lexicones de dominio específico*. Estos lexicones suelen ofrecer buenos resultados de exhaustividad, pero dependen de una ardua labor lexicográfica. Estos son creados *ex nihilo* (Loughran y McDonald, 2011), utilizados en combinación con un lexicón del dominio general (como el propuesto aquí) o adaptados desde un lexicón general para el dominio económico (Van de Kauter, et al., 2015). Al ser enfoque principal de este trabajo, se detalla una lista concisa de lexicones financieros.
- (3) *Métodos supervisados*. Solo existe una pequeña serie de técnicas de aprendizaje automático (NB, SVM, BOW, etc.) que ha sido aplicada en múltiples ocasiones con distintos objetivos. En O'Hare et al., (2009) se hace uso de *Naïve Bayes* y SVM para clasificar blogs. Li (2010) hace uso de un clasificador bayesiano para realizar una clasificación *bag-of-words* sin el uso de un lexicón. Algunos modelos incluidos en esta clasificación son los siguientes: Antweiler y Frank, (2004); Das y Chen, (2007); Huang, Zhang y Zheng, (2014); Malo et al., (2014); Van De Kauter, et al., (2015).

En esta década, se han propuesto métodos de generación de léxico específicos del dominio que ofrecen una mejora significativa en el rendimiento de los clasificadores. En esta última parte del apartado, se ofrece a continuación una revisión de la literatura relacionada con el desarrollo de lexicones de sentimiento. Entre los lexicones específicos más destacados están *LM* (Loughran *et al.*, 2016) y *Financial Polarity Lexicon* (FPL) (Malo et al., 2014). Estudios surgidos a partir de estos han obtenido mejoras considerables en la calidad de la detección del sentimiento (Li, Wang, et al., 2014; Li, Xie et al., 2014; o Schumaker y Chen, 2009; citados por Krishnamurthy, 2017).

Loughran y McDonald (2015) crearon manualmente un lexicón financiero con seis listas de palabras: positivas, negativas, litigiosas, con incertidumbre, de modal fuerte y modal débil. Estos términos estaban presentes en al menos el 5% de un corpus de 10.000 documentos económicos. El trabajo de Malo et al. (2014) es uno de los más destacados por el diseño de un banco de frases que incluye conceptos de dominio específico, una amplia lista de verbos y expresiones que ayudan a detectar la dirección del evento y una clasificación de la potencialidad de algunos términos dependientes de la direccionalidad. El sistema resultante es sensible a la direccionalidad de las entidades a través del *parsing*

de las oraciones. Por otro lado, Van de Kauter, et al. (2015) diseñaron un lexicón en neerlandés a partir de la modificación del lexicón general *Pattern y Duoman* (53.209 ítems) con el etiquetado manual de varios lexicones que incluyen 3.187 ítems, e indican que está diseñado para incluir tanto lenguaje explícito como el lenguaje implícito o factual. El trabajo de Oliveira, Cortez y Areal (2016) desarrolló un método mixto en el que se diseña un lexicón de sentimiento a partir de un corpus de mensajes de *Stocktwits* etiquetado previamente. Posteriormente, se identificaron las palabras de opinión en función de la fuerza de asociación con las clases de sentimiento (alcista y bajista). A pesar de la utilidad de este método con datos anotados, no se puede aplicar fácilmente en datos no etiquetados.

Los enfoques estadísticos o basados en aprendizaje automatizado utilizan *bag-of-words* o *n-gramas* como características de manera que aplican métodos generativos o discriminatorios para predecir los sentimientos. Koppel y Shtrimberg (2006) diseñaron un método que relacionaba los textos económicos con anotaciones del rendimiento de los valores del índice de Standard & Poor's. De esta manera, se extrajeron los términos más frecuentes y se calculó su sentimiento mediante los resultados obtenidos. Al contrario que en las malas noticias, se destaca la ausencia de palabras positivas cuando los documentos relataban valores bursátiles altos. En términos similares, el modelo diseñado por Zhang et al. (2014) utiliza un algoritmo que acumula las puntuaciones oracionales para ofrecer un sentimiento general a nivel documento. Asimismo, otros sistemas predijeron movimientos mediante, entre otros, el análisis de la gramática local en los textos financieros en varios idiomas (Ahmad, Cheng y Almas, 2006; Zhao y Ahmad, 2015).

Finalmente, esta revisión bibliográfica se resume mediante la Tabla 8 adaptada a partir de la recopilada por Kirshnamoorthy (2017). En ella se recogen los estudios principales de este dominio, el lexicón utilizado y el objetivo de estos.

<i>Estudio</i>	<i>Tipo de texto</i>	<i>Lexicón</i>	<i>Características</i>	<i>Método</i>	<i>Objetivo</i>
Tetlock	Noticias	GI	Léxico	Regresión	Rendimiento de las acciones
	Noticias	GI	Léxico	Regresión	Ganancias, rendimiento
Loughran	10K	LM	Léxico	Regresión	Rendimiento
Antweiler	Mensajes de Internet	-	Bag of Words	Regresión, NB	Volatilidad, rendimiento
Huang	Informes de analistas	GI, LIWC, LM	Bag of Words	Regresión, NB	Polaridad, ganancias
Li	10K	-	Bag of Words	NB	Polaridad, contenido de las informaciones
	Noticias, redes sociales	-	Sustantivos, palabras con carga semántica	Regresión	Tendencias en los mercados
	Noticias	GI, LM	Palabras de diccionario	SVM	Tendencias en los mercados
	Noticias, mensajes de Internet	GI, LM	Entidades, palabras con carga semántica	SVR	Tendencias en los mercados
O'Hare	Blogs	-	Términos y colocaciones	NB, SVM	Predicción de la polaridad
Mo	Noticias	SWN	Palabras de SWN	Regresión	Rendimiento en las acciones
RPS	Noticias	-	Sustantivos	SVR	Tendencias en los mercados
Malo et al.	Noticias	MPQA, GI, LM, DI	Entidades, estructura sintagmática	SVM	Predicción de la polaridad
Krishnamurthy	Noticias	PI LM DI	PI tags	ARM	Predicción de la polaridad

**Tabla 8** Resumen de los principales estudios de análisis de sentimiento relacionado con el dominio económico-financiero (Kirshnamoorthy, 2017)

## 5.8 Cambios en la orientación semántica y análisis de sentimiento

El *cambio semántico* (cf. Bybee, 2015; Bréal, 1897; Traugott y Dasher, 2002; Stern, 1931; Ullmann, 1957; 1962) es un fenómeno de cambio en el lenguaje con respecto a la evolución del uso de la palabra en distintos niveles, en muchas ocasiones hasta el punto de que el significado moderno es radicalmente diferente del uso original. Cada palabra tiene una variedad de sentidos y connotaciones, que pueden agregarse, eliminarse o variar con el tiempo. Stern (1931, p. 163) define el concepto de la siguiente forma:

the habitual modification, among a comparatively large number of speakers, of the traditional semantic range of the word, which results from the use of the word (a) to denote one or more referents which it has not previously denoted, or, (b) to express a novel manner of apprehending one or more of its referents.

El estudio del cambio de significado a través de sus relaciones no es un fenómeno exclusivamente contemporáneo y, ya hace medio siglo, Waldron (1967, p. 139) reconoce los cambios de significado a partir de variaciones entre las relaciones léxicas: “Change of meaning, then, essentially concerns a shift of relationships between words on the one hand and integral senses on the other hand”. Tanto Waldron como Ullmann (1970) consideran que detrás de cada cambio de significado hay una cadena de causalidad que puede observarse en varios niveles (cambios en las condiciones materiales, sociales o psicológicas, etc.).

El objetivo de este proyecto es tratar de encontrar los vínculos entre los modelos de semántica y el análisis de sentimiento. Cook y Stevenson (2010) consideran que los significados experimentan cambios en sus connotaciones y resultar el germen de nuevos sentidos de las palabras de manera que estos adquieren nuevos aspectos de significado y otros caen en desuso.

En los últimos años del siglo XIX, el filólogo francés Michel Bréal (1897), uno de los fundadores de la semántica moderna, observó que los cambios en las palabras no solo responden a un carácter inflexivo o sintáctico, sino que este también se producen cambios en su significado (Bybee, 2015; Traugott y Dasher, 2002; Ynduráin, 2008). Por ello propuso un modelo para la observación de los cambios onomasiológicos y semasiológicos en las palabras. Dos tipos de cambio semántico son la *mejora* y la *peyoración*. En estos procesos de cambio, el sentido de una palabra cambia para volverse más positivo o negativo, respectivamente. En relación con el sentimiento, son relevantes los cambios onomasiológicos que fueron organizados en tres ejes principales: *peyoración/mejora*, *restricción/expansión* y *metáfora/metonimia*. El eje más relevante para este trabajo es el primero ya que referencia a variaciones en el sentimiento de las unidades a lo largo de una línea temporal. El primer eje, según Traugott y Dasher (2002) se define de la siguiente forma:

- a) La *peyoración* es la tendencia aumentar las connotaciones más negativas al significado de una palabra. Por ejemplo, la palabra “bourgeois” pasó de definir a alguien con los valores de la clase media francesa a adquirir un significado similar a “philistine”, una persona que se caracteriza por su materialismo y desdén hacia los valores intelectuales. Otro caso sería el de “mistress” que pasó de referir a las mujeres cabeza de familia para pasar a referirse despectivamente a la pareja extramarital de una persona. Asimismo, esta tendencia puede invertirse a partir de la popularización de determinados nuevos usos de las palabras. En la última

década, se está produciendo una *resignificación* de dicho término con connotaciones más positivas en relación con el empoderamiento, a partir de su uso referencial a la mujer dominante en determinadas prácticas específicas de índole sexual. Un ejemplo claro de la influencia de los cambios sociales en el lenguaje es el observado por Louwrens (1993) con el uso de la palabra “apartheid”. En afrikaans, la palabra originalmente significaba “separación”, y nada más. El término se negativizó tras la implantación de políticas racistas del gobierno sudafricano durante gran parte del siglo XX que segregaban a la mayoría negra de la minoría blanca. Las connotaciones adquiridas del significado de una palabra tienden a “empujar” su sentido original a una posición subordinada.

- b) La *mejora* es la tendencia de sentimiento con connotaciones más positivas. Por ejemplo, la palabra latina “nescius” (ignorante) entró en el inglés a partir de la palabra del francés antiguo “ni(s)ce” y gradualmente desarrolló acepciones más positivas relacionadas con la timidez, para devenir en la palabra contemporánea “nice”, con un significado totalmente positivo similar a “pleasant”.

Muchos estos casos de peyoración y mejoramiento son resultados de contagio a través de la metonimia, de manera que a partir de una relación parte/todo se asocia a variaciones contextuales de la palabra, de manera que “gold”, originalmente un metal, se asocia a expresiones evaluativas positivas, mientras que “blue” suele relacionarse a elementos negativos relacionados con un estado de ánimo depresivo.

Estos cambios son asociados a distintos factores, principalmente (a) evitar dificultades, (b) buscar expresarse con mayor claridad, o lo contrario, (c) uso eufemístico, (d) pérdida de contenido semántico (p. ej., “awfully” significaba “in a way of inspiring awe”) y (e) cambios en las actitudes socioculturales de la población (p. ej., el caso de “bourgeois”).

Cook y Stevenson (2010) hacen una adaptación del sistema de detección de la orientación semántica de Turney y Litman (2003) para cuantificar cambios históricos en la orientación semántica de las palabras en los términos recogidos más arriba. Para ello compararon las valencias obtenidas en el análisis de sentimiento de determinadas palabras en tres corpus del inglés británico que hacen referencia a tres épocas distintas: el *Lampeter Corpus of Early Modern English Tracts* que recoge textos de los siglos XVII y XVIII, el *Corpus of Late Modern English Texts* con textos entre los siglos XVIII y XX y el corpus contemporáneo paradigmático, el *British National Corpus*. Los resultados muestran que



palabras como “artful” o “plainness” adquirieron connotaciones más negativas, mientras que palabras como “nice” o “luxury” adquirieron connotaciones significativamente más positivas.

Además de cambios a largo plazo, los estudios de sentimiento pueden observar tendencias o microcambios no necesariamente permanentes que van asociados a sucesos o eventos históricos influenciados parcialmente por la opinión pública y las últimas tendencias (Bifet et al., 2011; Ynduráin, 2008). Amiri y Chua (2012) propusieron un enfoque sensible al tiempo para la detección del sentimiento en terminología en el que mostraron que la acumulación de polaridad a través del tiempo da como resultado una inferencia de polaridad más precisa que los datos de polaridad obtenidos usando métodos no basados en el tiempo. Además, se evidenció que la extracción de palabras de opinión más recientes conduce a una mejora en el rendimiento de la clasificación de sentimiento. Con dos métodos de inclusión de *word embeddings*, Hamilton, Leskovec y Jurafsky (2016abc) usaron el estudio de los vecindarios locales y globales para encontrar la causa del movimiento en espacios semánticos entre dos períodos de tiempo. Un claro ejemplo presentado para esto es el mostrado por Zamora Maass (2016) que observa, mediante el estudio de colocaciones, cómo la prosodia semántica de la palabra “gay”, implicaba colocaciones como “brilliant” y “joyous” durante 1900, y hoy en día en su contexto se encuentran palabras como “homosexual” o “lesbian”.

En cuanto a los usos contemporáneos observados a través de las redes sociales, Bollen, Mao y Pepe (2011) realizaron un análisis de sentimiento y emoción en seis dimensiones de los mensajes publicados en Twitter en la segunda mitad de 2008 y compararon sus resultados con una línea de tiempo de eventos que tuvieron lugar en ese período de tiempo y encontraron una correlación con fluctuaciones significativas de los niveles de sentimiento del público. Finalmente, Giachanou y Crestani (2016ab) exploraron la usabilidad de los métodos de series de tiempo en el seguimiento de sentimientos. En él trazaron las frecuencias y los datos descompuestos de Twitter para identificar patrones, tendencias y estacionalidad de temas del momento como “Android Lollipop”, “Michelle Obama” o “Angela Merkel” a lo largo de un año. Además, identificaron los picos en la popularidad del tema y en el sentimiento e investigamos su utilidad para detectar eventos importantes o los motivos del cambio de opinión. Su metodología supone un buen comienzo para el desarrollo de métodos que puedan detectar de manera automatizada cambios en el sentimiento en combinación con otras técnicas de lingüística de corpus para detectar las razones que lo causaron.



En cuanto a implementaciones, existe un microservicio online, *Sentiment Time Series*<sup>35</sup> que, a través de una API, permite analizar una variedad de conjuntos de datos de texto no estructurado y devuelve una trama de una serie temporal de sentimiento. Esta combinación de análisis de series temporales con el procesamiento del lenguaje natural permite visualizar cambios el sentimiento de los datos de texto no estructurados a lo largo del tiempo e incluso prever de futuras tendencias.

## 5.9 Aplicaciones comerciales de análisis de sentimiento

Además de los múltiples estudios de corte académico, con vocación mayoritariamente experimental, existen soluciones de análisis de sentimiento aplicadas comercialmente que ofrecen interfaces amigables para la explotación de datos no estructurados y, subsiguientemente, obtener información procesable. De esta manera, compañías, agencias gubernamentales y empresas de sondeo de opinión, entre otras, pueden obtener datos cuantificables sobre la opinión de sus objetivos de manera rápida y cuantitativa y así eliminar tanto conjeturas como sesgos.

Convertir la opinión en cifras permite agilizar pruebas conceptuales sobre nuevas ideas (por ejemplo, prueba de nuevas líneas discursivas para los movimientos sociales o de estrategias comunicativas de una marca) de manera que se pueden cuantificar las reacciones de una manera más rápida que un análisis manual a través de distintas estrategias, como puede ser la comparación de la respuesta ante dos campañas distintas, o los estudios de evolución sobre determinados productos, factores o conceptos. Como se puede ver en la literatura (Vohra y Teraiya, 2013), se resumen aquí algunas de aplicaciones prácticas de la disciplina:

- Gestión política: el análisis de sentimiento ayuda a las organizaciones y agencias políticas de toda índole a evaluar su fortaleza y debilidades analizando las opiniones del público y la prensa de manera que es posible extraer las opiniones de los ciudadanos sobre nuevas ideas o medidas e identificar sus fortalezas y debilidades.
- Comercio en línea: uno de los usos más populares, tanto desde una perspectiva científica como comercial, del análisis de sentimiento está en el estudio de las

---

<sup>35</sup> Se puede consultar en esta web: <https://algorithmia.com/algorithms/nlp/SentimentTimeSeries>

actividades de comercio electrónico. Muchas plataformas de comercio online disponen de apartados en los que los usuarios pueden enviar su experiencia y valorar las características de los productos adquiridos en las que reseñan los principales pros y contras del producto y asignan una puntuación global o pormenorizada. Los clientes pueden ver fácilmente las opiniones y la información de recomendación sobre el producto completo según la experiencia de otros comprados, así como una evaluación de las características específicas del producto. De esta manera, el análisis de sentimiento permite convertir en promotores a los clientes satisfechos o insatisfechos mediante el análisis automatizado de sus reseñas. Además, la voz del cliente (*voice of the customer*, VOC) ayuda a identificar nuevas oportunidades para invenciones de productos pues permite identificar fácilmente los requisitos funcionales de los productos y algunos no funcionales, así como su rendimiento y costo.

- Voz del mercado (*Voice of the Market* o VOM): en este caso, en lugar de centrarse en el estudio de una marca propia, se realiza un análisis sobre el sentimiento atribuido por el público a la competencia, lo cual supone una ventaja competitiva para adelantarse en el desarrollo de nuevos productos, obteniendo la opinión del cliente en tiempo real a través de, por ejemplo, las redes sociales.

Un ejemplo sería el de Zhang, Xu y Wan (2013) que proponen un sistema de búsqueda de debilidades a través del análisis de sentimiento basado en aspectos a partir de las revisiones de los productos en las redes sociales.

- Gestión de reputación de marca: las plataformas de contenidos generados por el usuario permiten una horizontalización en la expresión y peso de las opiniones ya que las conversaciones de uno a muchos se llevan a cabo en línea a gran velocidad, no dependiendo de empresas de publicidad y relaciones públicas. Además de un desafío, esto es una fuente de oportunidades para que las organizaciones administren la reputación de la marca ya que los clientes pueden dañar o mejorar su reputación. De esta manera, el análisis de sentimiento permite determinar el desarrollo de la reputación del servicio de la organización a partir de la información obtenida en línea.
- Educación: el análisis de sentimiento también puede tener múltiples aplicaciones educativas. Las administraciones de los centros pueden detectar la parte emocional del trabajo escrito del alumnado de manera que se pueden anticipar a distintos conflictos, problemas de motivación o de autoestima. Asimismo, no solo

las escuelas, sino también las plataformas de tipo MOOC<sup>36</sup> pueden procesar las observaciones en las evaluaciones al profesorado y las autoevaluaciones de los alumnos (Fernández-Cruz, 2017).

Finalmente, en la actualidad existen bastantes herramientas que pueden ayudar a las empresas a extraer y analizar información de los blogs sobre tendencias a gran escala en las opiniones de los clientes sobre los productos. Aunque el análisis detallado de distintas soluciones comerciales al problema del análisis de sentimiento sale del marco de este trabajo, se hará una breve mención a herramientas comerciales disponibles en Internet. Este es el caso de *SenticNet*<sup>37</sup>, *Luminoso*<sup>38</sup>, *Factiva*<sup>39</sup>, *Attensity*<sup>40</sup> y *Converseon*<sup>41</sup>. Existen, además, algunas herramientas gratuitas disponibles al público como *TweetFeel* o *Socialmention* (Cambria, Schuller, Xia y Havasi, 2013).

## 5.10 Conclusiones

En el presente capítulo se ha revisado exhaustivamente la historia del desarrollo del análisis de sentimiento como disciplina, así como sus principales enfoques de trabajo. Existen tres niveles de análisis de sentimiento según la entrada (documento, oración y entidad) que son enfocados de manera distinta durante su proceso de asignación automática de orientación semántica. En cuanto a la granularidad de los resultados, existen dos enfoques principales: el primero llamado *Thumbs up/Thumbs down* en la literatura que ofrece resultados binarios o ternarios (positivo, negativo y, dado el caso, neutro) y el segundo tipo de sistemas ofrece resultados positivos dentro de una escala (p. ej., de 0 a 10).

La tercera clasificación de los sistemas es el método de análisis. Aquí concurren dos enfoques principales: uno basado en el aprendizaje automático basado en algoritmos

---

<sup>36</sup> Los MOOC (acrónimo en inglés de *Massive Open On-line Course*) son cursos en línea dirigidos a un número ilimitado de participantes a través de Internet según el principio de educación abierta y masiva (definición de Wikipedia, 2018).

<sup>37</sup> Puede consultarse en: <http://sentic.net>

<sup>38</sup> Puede consultarse en: <http://luminoso.com>

<sup>39</sup> Puede consultarse en: <http://dowjones.com/factiva>

<sup>40</sup> Puede consultarse en: <http://attensity.com>

<sup>41</sup> Puede consultarse en: <http://converseon.com>

estadísticos (supervisado, semisupervisado, no supervisado y reforzado), dependen en gran medida de conjuntos de datos de entrenamiento etiquetados previamente y el rigor de sus resultados dependerá, entre otros factores, del volumen de estos. El enfoque basado en léxico será el que se utilizará en este proyecto y su funcionamiento se fundamenta en recopilaciones lexicográficas de entradas y reglas contextuales etiquetadas según el sentimiento que denotan. Es relativamente frecuente encontrar sistemas híbridos de análisis que utilizan técnicas de ambos enfoques.

No existe una solución definitiva para el problema de la minería de opinión. Los sistemas de análisis de sentimiento son, en gran parte, sistemas diseñados para resolver problemas ad-hoc y no hay formas universales de evaluar el desempeño de estos mediante una comparación universal entre los distintos trabajos. Es por esto, que algunos enfoques tendrán mejor o peor rendimiento según la información de entrada. Sin embargo, desde una perspectiva lingüística, los enfoques basados en léxico permitirán sacar mejor partido al estudio de las distintas teorías sobre el lenguaje evaluativo. Gracias a estos, es posible implementar aplicaciones prácticas y estudiarlas desde una nueva perspectiva. Sobre esta relación, brevemente, se pueden establecer relaciones claras entre la teoría de la valoración de Martin y White (2005), los *Contextual Valence Shifters* de Polanyi y Zaenen (2008); y la gramática de patrones (Hunston et al., 2005) o implementaciones semánticas como *FrameNet* (Fillmore et al., 2002) revisados en este trabajo. Asimismo, las recopilaciones léxicas pueden ser reutilizadas inversamente para elaborar no solo productos terminográficos como tesauros o bases de datos terminológicas, sino que también tienen múltiples aplicaciones informáticas en los campos de la extracción de información o la *Web Semántica*, entre tantos.

Subsiguientemente, se ha abordado el problema del dominio a partir del estudio del dominio específico económico-financiero. Aquí se revisan los principales trabajos realizados sobre el dominio económico-financiero, objeto de esta tesis, así como sus principales características. El sentimiento es especialmente sensible al dominio, pues las valencias del análisis variarán dependiendo del tipo de lenguaje especializado (o no) que se use. Muchos estudios de este tipo están enfocados a buscar correlaciones entre el discurso especializado (informes o prensa especializada) y el rendimiento de las acciones. Si embargo este no es el caso de este estudio, que se centra en el estudio de la evolución del sentimiento de una serie de términos económicos, por lo que fue pertinente revisar las características del dominio y las distintas aproximaciones revisadas en la literatura en las

que observa longitudinalmente la evolución del sentimiento y los posibles cambios detectados en su sentimiento.

Se analizaron los modelos teóricos en que los significados experimentan cambios en sus connotaciones y resultar el germen de nuevos sentidos de las palabras de manera que estos adquieren nuevos aspectos de significado y otros caen en desuso. El eje *mejora/peyoración* dentro de los estudios sobre el cambio semántico es fundamental para este trabajo. En estos procesos de cambio, el sentido de una palabra cambia para volverse más positivo o negativo, respectivamente a lo largo de un intervalo de tiempo determinado.

El análisis de sentimiento es un campo de trabajo multidisciplinar con un carácter principalmente experimental. Sin embargo, existe una gran cantidad de aplicaciones prácticas de sus técnicas que pueden ayudar a resolver problemas cotidianos. De manera breve, se revisaron las aplicaciones más destacadas del análisis de sentimiento y se recopilaron algunas de las explotaciones comerciales más conocidas.

## 6. Métodos de investigación

- « En tout cas, c'est idiot de faire un travail que des machines pourraient faire.  
- Il faut construire des machines [...]. Qui le fera ?  
- Oh ! Évidemment [...]. Pour faire un œuf, il faut une poule, une fois qu'on a la poule, on peut avoir des tas d'œufs. Il vaut donc mieux commencer par la poule.  
- Il faudrait savoir [...] qui empêche de faire des machines. C'est le temps qui doit manquer. [...] S'ils avaient le temps de construire des machines, après ils n'auraient plus besoin de rien faire. [...] Ils travaillent pour vivre au lieu de travailler à construire des machines qui les feraient vivre sans travailler. »  
Boris Vian (1920-1959), *L'écume des jours* (1947).

### 6.1 Paradigmas y enfoques metodológicos

La complejidad teórica y metodológica que marcan los objetivos del proyecto requieren especificar los fundamentos e instrumentos de investigación para guiar el proceso científico de manera eficaz y adecuada a los objetivos propuestos. El diseño de la investigación debe ser acorde a las estrategias empíricas convenientes, y afin con el modelo conceptual sobre el que se asienta pues si esto fuera de otra forma, el diseño de una estrategia de actuación sin un modelo conceptual previo nos llevaría a una interpretación posterior poco determinada y mal fundamentada. Es pertinente entonces definir epistemológicamente el enfoque y el paradigma en el que se encuadra este trabajo.

Para abordar el problema metodológico se debe partir del enfoque conceptual que se considere más adecuado, pues el paradigma que se asuma condicionará los procedimientos de estudio en el proyecto (López Noguero, 2002). Un paradigma, según Kuhn (1970), se entiende como un conjunto de creencias y actitudes, esto es, una determinada visión del mundo, que es compartida de manera implícita por un grupo de científicos que aplican metodologías similares a la hora de buscar respuestas sobre los problemas. Son tres los paradigmas principales reconocidos en la tradición investigadora: positivista, interpretativo y sociocrítico o crítico. Según Ricoy Lorenzo (2006), (a) el *paradigma positivista* se califica de cuantitativo, empírico-analítico, racionalista, sistemático gerencial y científico-tecnológico, (b) el *paradigma interpretativo* se

considera como interpretativo-simbólico, cualitativo, naturalista, humanista y fenomenológico; y (c) el *paradigma crítico* se concibe como comprometido, naturalista y ecológico.

Existen dos enfoques metodológicos principales: cuantitativo y cualitativo. La distinción entre ambas metodologías va más allá del hecho de usar cifras o datos no cuantificables. Dörnyei (2007) expone que tras estas se confrontan múltiples factores: desde la orientación ideológica de fondo, a la recolección de muestras y la naturaleza de estas, al procesamiento de los datos. La investigación cuantitativa se centra en procedimientos para obtener datos numéricos que serán analizados mediante métodos estadísticos. Por ejemplo, una encuesta realizada mediante un cuestionario que posteriormente se analiza utilizando un paquete estadístico informatizado. Las investigaciones cualitativas implican la recopilación de datos de carácter no numérico y abierto (*open-ended*) que deben ser estudiados mediante métodos no estadísticos. Un ejemplo de esto serían entrevistas transcritas que posteriormente se analizan mediante una rúbrica. Además, como es el caso aquí, los métodos mixtos combinan investigaciones cualitativas y cuantitativas a la hora de recoger datos o analizarlos. Por ejemplo, una combinación de entrevistas y encuestas.

Si bien existen diferencias de calado, Dörnyei (2007) revisa la literatura sobre las oposiciones paradigmáticas (ideológicas, de categorización y de percepción de la diversidad) y concluye aseverando que, a fin de cuentas, ambos enfoques metodológicos habitan el mismo mundo y forman parte de una misma continuidad. Desde comienzos del siglo XX, los estudios en ciencias sociales utilizaban enfoques metodológicos de ambos tipos. A mediados del siglo XX, la filosofía de la ciencia se había concentrado en la viabilidad del método científico, de forma que se proponían justificaciones para alcanzar la veracidad de las teorías científicas y se intentaba descubrir las causas filosóficas del funcionamiento de la ciencia.

El enfoque cuantitativo, influido por el positivismo de Comté se basa en los fenómenos naturales mensurables y sus propiedades. Estos fenómenos son independientes de la percepción del sujeto. Las evidencias se recogen a través de datos de carácter cuantitativo que más tarde se interpretan a la luz de una serie de significados atribuidos (codificados) a los conceptos contextualizados a través de los cuales adquieren un sentido. De esta forma, la información obtenida a través de evidencias empíricas (hechos positivos) es interpretada *a posteriori* mediante la razón y la lógica, constituye la fuente exclusiva de todo conocimiento cierto. Los enfoques metodológicos cuantitativos

tienden a generalizar las tendencias en un universo de una forma soberanamente estructurada. La diversidad aquí (ya sea la realidad humana o el mundo físico) se aborda de acuerdo con la aceptación de la existencia de leyes generales. De esta manera, los enfoques cuantitativos deben recopilar muestras lo suficientemente amplias como para que los resultados reflejen los puntos en común que existen en los datos (Dörney, 2007).

Karl Popper, uno de los primeros opositores del positivismo lógico en el siglo XX, repudió la forma clásica de observación / inductivismo del método científico en favor de la falsación empírica. En las últimas cuatro décadas, las tradiciones científicas naturalistas han ganado peso en un mundo científico previamente copado por el paradigma positivista. El modelo cualitativo surge como alternativa al paradigma racionalista, dado que existen problemáticas, cuestiones y restricciones que no se pueden explicarse tan solo a través de un enfoque cuantitativo.

Según Rivas Flores (1990), las investigaciones cualitativas conceptualizan de manera distinta los objetos de estudio al situar la construcción del conocimiento en consideraciones plausibles en lugar de consideraciones demostrativas. Las investigaciones cualitativas deben entonces describir, desarrollar e interpretar los datos recogidos de manera que formen parte contextualizada del conocimiento de forma que se alejen de posturas deterministas donde el saber se entiende como una construcción compartida. En otras palabras, el método es el ente que se acerca a la realidad y no al contrario. Por esto, las estrategias de investigación requieren situar los elementos axiológicos, deontológicos y metodológicos de manera que estén desarrollados de una manera acorde con el fenómeno o realidad que se observa.

La irrupción de nuevos sistemas culturales y las posibilidades que actualmente ofrecen las tecnologías de la información requieren nuevos enfoques para comprender los pormenores desde distintas perspectivas y abriendo campo a enfoques metodológicos mixtos que combinan elementos cuantitativos y cualitativos. Si bien las herramientas utilizadas tradicionalmente por los métodos positivistas aportan información objetiva y mensurable de la realidad, evitando así sesgos apriorísticos, estos no permiten profundizar en la naturaleza del objeto de investigación y conocer facetas más complejas en una realidad versátil e interconectada. Por lo tanto, el conocimiento generado a través de mediciones puede interpretarse desde elementos demostrables de manera teórica y pasar a cumplir con los objetivos mediante una triangulación con información más interpretativa y subjetiva. Esta es la razón por la que los enfoques mixtos implican combinaciones de enfoques metodológicos cuantitativos y cualitativos a la hora de



recopilar datos o para llevar a cabo los análisis (Miles y Huberman, 1994). Entre los beneficios de los enfoques mixtos se encuentran: por un lado, un aumento de las fortalezas que, a su vez, reduce las debilidades de la investigación de manera que las generalizaciones cuantitativas pierden simpleza y ganan perspectiva a través de estudios de casos particulares en un contexto más amplio. Por otro lado, se produce una mejora de la validez de los estudios ya que se puede contrastar la evidencia a través de múltiples métodos que aumenten la generalización de los resultados.

Centrando el debate en la lingüística como disciplina científica, tanto el comportamiento humano como las distintas variables en las que se encuadran los fenómenos lingüísticos se hallan inmersas en una enorme complejidad y circunstancias cambiantes que ha generado discusiones paradigmáticas, especialmente en las últimas cinco décadas. La distinción basada en la teoría (*theory-driven*) en contraste con la distinción basada en datos (*data-driven*) pone de manifiesto el conflicto, aún caliente, entre el racionalismo y empirismo. En lingüística, el paradigma tradicionalmente utilizado fue el primero. Por ejemplo: a principios de la década de 1960 las teorías lingüísticas, destacando la *gramática transformativa-generativa* de Chomsky, apoyaron sus postulados con ejemplos basados en la intuición. Fue a partir del desarrollo de técnicas computacionales cuando se ha conseguido favorecer el empleo de enfoques más diversos y empíricos, inaugurando un amplio debate sobre la naturaleza del conocimiento en las ciencias lingüísticas. La polarización de estos dos enfoques fue caricaturizada por Fillmore (1992) e ilustrada a partir de dos personajes: el lingüista del sillón y el lingüista de corpus. Mientras que el lingüista del sillón piensa que lo que el lingüista del corpus está haciendo no es interesante, el lingüista del corpus cree que lo que el lingüista del sillón está haciendo es falso. Si bien Xiao y Hu (2015) reflejan excepciones extremistas de ambos campos<sup>42</sup>, la mayoría de las discusiones epistemológicas son conscientes de que los dos enfoques son complementarios entre sí (p. ej., Leech, 1992; Meyer, 2002; McEnery et al., 2006). El uso de estas herramientas mejora de sobremanera las posibilidades de interpretación que se pueden efectuar de los resultados recopilados en la

---

<sup>42</sup> Esto se caracteriza con dos afirmaciones antagónicas, la de Chomsky: “corpus linguistics doesn’t mean anything” (Andor, 2004) vs. nada significativo puede hacerse sin un corpus (Murison-Bowie, 1996).

investigación, obteniendo una mejor exploración de la realidad y aportando perspectivas originales para los hallazgos.

La motivación y buena parte de la influencia de este estudio parte de un interés en la crítica y la identificación del potencial de cambio social<sup>43</sup>. Por este motivo, la presente investigación se enmarca en la tradición *interpretativa* puesto que la finalidad no es generalizar sino comprender y dar explicación a las variaciones en la orientación semántica de la lengua mediante la interpretación cualitativa de una gran cantidad de datos codificados numéricamente. Este enfoque metodológico mixto combina elementos cuantitativos y cualitativos ya que se van a observar el sentimiento asociado a una serie de palabras desde un punto de vista eminentemente cuantitativo y, posteriormente, van a ser interpretados mediante la relación de estos con una línea temporal en la que ocurrieron determinados cambios sociales. Además, es necesario tener en cuenta la noción de serendipia, la cual es intrínseca a la extracción de expresiones multipalabra. Esta refuerza el carácter mixto (cuantitativo-cualitativo) pues a partir de los datos cuantitativos, se obtienen resultados imprevisibles, lo que impide la redacción de una hipótesis definida característica de una investigación cuantitativa definida por sus objetivos.

Para concluir esta sección, es importante poner sobre la mesa el papel de la propia investigación y los roles tanto del investigador como el objeto de estudio, esto es, los textos. La naturaleza del objeto de estudio y la investigación en sí deben entenderse no solo como el mero estudio de la evolución de la orientación semántica de una serie de palabras de carácter especializado en lengua inglesa, sino como una forma de aproximarse desde el análisis de sentimiento a una realidad histórica contemporánea, destacando los roles de los componentes políticos, sociales e ideológicos en los textos. De esta forma, se entiende la investigación como un uso de la ciencia al servicio del público, con vistas a observar una parcela de la realidad que favorezca, aun cuando esto sea realizado de una manera descriptiva, la comprensión del espíritu de nuestros tiempos a través del uso del lenguaje.

---

43 Toret Medina (2015) fue uno de los trabajos que inspiró esta tesis. Su trabajo fue publicado en una colección de trabajos coordinada por Manuel Castells en la Universitat Oberta de Catalunya.

## 6.2 Diseño de la investigación

El Capítulo 1 recogía el objetivo general y los cuatro objetivos específicos propuestos para este trabajo. Dichos objetivos se abordaron a través de un enfoque mixto de carácter interdisciplinar (lingüístico y computacional). El diseño de la investigación se puede resumir en el desarrollo de un modelo que involucra cuatro pasos:

1. Recopilación de un lexicón de dominio específico para el software de análisis de sentimiento:
  - 1.1. Diseño y construcción de un corpus de dominio específico.
  - 1.2. Extracción terminológica a través de distintas técnicas de lingüística de corpus.
2. Mediante series temporales, analizar la evolución longitudinal del sentimiento de una serie términos encontrados en oraciones procedentes del corpus
3. Determinar los factores extralingüísticos (sociales, políticos, económicos, etc.) que influyen los cambios.
4. Interpretar cualitativamente, con detalle, las variaciones sociolingüísticas diacrónicas de las condiciones de uso de los términos seleccionados (frecuencia, eventos, colocaciones, usos del lenguaje).

Para este estudio, se diseñó, recopiló y analizó un corpus de dominio específico de más de 24 millones de palabras: el denominado *Esmeraldas Corpus of Business English*, (ECBE). Esta recopilación incluyó las secciones económicas de *The Guardian* y *The New York Times* entre enero de 2007 y diciembre de 2015. Cada uno de los artículos se recogió en un fichero de texto plano individual y se indexó sistemáticamente por fecha y publicación. Este corpus fue la fuente principal de datos, pues a partir de él: (a) se extrajo la terminología de dominio específico para construir el lexicón, (b) se obtuvieron los textos que constituyen el *gold standard* que sirvió para evaluar el rendimiento del software de análisis de sentimiento, (c) se extrajeron las oraciones que incluían los términos cuyo sentimiento sea objeto de estudio en esta tesis.

Para llevar a cabo la construcción de un lexicón de dominio específico, se comenzó con la aplicación de *KeyWords*, el potente extractor de terminología incluido en la suite de *The Sketch Engine* que está basado en patrones gramaticales y frecuencia. Posteriormente se extrajeron todos los elementos detectados en el ECBE, tomando al corpus *EnTenTen15* como corpus de referencia. Como resultado, se extrajo un conjunto de candidatos a término que incluya 30 mil términos: 15.000 palabras y 15.000

expresiones multipalabra. Para refinar los resultados, con la ayuda de un experto, a la extracción automática de candidatos a término le siguió un filtrado manual conciso mediante distintas técnicas terminográficas. Este proceso se describirá con profundidad en el apartado 6.5.

Es necesario reseñar que no se trata de un trabajo terminográfico tradicional, pues no hay interés en extraer una lista completa de términos, sino la lista de términos que contengan sentimiento dentro el dominio económico-financiero cuando la orientación difiera del sentimiento que tienen en dominios distintos. Por ejemplo, términos como “sales” o “bank” son irrelevantes en este caso, pues ambos son neutros tanto en el dominio específico como en el dominio general. Sin embargo, las palabras con sentimiento neutro como “asset” o “fix” se añadieron a la selección, ya que los positivos en el dominio general son neutros aquí y deben ser recogidas. Por otro lado, los términos con una polaridad determinada que ya estén presentes en el lexicón base de *Lingmotif* no fueron contabilizados (por ejemplo, “recovery” o “crisis”). En pocas palabras, se descartaron todos los candidatos que resultaron no tener un significado especializado, junto con los términos neutros y los términos que no transmitan una orientación semántica diferente en el dominio económico-financiero.

A la recopilación final del *plug-in* lexicón se le llamó *LexiEcon* y se recogió en un documento de texto plano estructurado en formato CSV con codificación UTF-8, según las especificaciones requeridas por *Lingmotif*. Para evaluar su rendimiento, se desarrolló un *gold standard*. Con el fin de llevar esto a cabo, se extrajo una muestra aleatoria de 400 oraciones del corpus ECBE. Dos evaluadores fueron instruidos para que solo tomen en cuenta la información disponible en las oraciones y anoten las oraciones acordando una respuesta final de manera categórica (positivo/negativo/neutro). Una vez hecho esto, se procedió a analizar dichas oraciones en dos pasos, primero aplicando solo el lexicón general de *Lingmotif* (*Lingmotif-Lex*) y posteriormente los resultados analizados por el conjunto de *Lingmotif-Lex* y *LexiEcon*. Se compararon los resultados aportados por los evaluadores con cada análisis de sentimiento y se obtuvieron cifras de rendimiento y exhaustividad a partir de matrices de confusión.

El fin último de este trabajo es detectar y analizar las variaciones en el sentimiento de los términos a través de un estudio longitudinal a través de análisis estadísticos descriptivos. Se propone la anotación de los resultados del análisis de sentimiento a partir de un conjunto inicial de términos que fueron analizados según su aparición anual entre 2007 y 2015. Se anotaron las frecuencias, las colocaciones y los resultados del análisis de

sentimiento para poder realizar un análisis cualitativo adicional que permita de determinar los factores externos que intervienen en el cambio de actitud hacia estos elementos léxicos.

Los últimos pasos fueron, entonces, un análisis longitudinal de la evolución del sentimiento en una serie de términos del dominio económico-financiero que destacaron durante la crisis. Las investigaciones longitudinales son una serie imprecisa de métodos que comparten una cosa en común: la información se recopila en torno a unos objetivos durante una serie de puntos temporales. Estos datos tienen que ser de la misma naturaleza con vistas a ser comparados. El modelo longitudinal es frecuente en modelos mixtos, ya que requiere que, de manera cualitativa, se aporten introspecciones relacionadas con las tendencias estadísticas observadas.

Los objetivos de los trabajos longitudinales tienen dos propósitos principales: *describir* patrones de cambio y *explicar* las relaciones causales. En esta tesis se propone *describir* los cambios en las puntuaciones de sentimiento en palabras especializadas y *explicar* estos cambios a través de su relación con eventos sociopolíticos relacionados con la crisis. En este caso, se habla de estudios longitudinales transversales o de tendencias que analizan cambios a través del tiempo dentro de una población determinada simplemente a través de una toma de muestras repetida a lo largo del tiempo (Dörnyei, 2007).

En cuanto el análisis de datos, se propone el estudio del sentimiento a través de series temporales de manera que se busquen patrones de cambio en casos individuales y generar los hallazgos a través de periodos temporales. La modelización de la dinámica de la opinión pública mediante series temporales puede realizarse gracias a la agregación de la puntuación de sentimiento de texto obtenida para un conjunto preestablecido de fuentes de datos. La naturaleza y cantidad necesaria de estos textos dependerá en cada caso de las dinámicas concretas que se quieran abordar. La diferencia entre identificar correctamente un modelo o no radica principalmente en identificar los diferentes fenómenos concurrentes en la serie de tiempo, así como identificar cada uno de sus cambios estructurales y su naturaleza (Fernández-Cruz y Baixauli-Pérez, 2018).

Un cambio estructural, por ejemplo, un discurso político o una acción ciudadana, perturba la componente determinista de la serie temporal, provocando una pérdida de poder en la identificación de la dinámica subyacente. Atkinson et al. (1997) desarrollaron técnicas para identificar cambios de nivel y la estimación de parámetros ARIMA (*Autoregressive Integrated Moving Average*) y Harvey et al. (2002) consideran pruebas

para detectar *outliers* aditivos e innovacionales ante cambios estacionales en la media aritmética (Fernández-Cruz y Baixauli-Pérez, 2018). La dinámica de las puntuaciones de sentimiento, así como el estudio de cualquier dinámica en la naturaleza, está potencialmente influenciada por una gran variedad de fuerzas exógenas, las cuales provocarán cambios estructurales atípicos de diferente tipo, magnitud, posición y duración. Para observar estos cambios se seleccionó una serie de términos previsiblemente sujetos de haber sufrido cambios en la evolución del sentimiento (Tabla 9). Esta selección se realizó de manera intuitiva (estas palabras resultaban interesantes por la coyuntura histórico-económica), tras haber tomado como base la observación de las *palabras clave* del corpus.

Para observar la evolución de la orientación semántica de cada uno de los contextos de estos términos a lo largo de la serie temporal, se registraron las puntuaciones de salida (entre 0 y 100) obtenidas tras el ser procesadas por *Lingmotif* + *LexiEcon*. Para abordar esta tarea, se presenta un modelo completo de análisis de sentimiento en series temporales:

- (a) Del corpus ECBE se extrajo la totalidad de las oraciones que incluyan estos términos y se recopilaron diez conjuntos de datos (*datasets*), uno por cada término de estudio, estructurados en cuatro columnas: (ID) número de identificación, (TEXT) oración, (DATE) mes y año de publicación y (PAPER) nombre de la cabecera.
- (b) Posteriormente, estos conjuntos de datos fueron analizados automáticamente con *Lingmotif* más el *plug-in* de dominio específico *LexiEcon*. Este devolvió una puntuación de sentimiento total para cada oración en un rango de 0 a 100. La primera cifra hace referencia al valor negativo extremo del sentimiento dentro de la escala y el segundo al valor positivo extremo.
- (c) Para cada término, se llevó a cabo un análisis estadístico mediante la librería *Pandas* del lenguaje de programación *Python*. Las puntuaciones totales de sentimiento fueron agregadas estadísticamente mediante una media aritmética anual. De esta manera se pudo observar sin estacionalidad la variación anual de las puntuaciones del sentimiento del contexto de cada término objeto de estudio.
- (d) Mediante herramientas de procesamiento de corpus, se recopilaron los datos anuales de cada una de las unidades observadas: se registraron las veinte

colocaciones asociadas con cada una de ellas para cada año<sup>44</sup>, la cantidad anual de coocurrencias y el coeficiente *logDice* de cada una. En cada una de las tablas se destacaron las coocurrencias que tengan orientación semántica positiva o negativa. Posteriormente, se observaron ejemplos representativos en el corpus que sean correspondientes a la prosodia semántica de cada término y se trazaron líneas de correspondencia histórica que motiven estas variaciones. Se ha seleccionado el coeficiente *logDice* como indicador estadístico de colocación por ser un indicador que marca no solo la frecuencia, sino la exclusividad de colocación entre las dos unidades léxicas. *logDice* es una medida estandarizada (rango de 0-14) que, al contrario de otras medidas como *T-Score* o *MI*, evita el sesgo producido por el tamaño del corpus. De esta forma, es posible comparar las puntuaciones obtenidas entre los resultados obtenidos en los distintos subcorpus anuales que tienen tamaños muy distintos (Gablasova, Brezina y McEnery, 2017). Además de aportar mejores valores de transparencia y replicabilidad, permitió validar los datos mediante triangulación de los resultados cuantitativos y los cualitativos con el fin de establecer relaciones entre los datos y los acontecimientos históricos.

<i>Término de análisis</i>	<i>N</i>
Credit	6764
Debt	8636
Inflation	4844
Housing	3256
Markets	11377
Investor	14901
Public	9893
Reform	2670
Spending	6262
Company	81036

**Tabla 9** Lista de términos de análisis y número de oraciones analizadas.

De acuerdo con Neale y Flowerdew (2003) o Dörnyei (2007), el objetivo de implementar este diseño metodológico mixto es aumentar las fortalezas y reducir las debilidades que presentan algunos diseños experimentales de carácter longitudinal, de manera que se ejecute un modelo que combine y complemente técnicas extensivas e intensivas. El elemento extensivo sería el análisis cuantitativo del sentimiento (el trabajo

<sup>44</sup> Sustantivos comunes, adjetivos y verbos.



estadístico descriptivo) y el elemento intensivo que hace referencia al trabajo cualitativo relacionado directamente con los textos.

Esta triangulación hace referencia a lo que Dörnyei llama *validación mediante convergencia* y permitió asegurar la validez de la investigación. Si un hallazgo subsiste a una serie de pruebas mediante diferentes métodos puede considerarse más válido que una prueba analizada mediante un solo método. Como se verá en el Capítulo 7, esta triangulación no está libre de divergencias, y la exploración de sus aristas permitió profundizar en la observación de las complejas dinámicas tras los cambios en la orientación semántica de las unidades léxicas especializadas.

Para concluir, los estudios longitudinales son una oportunidad fantástica para observar el desarrollo de la orientación semántica a nivel micro y macro. Como resultado, se espera obtener pistas que indiquen los motivos históricos externos, o también *perfiles narrativos* contextualizadores (Dörnyei, 2007), relativos a dichos cambios en la orientación semántica. A través del análisis, se espera observar tendencias y factores (cambios de coyuntura económica, discursos, protestas, cambios en la agenda política, etc.) que queden reflejadas en las noticias y provocaron cambios en la orientación semántica de las unidades. Los *perfiles narrativos* son un medio accesible y potente de desarrollar un conocimiento completo del estado de la cuestión y llegar a conclusiones más convincentes y válidas.

### 6.3 Muestra: el Esmeraldas Corpus of Business English

En esta primera parte del capítulo se describe y justifica el diseño e implementación de un corpus de textos periodísticos de divulgación económica en lengua inglesa durante la crisis: el *Esmeraldas Corpus of Business English* (ECBE). El objetivo aquí es la definición de los criterios de selección de textos y explicar las técnicas de procesamiento, organización y limpieza desde su extracción desde páginas web hasta su producto final en texto plano. El corpus resultante incluye 42.193 artículos de noticias de la sección *Business* de The Guardian y The New York Times en archivos de texto plano etiquetados según su fecha de publicación en el intervalo entre enero de 2007 y diciembre de 2015. El corpus resultante incluye 21 millones de tokens, siendo esta una colección mediana<sup>45</sup>

---

<sup>45</sup> Calificación subjetiva, en términos contemporáneos.



de textos electrónicos de acuerdo con los criterios clásicos (p.ej., McEnery, Xiao y Tono, 2006) en la lingüística de corpus: electrónico, tamaño, autenticidad y criterios específicos.

El motivo principal de diseño de un corpus desde cero fue la no disponibilidad de un corpus textual que se ajuste a las necesidades específicas marcadas por el diseño de la investigación. El producto final debe servir como fuente de datos única para así permitir el cumplimiento de los objetivos de esta investigación: la construcción de un lexicón de palabras de sentimiento en el dominio de la economía en lengua inglesa y, a su vez, servir como plataforma textual sólida para observar la evolución del sentimiento en distintas unidades léxicas propias de dicho dominio.

En comparación con otros recursos, el uso de este corpus es representativo e idóneo para llevar a cabo el objetivo principal: el estudio longitudinal de la evolución del sentimiento de una serie de términos. Ya que se van a analizar las oraciones procedentes del propio corpus, lo ideal es que el lexicón de sentimiento se ajuste lo más ceñidamente a sus características textuales. Así, mediante el uso de técnicas de corpus se podrán extraer, por ejemplo, elementos contextuales y conocer qué otras palabras acompañan a los términos seleccionados para su estudio, de manera que se añadan al léxico una gran cantidad de expresiones multipalabra que trasciendan las colocaciones tradicionales y se dirijan a representar usos prosódicos del lenguaje (p. ej., “the burst of the housing bubble”, “the housing bubble has burst”). Se habla entonces de observar *clusters*, líneas de concordancia de la búsqueda de una palabra específica o una construcción gramatical, entre otras. Esta información se puede obtener a través del estudio contextualizado de los términos que permite el corpus, en detrimento de una observación de los elementos en aislamiento. De esta manera, un buen procesamiento del corpus aportará la cantidad de términos necesaria y tendrá una mejor contextualización del estudio (esto es, colocaciones, coligaciones, etc.) aportando, previsiblemente, mejores cifras de exhaustividad.

Por ello, es necesario un corpus lo suficientemente representativo y extenso que aporte el nivel de detalle necesario para conocer las especificidades lingüísticas de disciplina que se quiere estudiar. Según Bowker y Pearson (2002) en lo que respecta a la amplitud no existe un tamaño ideal prefijado pues el volumen de un corpus dependerá principalmente del propósito de su estudio. La mejora de la cantidad y calidad de la extracción de terminología viene motivada por una fuente textual mayor, siguiendo la máxima de Sinclair (1991) “bigger is always better”, siempre que venga condicionado por las características y los objetivos. No obstante, hasta un corpus pequeño puede ser un

recurso muy útil si este está bien diseñado. En especial, según los autores, se acepta que los estudios de corpus destinados a lenguajes para fines específicos pueden ser más pequeños que los utilizados para los estudios de carácter general.

Las bases del corpus de trabajo deben cumplir con cuatro requisitos principales. Según McEnery, Xiao y Tono (2006, p. 5):

an increasing consensus that a corpus is a collection of (1) machine readable (2) authentic texts (including transcripts of spoken data) which is (3) sampled to be (4) representative of a particular language or language variety.

Según los rasgos definitorios propuestos, la condición de *electrónico* es la primera noción definitoria: estos se almacenan digitalmente en maquinaria informática y son tratados mediante software de análisis de corpus. Que el corpus sea electrónico, en el siglo XXI, parece casi un hecho ordinario; sin embargo, resulta necesario reseñar que hasta hace escasamente dos décadas los estudios de lingüística de corpus eran realmente innovadores. El giro empírico que las técnicas de procesamiento informático del lenguaje natural han supuesto en los últimos 40 años ha permitido que ejemplos reales de texto puedan ser consultados de manera sencilla automática, en lugar de observar manual y linealmente montañas de textos impresos. Por esto, Leech (1992) señala que la investigación basada en corpus tiene un peso metodológico suficiente pues permite verificar empíricamente los resultados de manera objetiva. Además, los planteamientos realizados a través de la introspección que antes no podían evidenciarse con claridad ahora pueden ser contrastados mediante el acceso a miles de muestras de lenguaje real de cualquier especialidad imaginable. Aparte de alcanzar conclusiones más rápidas, permite observar el lenguaje desde perspectivas inéditas, enfocándose exclusivamente en las partes reseñables del texto. En el caso de este estudio, la consulta lineal de diez términos incluidos en 41.000 artículos periodísticos impresos hubiera requerido el trabajo de varios años de un equipo de investigación completo. Sin embargo, este estudio ha podido llevarse a cabo por una sola persona en una fracción del tiempo (y en un espacio considerablemente pequeño) que hubiese requerido esta investigación con técnicas de la *era analógica*.

La segunda noción definitoria sería la *autenticidad* de los textos. Un hecho primordial para cualquier investigación, especialmente en el caso presente, tiene que ver con el contexto de producción y la selección de los datos. Los datos auténticos generalmente se entienden como datos que ocurren naturalmente, es decir, no se crearon originalmente ni se obtuvieron con el propósito de realizar un análisis lingüístico. En las

discusiones previas a la presentación de este trabajo se baremó la naturaleza de las fuentes de las que se iba a obtener los textos del ECBE. Si bien, inicialmente se planteó utilizar distintas fuentes como textos procedentes de redes sociales o discursos políticos, finalmente se decidió utilizar el discurso de dos cabeceras reconocidas, principalmente por los siguientes criterios: (1) la homogeneidad de su lenguaje, (2) un nivel divulgativo/didáctico en cuanto al uso del lenguaje especializado, de manera que enlaza el discurso del especialista con el gran público y (3) la amplia disponibilidad de los textos en Internet.

Los corpus han de ser representativos (Biber et al., 1998) y aportar un modelo verosímil de las propiedades del lenguaje: un corpus representativo permitirá obtener conclusiones más sólidas. El objetivo marcado era que el corpus de trabajo fuese altamente representativo del lenguaje objeto de estudio de manera que aportase resultados concluyentes. Para que los textos reflejen el lenguaje objetivo, en nuestro caso, la evolución del discurso de la prensa económica durante la crisis, deben seleccionarse muestras relevantes para el estudio. Al tratarse de un corpus especializado es necesario establecer criterios estrictos en la selección de textos. Con vistas a conseguir una colección equilibrada y representativa del estudio se ha considerado la selección de la sección *Business* de periódicos generalistas como material principal del corpus. En este caso, se ha considerado idónea la versión web de dos de los periódicos generalistas con mayor representatividad en lengua inglesa<sup>46</sup>: The Guardian y The New York Times, de forma que queden cubiertos las dos variedades principales de la lengua inglesa: británico y estadounidense. Tanto los comités editoriales como los autores son expertos representativos del dominio.

Ambas publicaciones publican con la visión de influenciar solo a sus lectores directos, y además tienen un alcance internacional a través del impacto en otras cabeceras<sup>47</sup>. En términos de divulgación o vulgarización, la noticia periodística

---

<sup>46</sup> Según la agencia comScore (2012), 44 millones de personas en todo el mundo accedieron a sitios de periódicos en línea en octubre de 2012, lo que representa un 42,6% de la población total de Internet. El New York Times ocupó el segundo lugar con 48,7 millones de visitantes únicos en todo el mundo y The Guardian ocupó el tercer lugar con una audiencia de 38.9 millones.

<sup>47</sup> Tanto el New York Times, como The Guardian son citados por la prensa internacional, y son traducidos a otros idiomas. Un ejemplo de ello es The New York Times *en Español* lanzado en 2016 o el vigente convenio entre la cabecera digital española *Eldiario.es* y The Guardian para traducir al español sus artículos más influyentes.

económica traduce al público general la información codificada a través del lenguaje especializado de informes económicos y boletines de prensa de los gobiernos, agencias públicas y principales empresas privadas del mundo. Es indiscutible que la crisis financiera ha impactado de lleno y cambiado el estilo vida de los ciudadanos. Por esto, es necesario destacar la oportunidad que ofrece estudiar la prensa en una época extraordinaria como esta, donde se produjo un aluvión de informaciones y, además, como se puede observar en el apartado de descripción del corpus, multiplicó las líneas dedicadas a la economía en la prensa generalista. La construcción del ECBE permite obtener una panorámica sobre cómo los textos periodísticos llevan a cabo la construcción lingüística de la realidad social. Un argumento que sostiene la autenticidad y representatividad de los textos, de acuerdo con Samson (2001) y trabajos recientes como la tesis de Gadea Boronat (2017), es la proposición de utilizar corpus de este tipo. pues, por un lado, utiliza un lenguaje sobradamente representativo y, además, ofrecen gran cantidad de ocurrencias de los términos objeto de esta investigación contextualizados en una muestra de lenguaje en uso realista. Esto permite asegurar que existen evidencias suficientes para generar o probar hipótesis lingüísticas.

La selección de textos de carácter auténticos trasciende el estudio del sentimiento del lenguaje especializado en la prensa como polea de transmisión del discurso económico, ya que las noticias también sirven como ventanas de interpretación de la opinión pública. Entre los textos no solo se encuentran resúmenes de las ruedas de prensa<sup>48</sup> e informes económicos, sino que se incluyen sus interpretaciones: columnas de opinión, entrevistas o coberturas en vivo de distintos eventos de interés, así como comentarios periodísticos de las reacciones del público en las redes sociales.

El último motivo es eminentemente técnico: las webs de ambas cabeceras están disponibles de manera abierta y gratuita en Internet<sup>49</sup> a través de sendas ediciones digitales. Esta circunstancia favoreció el trabajo de recopilación de los textos. El libre acceso facilitó la descarga mediante distintas técnicas computacionales de extracción de información de manera que los textos quedasen codificados y catalogados con sencillez.

---

<sup>48</sup> Por ejemplo, las ruedas de prensa mensuales de la presidencia del Banco Central Europeo, que también se analizan en Fernández-Cruz y Ortega Palomo (2019, en prensa).

<sup>49</sup> Otras opciones iniciales, también orientadas al gran público, como *Financial Times* o *The Washington Post* son medios de pago y no tienen acceso libre al público.

Dada la facilidad para obtener los datos, se decidió extraer la sección *Business* al completo en ambas cabeceras, de manera que se pudiese trabajar con la población completa.

El cuarto factor son los *criterios específicos*. Aquí, además de hacer uso de lenguaje auténtico dentro del discurso de la prensa económica, el otro criterio estructural necesario para cumplir los objetivos es el de longitudinalidad. Se toma como punto de partida el mes de enero de 2007, marcado por un momento previo a la coyuntura de inseguridad surgido a partir de la crisis de las hipotecas *subprime* y el posterior colapso del gigante bancario Lehman Brothers en septiembre de 2008. A partir de este momento, se encadenaron una serie de factores económicos, políticos, y sociológicos cuyas repercusiones sobre la población siguen vigentes una década después. Así, se pueden extraer muestras textuales por fecha de publicación y realizar estudios descriptivos a través de series temporales. El estudio cuantitativo y cualitativo del lenguaje del corpus textual permitirá estudiar la evolución de unidades clave seleccionadas a partir del discurso periodístico económico.

Todos los textos del corpus fueron anotados sintácticamente de manera automática a través de los conjuntos de etiquetas *Penn Treebank*<sup>50</sup> (Marcus, Santorini y Marcinkiewicz, 1993), utilizadas por *Lingmotif*. Esta anotación es un requisito necesario para el siguiente paso: la extracción automatizada de candidatos a términos. Este asigna a cada palabra etiquetas referente a los tipos de palabra (etiquetas POS para abreviar), es decir, etiquetas utilizadas para indicar la parte de voz y, en ocasiones, también otras categorías gramaticales (caso, tiempo, etc.) de cada *token* en un corpus de texto. Una etiqueta POS representa un conjunto único de características morfosintácticas y una palabra puede tener varias etiquetas POS. Por ejemplo, la cadena “books” generalmente puede tener dos lecturas: en el sintagma “he calls the hotel” la palabra “calls” es un verbo de la tercera persona (usaría la etiqueta VBZ) pero en el sintagma “there have been three missed calls” sería un sustantivo plural (etiquetado como NNS).

Finalmente, aunque no es una parte esencial para el cumplimiento de los objetivos de esta investigación, desde el primer momento se ha tenido en cuenta el reaprovechamiento del corpus para futuras investigaciones. Por ello, se ha priorizado el

---

<sup>50</sup> En este caso el conjunto de etiquetas English Modified Penn Treebank: <https://www.sketchengine.eu/modified-penn-treebank-tagset/>

uso de formatos universales<sup>51</sup> y estándares de codificación abiertos según la filosofía *OpenData*<sup>52</sup>.

### 6.3.1 Recopilación del corpus

Como se ha visto en la sección anterior, los textos fueron extraídos a través técnicas computacionales. A pesar de la existencia de gran cantidad de *scrapers* y otros extractores de información, no se encontraron herramientas que satisficiesen enteramente las necesidades de extracción de la información. Por ello, se diseñó el proceso siguiente a cabo a través de cuatro pasos:

1. Extracción de las URL de cada noticia mediante técnicas mixtas.
2. Descarga organizada de los ficheros HTML.
3. Extracción y limpieza de los textos.
4. Clasificación y etiquetado y empaquetado de los ficheros del corpus.

En el primer paso se extrajeron todas las URL públicas de la sección *Business* de ambos periódicos digitales. Para ello, a través la herramienta de *búsqueda avanzada* del buscador *Google*, con la ayuda de un extractor de URL, se extrajeron todas las direcciones que se incluyesen el patrón “/business/” dentro de dicha dirección, conforme con las especificaciones organizativas de ambas webs. Se realizó, entonces, una búsqueda avanzada por fecha para cada uno de los meses entre 2007 y 2015. Dadas las medidas de seguridad de Google, esto se debe realizar de manera manual. Si bien esto es una labor que requiere un trabajo intensivo, los resultados ofrecidos son muy refinados. Las URL se recopilaban en ficheros de texto plano por meses y cabeceras.

El segundo paso fue descargar la información al completo. Dado que esta venía codificada en formato HTML, se pudo descargar con relativa facilidad. Se desarrolló un script de *shell* en *Ubuntu* que incluía los comandos UNIX necesarios para realizar esta labor de manera que descargase los ficheros en directorios distintos organizados jerárquicamente por año, mes y cabecera. Tras una serie de medidas de control que permitió eliminar ficheros irrelevantes por su tamaño o repetidos.

---

<sup>51</sup> Por ejemplo, el uso de codificación universal UTF-8 en los archivos de texto.

<sup>52</sup> Datos abiertos según las características de *disponibilidad y acceso, reutilización y redistribución y participación universal*, según la Fundación Open Knowledge International. Más información en: <https://okfn.org/opendata/>

Después, se debió separar el texto plano del resto de la información irrelevante para este trabajo, principalmente: código estructural HTML, hojas de estilo CSS y scripts en *Java*. Así, se obtuvo el texto plano que necesitábamos a partir de sus respectivas etiquetas HTML: titulares (<h1>), subtítulos (<h2>) y cuerpo de texto (<p>). Si bien los documentos resultantes presentaban generalmente buenos niveles de limpieza; tras su revisión manual, fue necesario eliminar patrones textuales repetitivos asociados, principalmente, con los menús y otros elementos estructurales de las páginas. Para realizar esta labor de limpieza de elementos redundantes, se debió ejecutar un sencillo proceso de limpieza mediante expresiones regulares.

### 6.3.2 Descripción del corpus

La extracción del ECBE conforma ya el primer resultado del trabajo de recopilación de textos periodísticos de carácter longitudinal de manera que, con una sensibilidad temporal mensual, se abarca el periodo histórico 2007-2015. La totalidad de los textos son artículos periodísticos pertenecientes a la sección *Business* en lengua inglesa de los periódicos *The Guardian* y *New York Times*, cuya tipología textual corresponde a un corpus de carácter especializado. Al finalizar la recopilación de textos procedentes de ambos medios de comunicación, se han contabilizado 24.871.005 palabras y 21.267.386 tokens pertenecientes a 42.193 textos periodísticos, 21.312 correspondientes a *The Guardian* y 20.881 al *New York Times*. En cuanto a las características del léxico, el corpus incluye 241.831 palabras diferentes, de las cuales, 187.510 son lemas (paradigmas morfológicos) distintos agrupadas en 942.703 oraciones. La Tabla 10 presenta un resumen descriptivo del corpus.

<i>Año</i>	<i>Tokens</i>	<i>Palabras</i>	<i>Textos</i>	<i>%</i>
2007	1.720.623	1.471.317	2.921	6,91
2008	1.856.425	1.587.443	3.319	7,46
2009	1.991.655	1.703.079	3.514	8,00
2010	2.246.883	1.921.326	3.782	9,03
2011	2.706.020	2.313.938	4.409	10,88
2012	2.489.847	2.129.087	4.718	10,01
2013	3.595.273	3.074.345	5.585	14,45
2014	3.923.744	3.355.223	6.673	15,77
2015	4.340.535	3.711.624	7.408	17,45

**Tabla 10** Datos del ECBE, desglose por año de los datos sobre el número total de palabras y de artículos recopilados

La Tabla 11 resume los datos finales resultantes de la extracción de artículos entre 2007 y 2015 en los que hemos subdividido la investigación. Si se observan los datos



anuales del ECBE, se puede ver claramente el despegue de las informaciones publicadas en línea en ambos medios, según avanzan los años. Concretamente, es a partir de 2012 cuando las publicaciones de la sección *Business* de ambos periódicos aumentan cuatro puntos con respecto al año anterior. Esto viene motivado, principalmente, por una progresiva mejora en la indexación de los contenidos en línea, sumado a una mayor implantación de los medios web de manera que a lo largo de la crisis se fueron generando contenidos exclusivos para la web y un número creciente de usuarios. Asimismo, la situación histórica en términos económicos requería de una mayor cobertura con, por ejemplo, reportajes especiales y análisis al minuto de eventos clave.

<i>Fuente</i>	<i>Tokens</i>	<i>Palabras</i>	<i>Textos</i>	<i>%</i>
The Guardian	13.197.301	11.285.112	21.312	53,06
The New York Times	11.673.704	9.982.273	20.881	46,93

**Tabla 11** Descripción del corpus ECBE según fuente

Para una búsqueda eficiente que permita observar longitudinalmente el contexto de los términos clave era necesario realizar una buena y sencilla catalogación de los textos. Para ello, cada uno de los archivos del ECBE fue denominado de manera estándar de forma que incluyesen los siguientes metadatos codificados:

- La fecha de publicación de los textos. Van por orden desde enero de 2007 hasta diciembre de 2015. La fecha se codificó en cuatro cifras, siendo las dos primeras relativas al año y las dos últimas relativas al mes de publicación. Por ejemplo, “07” para el año 2007 y “15” para 2015 en el caso del año, y “01” para enero y “12” para diciembre. De esta forma, la fecha de un fichero publicado en agosto de 2013 se codificaría de esta forma: “1308”.
- El nombre del periódico que hemos analizado. En este caso son dos The Guardian (“GU”) y The New York Times (“NYT”).
- Un código ID numérico para identificar cada artículo de manera que cada uno de los ficheros de texto recibió un código único de identificación. Este sirvió como nombre de archivo.

Un ejemplo de nombre de archivo sería el siguiente: “1303NYT103.txt”, que correspondería a un artículo publicado en el año 2013, mes de marzo, bajo la cabecera del New York Times y el 103 como un número de identificación.

Finalmente, se seleccionó la suite online *The Sketch Engine* para su gestión, procesamiento y divulgación. Esta suite ofrece la utilización de 500 corpus en más de 90 idiomas y múltiples especializaciones, algunos de ellos con tamaños que sobrepasan los



30 mil millones de palabras, de manera que se dispone una amplia representatividad de las distintas variedades de la lengua.

## 6.4 Herramientas

### 6.4.1 *Lingmotif*, un software de análisis de sentimiento basado en léxico

#### 6.4.1.1 Descripción del software

*Lingmotif* es un sistema de análisis de sentimiento basado en léxico en español e inglés desarrollado por el grupo Tecnolengua en la Universidad de Málaga cuya conceptualización original se remonta a 2010. La investigación en este campo, inicialmente dedicado al análisis de sentimiento en español, se inició con el proyecto *Sentitext* (Moreno-Ortiz et al., 2010) y tras el lanzamiento de *Lingmotif* se añadió un amplio lexicón en inglés diseñado ad hoc junto con un conjunto propio de reglas contextuales. El sistema ha sido publicado en varias presentaciones, una versión multiplataforma de escritorio para *Windows*, *MacOS* y *GNU-Linux* desarrollada en *Python*.

El funcionamiento de *Lingmotif* depende de un conjunto muy amplio de fuentes léxicas y una base de datos de reglas de análisis de contexto que son utilizadas en conjunto para detectar segmentos textuales que incluyan sentimiento con el objetivo de generar dos puntuaciones en una escala de 0 a 100. La primera (*sentiment score*) cualifica un texto según su sentimiento o polaridad, mientras, la segunda lo cualifica aporta una medida de la intensidad del texto, siendo los textos con puntuaciones más bajas consideradas de carácter factual y altamente evaluativas las puntuaciones del otro extremo (Moreno-Ortiz, 2016). En resumen, *Lingmotif* separa el texto objetivo en sus oraciones constituyentes y posteriormente detecta y asigna una valencia a las palabras y expresiones multipalabra que incluyen sentimiento según sus fuentes de datos, teniendo en cuenta los posibles modificadores contextuales. Para cada idioma el sistema *Lingmotif* contiene un léxico central, un conjunto de reglas contextuales y, para textos de dominio específico, un lexicón complementario (*plug-in*), un etiquetador de categorías gramaticales y un lematizador (Moreno-Ortiz, 2017).

*Lingmotif* analiza los textos a nivel oración. Si bien existen muchos indicadores evaluativos que operan de manera supraoracional, los análisis a este nivel son muy eficaces a la hora de clasificar documentos cortos, p.e reseñas de usuarios en webs

especializadas o *microblogs*. Además, existen implementaciones multinivel de *Lingmotif* (Moreno-Ortiz, 2016) que analizan el sentimiento de manera pormenorizada, de manera que es posible observar tanto los segmentos evaluativos como obtener una puntuación general de sentimiento para un texto que incluya múltiples oraciones (Moreno-Ortiz, 2017).

Como se revisó en el Capítulo 5, la detección de elementos léxicos con una orientación semántica determinada es insuficiente sin el estudio de los factores contextuales ya que puede ser intensificado, atenuado, o invertido. La primera implementación de las reglas contextuales fue la propuesta de *modificadores contextuales de valencia* (*contextual valence shifters*) de Polanyi y Zaenen (2006) que propusieron un sistema que englobase la modificación de la orientación semántica de una unidad léxica determinada por medio del contexto que la rodea.

Los textos que han de ser analizados deben proporcionarse a *Lingmotif* recopilados en un fichero CSV o como una lista de archivos de texto. En el primer caso, un texto de entrada requiere al menos dos columnas separadas por tabulaciones *id* y *text* (identificador y texto de entrada). En este trabajo, además, se han utilizado dos campos más *file* y *date*, para identificar el nombre del fichero analizado y la fecha de publicación. Un ejemplo utilizado en este trabajo sería el siguiente:

<i>id</i>	<i>file</i>	<i>text</i>	<i>date</i>
1	0701NYT100	In late December, Standard & Poor's lowered its rating for the Times Company's long-term corporate debt from A- to BBB+.	2007-01-28
2	0701NYT126	In one case, a married couple embezzled more than \$50 million from the state-run Bank of China to pay off gambling debts in Macao.	2007-01-28

**Tabla 12** Ejemplo de las dos primeras entradas del dataset de entrada del término “debt”

Los resultados de *Lingmotif* se obtienen mediante un proceso de detección de las palabras y expresiones multipalabra que contengan sentimiento, para posteriormente analizar sus contextos a través de reglas contextuales y, finalmente, ponderar el sentimiento a través de una fórmula que también tiene en cuenta las unidades que no contienen sentimiento. El proceso de análisis es el siguiente:

1. Preprocesado: el texto objetivo se escanea en busca de abreviaturas comunes, contracciones y errores ortográficos.
2. *Tokenización* a nivel oración y a nivel palabra.

3. Identificación de expresiones multipalabra: los n-gramas se comparan con la lista de palabras contenidas primero en el lexicón *plug-in* de dominio específico (si se da el caso) y, posteriormente, en el lexicón principal. Una vez se detectan las expresiones multipalabra, se les asigna la valencia a la que se asocia en el lexicón.
4. Identificación de palabras individuales: estas se buscan primero en el lexicón *plug-in* y en el lexicón principal. Si se encuentran, se les asignará la valencia.
5. Coincidencia de reglas contextuales: las palabras de polaridad y expresiones multipalabra identificadas se comparan con la lista de reglas de contexto. Si el contexto de la palabra coincide con la estructura incluida en una regla, se marca todo el segmento de texto y, posteriormente, se etiqueta como una unidad de tipo regla contextual. Este proceso se lleva a cabo en varios barridos para detectar las secuencias de modificación contextual acumuladas, como “*very very good*”.
6. Cálculo de puntuación de sentimiento y puntuación de intensidad. Para una oración determinada, debe tenerse en cuenta el número de unidades positivas (*pos*), negativas (*neg*) y sin sentimiento (*none*). En este proyecto el cálculo de puntuaciones de sentimiento se llevó a cabo de la siguiente forma (Figuras 10 y 11). Si el número de unidades positivas es superior a las negativas:

$$S = \frac{1 - \sum neg}{\sum pos + \sum neg + 0,15 * \sum none}$$

**Figura 10** Fórmula de cálculo del sentimiento de Lingmotif cuando las unidades positivas superan a las negativas

Si, por el contrario, el número de unidades negativas supera a las unidades positivas, se aplicará la siguiente fórmula:

$$S = \frac{\sum pos}{\sum pos + \sum neg + 0,15 * \sum none}$$

**Figura 11** Fórmula de cálculo del sentimiento de Lingmotif cuando las unidades negativas superan a las positivas

Tras el análisis de los textos, el sistema generará un conjunto de datos en formato CSV con la información relativa a los resultados. Esta información está agrupada principalmente en métricas relativas al sentimiento e información estadística de carácter textométrico:

- a) Métricas relativas al sentimiento:
  - a. *polarity*: polaridad (*pos*, *neg*, *neu*, *none*).

- b. *tss*: *Total Sentiment Score*, puntuación total de sentimiento en una escala de 0 a 100. Siendo 0 el polo negativo, y 100 el polo máximo positivo.
  - c. *tsi*: *Total Sentiment Intensity*, puntuación total de la intensidad de en una escala de 0 a 100. Siendo 0 una oración sin marcadores evaluativos y 100 una oración con gran cantidad de marcadores evaluativos de cualquier tipo.
  - d. *pos\_score*, *neg\_score*, *neu\_score*: respectivamente la puntuación positiva, negativa y neutra.
  - e. *pos\_items*, *neg\_items*, *neu\_items*, *none\_items*: respectivamente, recuento numérico de elementos positivos, negativos, neutros o sin polaridad.
- b) Estadísticos textométricos:
- a. *sentences*: número de oraciones.
  - b. *tokens*: número de unidades léxicas entre espacios.
  - c. *tt\_ratio*: ratio type/token.
  - d. *lex\_words*, *function\_words*: cantidad de palabras léxicas y palabras gramaticales.
  - e. *single\_words*, *multi\_words*: cantidad de palabras y expresiones multipalabra.
  - f. *Redups*, *all\_caps\_words*: respectivamente, palabras reduplicadas y palabras que están en mayúsculas.
  - g. *p\_nouns*, *c\_nouns*, *verbs*, *adjectives*: cantidad de nombres propios y sustantivos comunes, verbos y adjetivos.
  - h. *q\_marks*, *x\_marks*: número de signos de interrogación y exclamación.
  - i. *Handles*: el número de *handles* (o nombres de usuario únicos) que aparece en el texto. Estos comienzan siempre con “@”, por ejemplo @tecnolengua.
  - j. *urls*: cantidad de direcciones URL.
  - k. *emojis*: número de emojis.

#### 6.4.1.2 El lexicón nuclear de sentimiento

El lexicón central es el recurso más importante del sistema *Lingmotif* y su rendimiento es directamente proporcional a la calidad y detalle de sus recursos léxicos. Por ello, el objeto de su equipo de desarrollo ha sido un lexicón base robusto y eficiente. Todos los

elementos incluidos en el léxico de *Lingmotif* (ya sea una palabra o expresión multipalabra) deben especificar cuatro elementos: (1) una entrada, (2) su categoría gramatical, (3) una etiqueta de polaridad y (4) la intensidad de esta.

Una entrada puede ser una cadena (p. ej., *clientelism*) o un lema representado entre paréntesis angulares en representación de todas las inflexiones de un paradigma morfológico completo (p. ej., <bad> englobará a tres formas: *bad*, *worse*, *worst*). En cuanto a la especificación de la categoría gramatical, *Lingmotif* también utiliza las etiquetas *Penn Treebank* expuestas anteriormente, junto con un comodín (ALL) para los casos en que todas las categorías gramaticales posibles para ese lema compartan la misma valencia. La polaridad hará referencia a tres posibilidades: *pos* para positivo, *neg* para negativo y *neu* para neutro. El último valor, la intensidad, será una puntuación en un intervalo del 0 al 3. Como se puede ver en la Figura 12 los datos son recogidos en un *dataset* estructurado con los siguientes datos: entrada, categoría gramatical, polaridad e intensidad.

<bad>                      JJ    neg    2

**Figura 12** Una entrada típica según las especificaciones de *Lingmotif*

Los lexicones de *Lingmotif* son el producto de un intenso trabajo lexicográfico multidisciplinar. El proceso de creación del lexicón central de *Lingmotif* se basó en la fusión de otros lexicones previos disponibles y la posterior adaptación de la lista fusionada al formato del sistema y su posterior refinado mediante técnicas de análisis de corpus y heurísticas integrales. Para la lengua inglesa, fusionaron los principales lexicones generalistas revisados en el Capítulo 5: el *General Inquirer* (Stone y Hunt, 1963), MPQA (Wilson, Wiebe y Hoffmann, 2005) y el lexicón *Bing Liu* (Hu y Liu, 2004). Como complemento, estos recursos se expandieron utilizando un diccionario de sinónimos mediante reglas de generación derivada. Todos estos recursos se caracterizan por no incluir expresiones multipalabra por lo que se etiquetaron varias fuentes de datos p. ej., los modismos comunes de *Wiktionary* (Moreno-Ortiz, 2017), que posteriormente se expandieron mediante el uso de corpus.

En la actualidad, el lexicón nuclear en lengua inglesa contiene más de 68.000 entradas y aproximadamente 760 reglas contextuales, mientras en español contiene 294.000 entradas y más de 770 reglas de contexto.

La desambiguación entre formas textuales que contienen unidades idénticas se establece a través de las características formales de las palabras, esto es, mediante la detección automática de las partes de la oración y el uso intensivo de expresiones

multipalabras que contiene, principalmente, giros idiomáticos y otras construcciones gramaticales. Estas unidades incluirán palabras que pueden tener una polaridad distinta a la expresión monoléxica y así resolver en muchos casos la desambiguación. Se ilustrará el primer caso de *fit* (p. ej., “The regression program FIT a line to the data.”) como verbo de carácter neutro y *fit* como adjetivo, con una polaridad positiva (“I am FIT enough to climb the Tungurahua”). Para el segundo caso se propone aquí la palabra *blue*, la cual se clasifica como neutra; por otro lado, colocaciones como “blue chip”<sup>53</sup> tienen una polaridad positiva. En el caso de que la desambiguación no pueda llevarse a cabo debido a factores como la polisemia o usos muy contextualizados, las opciones disponibles serían añadir la entrada etiquetando la polaridad estadísticamente más probable. En otros casos, se propone excluir la entrada del lexicon cuando existan posibilidades similares de que el elemento tenga las dos polaridades (Moreno-Ortiz, 2017).

#### 6.4.1.3 Reglas contextuales

Las reglas contextuales (*context rules*) son el conjunto de patrones gramaticales que utiliza *Lingmotif* como mecanismo para gestionar variaciones en el sentimiento de las entradas encontradas en el lexicon. Básicamente, se utiliza una implementación del enfoque propuesto a través de los denominados *Contextual Valence Shifters* de Polanyi y Zaenen, (2006), que funcionan detectando, además, otras palabras o expresiones en las inmediaciones de la palabra identificada. Otros sistemas que han implementado este enfoque han sido Kennedy e Inkpen, (2006), Taboada et al., (2011), y Moreno-Ortiz et al., (2010; 2017).

Dada una combinación de palabra de sentimiento y regla contextual, el sistema considerará el segmento de texto resultante como una sola unidad y posteriormente modificará los valores de sentimiento. Según las especificaciones de la regla contextual, se recalculará la puntuación de sentimiento mediante sumas o restas de puntuación y se le asignará una nueva valencia al conjunto. El juego de reglas contextuales que utiliza el sistema de *Lingmotif* se recopiló mediante análisis de corpus extensivo. En él, se observaron las concordancias de palabras de polaridad comunes de cualquier categoría

---

<sup>53</sup> “Blue chip” se utiliza para referirse a empresas bien establecidas, que tengan ingresos estables, valores sin grandes fluctuaciones y que no precisen de grandes ampliaciones de su pasivo.

gramatical en combinación con distintas reglas textuales con vistas de obtener un conjunto bien refinado de reglas (Moreno-Ortiz, 2017).

El sistema permite detectar reglas contextuales con alto nivel de detalle. Según Moreno-Ortiz et al. (2013), el sistema de reglas contextuales se implementó siguiendo la siguiente estructura de datos:

- (a) Formato de la unidad: definición morfosintáctica, según los estándares de *Freeling*, de los elementos modificados (p. ej.: "AQ" para adjetivos calificativos).
- (b) Signo de la unidad: polaridad del elemento modificado (p. ej., "+" o "-").
- (c) Definición de regla contextual: define el punto de partida del modificador (p. ej.: "very").
- (d) Posición de la regla contextual: posición del modificador (p. ej., "L" para realizar una búsqueda a la izquierda y "R" para realizarlo a la derecha).
- (e) Distancia a la regla contextual: distancia, es decir, número máximo de palabras posible entre el modificador y la palabra con polaridad.
- (f) Resultado: valencia resultante de la modificación, este resultado puede expresarse como un operador o como una valencia final ya establecida. El argumento  $n$  que se observa a continuación en los últimos dos operadores es el grado numérico que se va a aplicar o una valencia determinada. Los tres operadores son los siguientes:
  - INV (inversión de polaridad). Este tipo de regla contextual hará que la polaridad del adjetivo negativo se invierta, por ejemplo:
    - (4) `<be> not at all + Adj neg:` "He's not at all dumb/stupid/...".
  - INT $n$  (intensificación de  $n$ ). En este caso, las construcciones intensificarán la valencia del adjetivo negativo. Por ejemplo:
    - (5) `<be> a total + adj neg:` "He's a total loser"
  - DOW $n$  (atenuación de  $n$ ). Las construcciones de este tipo atenuarán la valencia del adjetivo negativo.
    - (6) `<be> a bit + adj neg:` "He's a bit stupid"

La implementación de este tipo de reglas aporta mayor flexibilidad que si se estableciese un repositorio extremadamente detallado de expresiones multipalabra. Sin reglas contextuales, sería muy difícil de llevar a cabo el procesamiento de la opinión de los textos que contengan combinaciones de este tipo, ya que requeriría explicitar redundantemente todas y cada una de estas combinaciones y repercutiría negativamente



en el rendimiento del sistema. El fenómeno lingüístico que trasciende aquí es la *prosodia semántica* (p. ej., Hoey, 2005) que en determinados patrones léxicos provoca una polaridad específica (Moreno-Ortiz, et al., 2013).

#### 6.4.1.4 Lexicones *plug-in*

*Lingmotif* permite utilizar el lexicón general en combinación con un léxico añadido por el usuario que permita cubrir las necesidades específicas al dominio. Como se ha revisado anteriormente, varios autores han tratado la dependencia del dominio específico en el análisis de sentimiento (Aue et al., 2005; Pang et al., 2008; Turney 2002). Dependiendo de la entidad tratada, atributos propios de estas, como el tamaño, el peso o la ubicación, por ejemplo, pueden ser considerados positivos o negativos o neutros. El sentimiento asignado a un atributo dado, como el *tamaño*, dependerá del tipo de entidad tratado p. ej., un *smartphone*, una *pizza* o un *patinete eléctrico*.

Para muchos casos, el lexicón general de *Lingmotif* es lo suficientemente amplio y flexible como para adaptarse a los distintos dominios específicos sin la necesidad de diseñar lexicones especializados. Sin embargo, dado el caso de existir ambigüedades relacionadas con el dominio, el lexicón general aportará el sentimiento más frecuente entre dominios. Al utilizar un lexicón complementario, la información léxica contenida en el mismo anulará las valencias asignadas inicialmente por el lexicón central de *Lingmotif* o parta otras para palabras inexistentes en él. De esta manera, se realizará primero la búsqueda entre las entradas del *plug-in* y los elementos a los que se les otorgue una valencia no se buscarán posteriormente entre las entradas del léxico principal. Dado el caso de que en el léxico específico una palabra determinada no tenga polaridad entre las entradas del lexicón *plug-in* (p. ej., como el caso de *debt* en el dominio financiero) se pueden añadir valores que anulen la valencia que asigna el léxico principal.

En cuanto al formato, los *plug-in* siguen exactamente las mismas especificaciones que las incluidas en el léxico central. Para ello, al usuario le bastará con crear un fichero de texto plano CSV con codificación universal UTF-8 que recoja la información del dominio específico separada por tabulaciones (entrada, categoría gramatical, polaridad e intensidad) con saltos de línea UNIX. Para obtener el léxico específico el usuario deberá, a través de distintas técnicas terminográficas, extraer los elementos léxicos cuyo sentimiento varíe en el dominio y asignarle valores de sentimiento según sus propios criterios. Como resultado, el rendimiento aumentará con el nivel de detalle. Así, se



dispone de un mecanismo muy flexible que le permite importar léxicos complementarios de sentimiento totalmente personalizables según los requisitos textuales del dominio. Un ejemplo de una entrada de léxico *plug-in* sería el mostrado en la Figura 13:

ITEM	POS	SENTIMENT	INTENSITY
<reduce>_3_risk_premium	VB	pos	2

**Figura 13** Ejemplo de una entrada del léxico *plug-in*

#### 6.4.2 *The SketchEngine*, una suite de gestión de corpus y análisis textual

Sketch Engine es un software *online* de análisis de texto y una suite de gestión de corpus desarrollado por *Lexical Computing Ltd* lanzada por primera vez en 2003. Su propósito es permitir que las personas que estudian el comportamiento del lenguaje (lexicógrafos, investigadores en lingüística de corpus, traductores o aprendices de idiomas) busquen grandes colecciones de texto de forma compleja y lingüística. La diferencia con otras herramientas de análisis de corpus está no solo en su rápido sistema *Manatee* en la nube, y en su cómoda interfaz web, sino en una oferta de corpus de todo tipo directamente disponibles en más de 90 idiomas. Incorpora un software de diseño, mantenimiento y ampliación de corpus. Combina *WebBootCat* (Baroni et al., 2006) que permite descargar y compilar corpus a través de la web y, posteriormente, *lematiza* y *tokeniza* los textos descargados.

La interfaz incluye una amplia oferta de análisis de corpus: *Concordancer*, *WordSketch*, *KeyWords* o *Palabras Clave*, *Tesaurus*, *Listas de palabras*, *N-gramas*, *Tendencias* (para búsquedas diacrónicas) y un comparador de corpus paralelo. Principalmente se utilizaron las tres primeras herramientas. Las posibilidades que ofrece esta suite son idóneas para el tratamiento del corpus ECBE y pertinentes para el cumplimiento del objetivo principal: (1) se utilizará *KeyWords* herramienta de extracción terminológica para obtener la base del léxico de sentimiento y (2) se utilizará el *concordancer* para recopilar las colocaciones del corpus y observar sus ejemplos.

## 6.5 La construcción de *LexiEcon*, un lexicón de dominio económico financiero para *Lingmotif*

### 6.5.1 Extracción terminológica

La extracción de un lexicón *plug-in* completo que abarque las particularidades del sentimiento de un dominio específico con vocación de ofrecer altos valores de precisión y exhaustividad requiere una labor terminográfica extensiva que combine técnicas de distinta naturaleza. Para este fin, se ha contado con una selección textual acorde con los objetivos (esto es, realizada a partir de los mismos textos que van a ser analizados). Para obtener una lista extensiva de candidatos a término se utilizaron distintas utilidades de *The Sketch Engine* principalmente el extractor de terminología y el analizador de concordancias. Para facilitar la clasificación de los candidatos a término se contó, además, con la perspectiva cualitativa de un especialista en el dominio. De esta manera, se pudo realizar un estudio extensivo de las producciones lingüísticas de carácter económico-financiero que tengan relación con la expresión del sentimiento.

Las limitaciones de *Lingmotif* para textos de dominio específico se evidenciaron y discutieron en trabajos anteriores (Moreno-Ortiz et al., 2015) donde se definieron por primera vez los requisitos para diseñar un léxico de dominio económico / financiero. Por aquel entonces ya se había propuesto un método simple para extraer elementos léxicos de dominio específico a partir de corpus. Para llevar a cabo esta tarea, se diseñó un método simple de 3 pasos para extraer los términos relacionados con el sentimiento del dominio específico del sentimiento, de la siguiente manera:

1. Extraer los términos candidatos automáticamente.
2. Hacer coincidir los candidatos a término con la base de datos de polaridad de lenguaje general en el lexicón *Lingmotif-Lex* (Moreno-Ortiz et al., 2018).
3. Refinar los resultados a través del trabajo de corpus para su posterior asignación de valencias de acuerdo con el trabajo conjunto con un especialista en el dominio.

El motivo de utilización de un corpus como base textual del trabajo terminográfico en detrimento de otras herramientas o recopilaciones de terminología, se justifica de acuerdo con una amplia literatura en terminología (p. ej., Cabré, 1993; Gamper y Stock; 1998; Sager, 1990) y análisis de sentimiento (p. ej., Hatzivassiloglou y McKeown, 1997; Turney, 2002; Taboada et al., 2011), pues la identificación y recopilación terminología debe realizarse no solo mediante recopilaciones lexicográficas tradicionales, sino a partir

de textos que incluyen *lenguaje real* pertenecientes al dominio de trabajo. De hecho, Sager (1990) fue uno de los principales defensores del uso de fuentes textuales realistas en la investigación terminológica: “[s]erious terminology compilation is now firmly corpus-based” (p. 154). Tomando esta afirmación como premisa, este trabajo no puede ser eminentemente onomasiológico. De acuerdo con Ten Hacken (2010), mientras que esta perspectiva terminográfica pretende organizar la información, en este trabajo solo interesa la información de sentimiento que estos términos puedan ofrecer por lo que debe trabajarse desde una perspectiva semasiológica. Es decir, solo son útiles los elementos que tengan una denominación preexistente y con un uso diferente al del dominio general. Para su clasificación, los términos se considerarán prototipos idealizados según convenciones aproximadas y escalares y se clasificarán *a posteriori*.

La extracción terminológica, además de ser utilizada ampliamente en campos de estudio como la terminografía o la traducción (Periñán-Pascual, 2015; León-Arauz, et al., 2018), se utiliza también en el campo del análisis textual para realizar tareas de minería de datos y extracción de información a partir de textos no estructurados. Los avances aportados por el desarrollo de las herramientas y técnicas de procesamiento del lenguaje natural han permitido disponer de gran cantidad de recursos terminológicos de gran riqueza de manera automática o semiautomática de manera que hoy en día son una herramienta indispensable en este tipo de trabajos.

En el caso de este trabajo, la base primordial son los textos de periodismo económico incluidos en el corpus ECBE que posteriormente fue procesado con la herramienta *KeyWords* (Jakubíček y Šmerk, 2016) integrada en la suite *The Sketch Engine*. Esta utilidad detecta candidatos a términos monoléxicos y poliléxicos según relaciones de frecuencia y patrones de gramática local. De esta manera, además de ofrecerse una cantidad exponencialmente mayor de información en una fracción de tiempo que la posibilitada por técnicas tradicionales, se permite llegar a niveles mayores de introspección gracias al acceso de información otrora inaccesible.

Un extractor terminológico eficiente devolverá una lista de términos con alta exhaustividad y escasos niveles de ruido, de manera que requiera escasos esfuerzos de limpieza manual. En este caso, el corpus de referencia fue el *English Web 2015* o *enTenTen15* un corpus enorme de 30 mil millones de palabras incluido en *The Sketch Engine*, que contiene muestras de textos relativos a cualquier temática (ver Jakubíček et al., 2013 para una descripción de su proyecto). Tras comparar las frecuencias relativas, el sistema ofrece una puntuación indicando el nivel de *especialidad* en dos listas: una de

palabras y otra de expresiones multipalabra. El sistema permitió extraer con rapidez 30.000 candidatos a términos (15.000 de cada tipo) y exportarlo a un fichero de texto plano. Fue esencial contar con la asistencia de un especialista<sup>54</sup> en el dominio con el fin de seleccionar los candidatos a término que deben ser incluidos en el lexicón. Esto resultó ser una inversión muy provechosa ya que no solo permitió agilizar el trabajo, sino que permitió aumentar el rigor científico de acuerdo con las particularidades de la disciplina. El experto ayudó a establecer criterios de clasificación y convenciones en el trabajo de etiquetado. De esta forma, se minimizaron las arbitrariedades, se ganó en introspección en los distintos usos de los términos y, especialmente, permitió profundizar cualitativamente en el significado y sus connotaciones.

### 6.5.2 Justificación y fundamentos de *LexiEcon*

Para cumplir con el objetivo de diseñar un lexicón de amplia exhaustividad y eficiencia, el problema aquí será trabajar de manera efectiva con los umbrales de ambigüedad inherente a los dominios especializados. Este asunto se ha discutido ampliamente en los estudios de lenguas de especialidad, como se ha revisado en el Capítulo 4. Determinar el estatus de término de una palabra es una tarea difícil a nivel teórico, ya que la condición de término es un problema gradual y no todos los términos son del mismo tipo ni lo son al mismo nivel. En filosofía, Foucault (1969) reflexiona sobre cómo los lenguajes especializados operan de manera diferente al lenguaje general, principalmente debido a su falta de ambigüedad y polisemia. Del mismo modo, la terminología tradicional, por ejemplo, Wüster o incluso autores contemporáneos como Cabré (1992) se centraron en la tarea de denominar conceptos en un sistema a través de sistemas terminológicos unívocos. Tal designación unívoca en la terminología ha sido, especialmente en las últimas décadas, muy discutida. Un referente del cambio de paradigma es el de Temmerman (2000), que critica este principio porque, en términos prácticos, la univocidad no se cumple completamente ya que la polisemia y la sinonimia no son en absoluto un fenómeno inusual en el uso del lenguaje especializado.

---

<sup>54</sup> Reiterar el agradecimiento por el trabajo incansable y las reflexiones aportadas por Ander Mujica Uranga que cuenta con años de experiencia gestionando productos financieros para distintas firmas internacionales y también como colaborador en iniciativas financieras populares.

Si bien estas discusiones no son primordiales en un enfoque abiertamente pragmático como es el caso del procesamiento del lenguaje natural y, más concretamente, del análisis de sentimiento, es necesario tener en cuenta estos factores a la hora de trabajar con el sentimiento en el dominio especializado. Por lo general esto no es problemático en el caso de que las unidades léxicas del dominio general y las unidades del dominio específico compartan la misma polaridad. Los *plug-in* de dominio específico de *Lingmotif* actúan cuando se da el caso de que el sentimiento en el dominio general y en el dominio específico sea distinto. Cuando esto ocurre, el *plug-in* de dominio específico sustituye la polaridad del lexicón general con la correspondiente al dominio especializado. Esto puede parecer una estrategia demasiado simplista desde una perspectiva terminológica pero, como se verá más adelante, es muy eficiente, ya que solo es necesario tratar los casos en los que la polaridad entre ambos dominios sea diferente.

La necesidad de utilizar léxicos especializados separados se manifiesta en el dominio económico / financiero en palabras como “tight” o “axe”, que en la gran mayoría de los casos son negativas en dicho dominio, mientras que su polaridad en el dominio general depende del contexto: es decir, “An elegant night dress must be tight” vs “tighter household budgets”. Esta dependencia del contexto se maneja con éxito combinando el uso de lexicones de dominio específicos, dominio general y reglas de contexto, derivadas de la observación minuciosa del uso del lenguaje real (Moreno-Ortiz, et al., 2015).

En este dominio, ciertas palabras o expresiones multipalabra mantienen la orientación que muestran en los textos de lenguaje general (p. ej., “crisis”, “collapse”) mientras que otros adquieren una orientación específica (p. ej., “bailout”, “growth”) y otros términos son técnicamente considerados neutros cuando están aislados (p. ej., “sales” o “debt”), pero transmiten una fuerte orientación semántica que se revela cuando se combinan con elementos léxicos direccionales, principalmente verbos y sustantivos. Se han recopilado algunas de las características específicas de este dominio de acuerdo con lo observado en el Capítulo 5 (Liu, 2015; Malo et al., 2014):

- (a) *Indicadores de rendimiento*: en términos de seguimiento o liderazgo (indicadores rezagados / líderes) en el mercado. Estos indicadores reflejan los resultados de la actividad de una empresa o institución: participación de mercado (“shares”), ganancia operativa (“operating profit”), costo operativo (“operating cost”), etc. Por otro lado, los indicadores de liderazgo hablan de posibles retornos futuros de la empresa: p. ej., nuevas tiendas (“new shops”), nuevos clientes (“new customers”), aumento en la capacidad de producción (“increase in production

capacity”), etc. Estos términos se caracterizan por no tener un sentimiento per se, sino que serán positivos o negativos según aumenten o disminuyan. Los indicadores de rendimiento se pueden dividir en dos subcategorías:

- Indicadores potencialmente positivos: índices al alza cuando la economía está en crecimiento (p. ej., “production”, “exports”).
- Indicadores potencialmente negativos: índices que tienen una evolución positiva durante un periodo de declive económico (p. ej., “unemployment”, “risk”).

(b) Direccionalidad de los indicadores de seguimiento y liderazgo. Estos son elementos léxicos, principalmente verbos o sustantivos, que describen la dirección de los eventos, es decir, si hay una tendencia creciente o decreciente. Palabras tales como “rise” y “rocket” o “fall”, y “collapse” marcan tanto la dirección como la intensidad de sus respectivos indicadores de rendimiento. Por lo tanto, la polaridad de los diferentes conceptos se basa en la conveniencia de los eventos (por ejemplo, para “sales” el resultado es positivo cuando aumentan, pero neutro o negativo cuando disminuyen). Algunos ejemplos de direccionalidad positiva y negativa son:

- Verbos: “increase”, “skyrocket”, “decrease”, “plummet”, “collapse”.
- Sustantivos: generalmente, sustantivos deverbales p. ej.: “a rise”, “a drop”, etc.

Esto se implementa fácilmente en el lexicón por medio de un conjunto de sustantivos y verbos (es decir, “drop”, “dip”, “doublé”, “rise”) que se combinan de manera redundante con diferentes indicadores positivos o negativos, en un lapso de tres palabras, como puede verse en la Tabla 13.

<i>Unidad</i>	<i>POS</i>	<i>Sent</i>	<i>Intens</i>
unemployment_3_<fall>	VB	pos	3
unemployment_3_<dip>	VB	pos	2
unemployment_3_<double>	VB	neg	3
unemployment_3_<grow>	VB	neg	2
employment_3_<fall>	VB	neg	3
employment_3_<dip>	VB	neg	2
employment_3_<double>	VB	pos	3
employment_3_<grow>	VB	pos	2

**Tabla 13** Ejemplos de direccionalidad en el indicador “employment”

También es importante mencionar que otras unidades de una sola palabra o multipalabra no forman parte necesariamente del conjunto de indicadores /



direccionalidad, ya que también transmiten un sentimiento específico del dominio. Entre ellos, se incluyen también términos, argot o palabras no especializadas que tienen una polaridad diferente en el dominio económico. Por poner un ejemplo de esto último, “flooding” tiene una connotación positiva en el dominio específico, mientras que frecuentemente tiene una connotación negativa para el dominio general:

- a. The ECB's interventions in the bond markets, its FLOODING the markets with liquidity and cheap credit.
- b. Our houses are FLOODING every two weeks because the city sewer system can't handle the rains.

### 6.5.3 Etiquetado de entradas

El siguiente paso consiste en la selección de las entradas en el lexicón *plug-in* a partir de una amplia lista de candidatos a términos. Para ello, se realizó un preprocesado de los datos proporcionados por el extractor terminológico, seguido de un proceso de cribado y asignación de valencias.

Primero, los 30.000 candidatos a término extraídos fueron convertidos al formato requerido por *Lingmotif*. De esta manera, se descartaron nombres propios y variaciones flexivas redundantes (p. ej., “creditor” y “creditors”), y se añadieron elementos estructurales según los requisitos del clasificador. Principalmente, se marcaron los lemas mediante los caracteres “<” y “>” al inicio y al final de palabra y, posteriormente, se sustituyeron los espacios entre palabras por el carácter “\_”. A continuación, se procesaron las palabras según el estándar *PennTreebank* de manera que cada una de las entradas disponga de una etiqueta de categoría gramatical. En los casos en los que la categoría gramatical no sea relevante, además de etiquetas como “NN” para sustantivos o “VB” para verbos, puede utilizarse la etiqueta universal “ALL”.

El siguiente paso fue comparar la polaridad de las entradas con las valencias presentes en *Lingmotif-Lex*. Es necesario recalcar que no existe interés en la extracción de una lista completa de términos, sino en recopilar una lista sucinta de términos que indiquen positividad o negatividad cuando la orientación difiere de la opinión que tienen en diferentes dominios. En este caso, términos como “sales” o “bank” son irrelevantes, ya que ambos son neutros en los dominios específicos y generales. Sin embargo, los términos neutros de una sola palabra como “asset” o “fix” se deben incluir en el lexicón específico, ya que ambos elementos son positivos en el dominio general mientras que son



neutros en el lenguaje especializado. Este es el caso de la contraposición del primer ejemplo con el segundo:

- a. The currency fixes are based on actual trades.
- b. There is no easy fix: removing the wires can be dangerous...

Los términos con carga de polaridad ya presentes en el lexicón central no tienen que ser contabilizados (p. ej., “recovery” o “crisis”). Por lo tanto, se descartaron todos los candidatos no especializados, los términos neutros y los términos que no transmitían una orientación semántica diferente en el dominio específico. En el caso de los usos complejos del léxico, los términos fueron refinados con la ayuda de un experto en el dominio, de manera que se pudo comprender con mayor profundidad la positividad o negatividad de enunciados y usos del lenguaje muy específicos. Se tuvo como principio el mantener la neutralidad ideológica dentro de lo posible, prefiriendo etiquetar los términos con un sentimiento muy dependiente de la ideología como neutros (p. ej., el caso de “eurosceptic”), dejando que se determine la polaridad a través de otras palabras contextuales. Dado que en economía una misma situación puede ser observada como positiva o negativa según la perspectiva material o ideológica del experto, en los casos en los que se consideró necesario tomar una posición a la hora de asignar valencias, se establecieron los siguientes criterios:

- *Lógica crecientista*: dada su *cuasi* universalidad, se considera que el crecimiento económico es un hecho positivo, en detrimento de otras corrientes como las teorías económicas relacionadas con el *decrecimiento*.
- Se consideran los términos relacionados con lo ecológico como positivos p. ej., “green investment”.
- En el caso de que el sentimiento de un elemento dependa del destinatario, se tomará como positivo el interés general o, en segundo lugar, el sentimiento prevalente de su uso en el corpus, p. ej., “privatization” es considerada como negativo.
- Se consideran las negociaciones y renegociaciones como positivas, ya que permiten resolver problemas, p. ej., “refinancing”.
- Se considera la transparencia como un valor positivo p. ej., “transparency” y el secretismo como algo negativo, p. ej., “tax haven”.
- Las reducciones repentinas de presupuesto (p. ej., “cuts”) son consideradas negativas.



Con la asistencia de un experto en el dominio, estos resultados fueron refinados mediante la comparación de candidatos con los textos del corpus a los que, de manera manual, recibieron puntuaciones de sentimiento. El flujo de trabajo para este paso fue el siguiente:

- a. Selección de término y asignación de puntuación de sentimiento inicial a palabras de polaridad claramente identificadas. Se verificó la especulación previa con los ejemplos encontrados en el corpus, mientras que se buscaron usos alternativos y patrones de gramática locales. Además, todos los fallos y omisiones obvias (como la ausencia de colocaciones) se añadieron al listado de manera manual. Por ejemplo, un término aparentemente no candidato que se encontró que, sin embargo, fue bastante frecuente durante los años de la crisis europea fue “haircut”:

The Troika want HAIRCUTS of over 50% for wealthy savers (€100k+) at Bank of Cyprus.

- b. Asignación de etiquetas finales de sentimiento e intensidad para palabras y expresiones multipalabra específicas según el formato de *Lingmotif*.
- c. En el dominio de la economía y las finanzas fue, además, necesario realizar una recopilación relacional de elementos direccionales e indicadores económicos. Para ello, se recopiló una lista coherente de palabras relacionadas con la direccionalidad (entre ellas, *rise*, *fall*, *decline*, *duplicate*, etc.) y una lista de indicadores económicos potencialmente positivos / negativos a los que se asignó un sentimiento de acuerdo con su naturaleza. Por ejemplo, *rise in unemployment* sería positivo, mientras que a *fall in unemployment* se le asignaría un valor negativo. Aprovechamos que la sintaxis del léxico en *Lingmotif* permite establecer rangos de distancia en las unidades multipalabra de manera que: `<fall>_3_unemployment` permite una ventana de tres palabras entre las cadenas *fall* y *unemployment* de manera que quedan cubiertas expresiones como “fall in unemployment” o “fall in Turkmenistan’s youth unemployment”.

En aislamiento, los indicadores están marcados como neutros con intensidad. Como un elemento potencialmente negativo (p. ej., *unemployment*, se marcará como negativo si va acompañado de un elemento léxico *ascendente* (p. ej., *rocketed*, *grow*) o como positivo si está en relación con un elemento léxico decreciente (p. ej., *fall* o *plummet*). La Tabla 14 aporta un resumen estadístico de la recopilación final.

<i>Tipo</i>		<i>Sentimiento</i>	
Palabras	1388	Positivos	1734
Expresiones multi palabra	5455	Neutros	2445
		Negativos	2664
<b>TOTAL</b>	<b>6843</b>		<b>6843</b>

**Tabla 14** Resumen estadístico de *LexiEcon*

#### 6.5.4 Evaluación de *LexiEcon*

Con el fin de evaluar el rendimiento del *plug-in*, se desarrolló un *gold standard* de dominio específico. Para elaborar un *gold standard* que permita evaluar el rendimiento del *plug-in LexiEcon*, se extrajo una muestra aleatoria<sup>55</sup> que incluye 400 oraciones aleatorias del ECBE. Posteriormente, dos evaluadores externos expertos en la materia clasificaron conjuntamente cada oración en tres categorías: positivo, negativo o neutro. Se les indicó que solo tuvieran en cuenta la información disponible en las oraciones y que anotaran las oraciones acordando una respuesta final. De manera similar a la propuesta de Wiebe et al. (2005), reproducida también por Malo et al., (2014), se instruyó a los anotadores para que consideren las siguientes pautas principales al anotar las frases:

- No hay reglas fijas sobre cómo se deben anotar palabras particulares.
- Dado el caso, evitar los sesgos basados en el conocimiento previo de la empresa o institución. Por lo tanto, cada oración de muestra debe anotarse utilizando la información que está explícitamente disponible.
- Se pide a los anotadores que sean coherentes con respecto a sus propias anotaciones.

Las tablas 15, 16 y 17 muestran los resultados de rendimiento (con y sin aplicar el léxico *plug-in*) que obtuvimos para el análisis del *gold standard* para cada una de las clases (positivo, negativo y neutro). Las cifras de rendimiento muestran que el *plug-in* mejora significativamente la precisión y la exhaustividad para oraciones positivas, negativas y neutras. Si bien las cifras de precisión son buenas para las oraciones negativas y neutras, una mejor precisión en las oraciones positivas parece estar motivada por el uso de recursos lingüísticos más complejos, como comparaciones y estructuras adversativas

<sup>55</sup> Intervalo de confianza de la muestra: 4,9. Nivel de confianza: 95%.

que incluyen elementos negativos, por lo tanto, las puntuaciones finales tienden a adquirir un sentimiento neutro y, en muchos casos volverse negativas. Por ejemplo:

- a. However, the proportion of the overall labour force that feels overemployed has fallen since the beginning of the century.
- b. Officials with the accounting board say the changes would bolster investor confidence by requiring the banks to more quickly recognize their losses.

	<i>Con plug-in</i>			<i>Sin plug-in</i>		
	Positivo	Negativo	Neutro	Positivo	Negativo	Neutro
Positivo	109	22	33	74	38	47
Negativo	13	135	19	17	110	78
Neutro	4	8	57	15	14	53

**Tabla 15** Matrices de confusión (fila = referencia; columna = test)

	<i>Con plug-in</i>				<i>Sin plug-in</i>			
	Positivo	Negativo	Neutro	Total	Positivo	Negativo	Neutro	Total
VP	109	135	57	301	74	110	53	237
FP	55	32	12	99	85	95	29	209
VN	219	203	279	701	255	189	239	683
FN	17	30	52	99	32	52	125	209

**Tabla 16** Resumen de verdaderos positivos, falsos positivos, verdaderos negativos y falsos negativos

	<i>Con plug-in</i>				<i>Sin plug-in</i>			
	Positivo	Negativo	Neutro	Media	Positivo	Negativo	Neutro	Media
Precisión	0,665	0,808	0,826	0,766	0,465	0,537	0,646	0,549
Exhaustividad	0,865	0,818	0,523	0,735	0,698	0,679	0,298	0,558
F1	0,752	0,813	0,640	0,735	0,559	0,600	0,408	0,522

**Tabla 17** Cifras de rendimiento

Por otro lado, las bajas cifras de exhaustividad para oraciones neutras parecen estar motivadas por razones similares, ya que 33 oraciones de carácter factual se consideraron positivas. En muchos casos, las oraciones de hecho presentan elementos léxicos positivos o negativos en combinación con estructuras adversativas o modos *irrealis*. Por ejemplo:

- a. As this continues, leadership, institutional strength and the ability to think in the long term – rather than focusing solely on short-term development – will determine how sustainable a city can become.
- b. The rest will be paid into an escrow account and clawed back if the bank makes a loss in a subsequent year.

De estos datos se obtienen dos conclusiones principales: (1) existe una gran diferencia entre la expresión de polaridad entre el dominio general y el dominio específico y (2) los niveles de precisión son altos debido a una menor ambigüedad en el uso del lenguaje.

En relación con la primera conclusión, se puede observar que se produce un aumento en torno al 20% en los datos de precisión y exhaustividad cuando se aplica *LexiEcon* en el corpus. El gran tamaño del *plug-in* permite una amplia cobertura de los usos especializados del léxico en este dominio. Sin embargo, en el dominio económico-financiero no solo influye el factor léxico, sino que también es muy relevante la presencia de patrones propios de la naturaleza del dominio (unidades direccionales y términos potencialmente positivos/negativos) que no quedan cubiertos con el lexicon nuclear. *LexiEcon* potencia especialmente el análisis de la opinión en oraciones factuales pero que claramente contienen rasgos evaluativos (p. ej., las valoraciones de los indicadores macroeconómicos, sus tendencias y predicciones), los cuales son un factor característico de este dominio. Por ejemplo:

- a. The index fell for the third week in succession for the first time since March.
- b. Sales of food, which accounts for 63% of the total, fell by 0.3%.

Se puede concluir que el estudio de las particularidades lingüísticas propias del dominio y su posterior implementación técnica permite alcanzar unos niveles óptimos en el rendimiento del sistema de análisis de sentimiento.

Por otro lado, los resultados de precisión y exhaustividad son notables, superando la media del 70%, destacado los buenos resultados especialmente en las oraciones negativas y neutras. Si bien se considera que el uso de los términos no es enteramente *unívoco*, de acuerdo con los desarrollos teóricos contemporáneos en terminología (p. ej., Cabré, 2000), los buenos resultados se justifican claramente la propia naturaleza del lenguaje especializado. Esto es un factor clave, ya que la expresión de la opinión en este dominio tiende a un bajo nivel de ambigüedad. Si bien en el discurso de la prensa semiespecializada se dan ambigüedades, redundancias, y una falta de precisión estricta, los niveles de precisión en el uso del lenguaje son mucho mayores en comparación con el dominio general. Asimismo, por lo general, los textos analizados están libres de usos estéticos o *creativos* del lenguaje los cuales son más frecuentes en el lenguaje general (evaluaciones implícitas, ironías, dobles sentidos...) y son más difíciles de analizar automáticamente. Esto se debe a que la redacción de los artículos económicos antepone

la claridad de la información con el fin de facilitar al público la comprensión de la información especializada. Por lo tanto, la naturaleza especializada de los textos facilita el análisis de sentimiento basado en léxico el cual, tras un amplio trabajo terminográfico, ofrece un mejor rendimiento que el obtenido en el análisis de textos del dominio general.

Como conclusión final, parece claro que *LexiEcon* cumple eficientemente como *plug-in* del lexicón nuclear de *Lingmotif* y abre una ventana de oportunidad al análisis de la expresión del sentimiento en el dominio económico en lengua inglesa. Estos resultados garantizan un análisis de sentimiento fiable de las oraciones que conforman la muestra y, por tanto, podemos proceder a la segunda parte de este trabajo: el análisis longitudinal del sentimiento.

## 6.6 Conclusiones

En este capítulo se han descrito la metodología, las fuentes de datos y las herramientas necesarias para llevar a cabo la tarea de analizar el sentimiento del contexto de una serie de términos durante la crisis global. Está dividido en cinco partes: una revisión de los paradigmas y enfoques metodológicos relacionados con el trabajo, el diseño metodológico de la tesis, el diseño y construcción de un corpus de dominio específico, la descripción de las herramientas externas utilizadas (el software de análisis de sentimiento *Lingmotif* y la suite de corpus *The Sketch Engine*) y la construcción de *LexiEcon* un gran lexicón de dominio específico en lengua inglesa para *Lingmotif*.

Los dos primeros apartados del capítulo han fundamentado un diseño metodológico mixto en cuatro pasos que permita extraer y clasificar la información sobre lenguajes especializados desde una perspectiva que generalmente no recibe mucha atención en la literatura: la expresión de la opinión y las emociones, un aspecto del lenguaje que, como parece evidente, empa la totalidad del uso del lenguaje.

La tercera parte de este capítulo trató sobre el diseño y recopilación de la fuente de datos: los textos que formaron el *Esmeraldas Corpus of Business English* de +24M de *tokens*. Este corpus está recopilado según los principios de equilibrio y representatividad tradicionales de la disciplina de lingüística de corpus. La colección de textos clasifica de manera cronológica todos los textos de la sección *Business* de las cabeceras *The Guardian* y *New York Times* a lo largo de los años en los que tuvo lugar la crisis (2007 y 2015). Para su extracción se utilizaron *arañas web* y los textos fueron clasificados de manera

semiautomática. El producto final fue procesado y empaquetado en *The Sketch Engine* y está disponible bajo demanda.

La cuarta parte del capítulo describió las dos herramientas externas utilizadas para cumplir con los objetivos: la suite de corpus *The Sketch Engine* y la aplicación de análisis de sentimiento *Lingmotif*. La primera es un administrador de corpus y un software de análisis de texto *online* que ofrece múltiples posibilidades (concordancia, extractor de terminología, *word sketches*, entre otras) y además ofrece una gran variedad de corpus de múltiples dominios y lenguas. Por otro lado, *Lingmotif* es una aplicación que analiza textos desde la perspectiva del análisis de sentimiento. En resumidas cuentas, permite determinar la orientación semántica (si es positivo o negativo y en qué grado) de un texto mediante la detección de expresiones lingüísticas que indican una determinada polaridad. Incluye un lexicón muy detallado en inglés y en español y, a partir de *plug-ins*, permite adaptar fácilmente el análisis de sentimiento a los distintos lenguajes especializados.

La quinta parte del capítulo ha servido para cumplir con el primer objetivo específico, describiéndose la recopilación de un lexicón de sentimiento de dominio económico-financiero de gran tamaño que se ajuste a las especificaciones de *Lingmotif*. El lexicón se recopiló de acuerdo con la base teórica revisada sobre el lenguaje evaluativo en el Capítulo 3, las especificidades del dominio económico-financiero y del dominio periodístico revisadas con profundidad en el Capítulo 4 y las características propias de la expresión del sentimiento de este dominio en el apartado 5.7.2. Para ello, se han combinado técnicas cuantitativas y cualitativas de extracción de terminología. Los candidatos a término se obtuvieron principalmente a partir de técnicas de extracción terminológica basada en corpus en las que, a través de *KeyWords* de *The SketchEngine*, se extrajeron del ECBE 30.000 candidatos a término. Con la ayuda de un experto en el dominio, cada uno de los candidatos fue comparado con los valores de las entradas del lexicón de sentimiento de dominio general. Se seleccionaron exclusivamente aquellas entradas cuyo sentimiento en el dominio específico difiriera del léxico base y a cada una de ellas se le asignó la valencia correspondiente al dominio específico.

Como resultado se presenta *LexiEcon*, un lexicón *plug-in* de dominio específico para la lengua inglesa que en la actualidad incluye 6.843 entradas. Los resultados de su evaluación fueron muy satisfactorios. Tras analizar un *gold standard* que contiene 400 oraciones aleatorias de dominio específico extraídas del corpus ECBE a través de dos barridos (primero aplicando tan solo con el lexicón general y, posteriormente, en conjunto con el *plug-in*), se observan excelentes niveles de precisión en el dominio. Las cifras de

rendimiento son notables pues utilizándose en conjunto con el lexicón de *Lingmotif* obtiene un *F1* medio de 0,735. Esta cifra supone alrededor de un 20% más que los resultados que ofrece *Lingmotif* sin *plug-in* cuando clasifica los textos del dominio económico-financiero.

Tras observar los resultados del análisis de las oraciones anotadas se evidenciaron las profundas diferencias que existen en la expresión del sentimiento entre el dominio general y el dominio específico. En efecto, a partir de la observación de estos resultados, se concluye que (1) existe una gran diferencia entre la expresión de polaridad entre el dominio general y el dominio específico y (2) los niveles de precisión son altos debido a una menor ambigüedad en el uso del lenguaje.

## 7. Análisis de datos

« En effet l'état social change ; des institutions s'en vont, d'autres viennent ; les sciences font des découvertes ; les peuples, se mêlant, mêlent leurs idiomes : de là l'inévitable création d'une foule de termes.

D'autre part, tandis que le fond même se modifie, arrivant à la désuétude de certains mots par la désuétude de certaines choses, et gagnant de nouveaux mots pour satisfaire à des choses nouvelles, le sens esthétique, qui ne fait défaut à aucune génération d'âge en âge, sollicite, de son côté, l'esprit à des combinaisons qui n'aient pas encore été essayées. »

Emile Littré (1801-1881), prefacio del *Diccionario de la lengua francesa* (1863-1872)

Este capítulo es central a este trabajo, ya que en él se describe de manera pormenorizada el análisis realizado de las unidades léxicas seleccionadas y sus contextos. Para cumplir con los objetivos restantes de este trabajo se seleccionaron diez palabras que fueron clave durante la crisis (“credit”, “debt”, “inflation”, “housing”, “markets”, “investor”, “public”, “reform”, “spending” y “company”). Cada análisis consta de tres partes: (a) una tabla de resultados estadísticos descriptivos longitudinales relativos a las puntuaciones de sentimiento, (b) una tabla anual de colocaciones y (c) una discusión sobre los hallazgos en el corpus.

Es pertinente explicar de manera concisa aquí el proceso aplicado para cumplir con el segundo objetivo específico: “Definir un subconjunto de datos longitudinal en forma de oraciones que contienen los términos de estudio y que serán el input del análisis de sentimiento”. A través del extractor de oraciones que ofrece la herramienta *Concordance* de *The Sketch Engine* se extrajeron del corpus las oraciones que incluían los términos clave seleccionados. De acuerdo con las especificaciones de Lingmotif, se creó un conjunto de datos para cada uno de los términos de estudio y cada uno de ellos contiene cuatro columnas (*id*, *file*, *text* y *date*). Como resultado, se generaron diez ficheros CSV estructurados de manera que cada línea incluía una oración según el formato descrito previamente en la Tabla 12. Estos ficheros están codificados según los estándares y están disponibles para la comunidad bajo demanda.

Cada uno de los ficheros de entrada fueron procesados posteriormente con la suite de análisis de sentimiento *Lingmotif* que aplicó el *plug-in* de dominio específico



*LexiEcon*, el cual devolvió sendos ficheros CSV estructurados según el formato descrito en el apartado 6.4.1. En la Figura 14 se puede observar un ejemplo ilustrativo de las distintas columnas incluidas en cada conjunto de datos. A partir de estos datos, un resumen estadístico de cada uno de los conjuntos de datos se verá al principio de cada apartado de este capítulo.

id	file	text	date	polarity	tss	tsi	pos_score	neg_score	neu_score	pos_items	neg_items	neu_items	none_items	sentenci	tokens	tt_ratio	lex_words	function_words
1	0701NYT100	In late Decem	28/01/2007	NONE	50	29	0	0	5	0	0	2	5	1	21	1.0	12	6
2	0701NYT126	In one case, a	28/01/2007	NEG	14	30	1	6	0	1	2	0	7	1	23	0.95700	12	10
3	0701NYT128	A forced merj	28/01/2007	NEG	0	38	0	4	2	0	2	1	5	1	21	0.905	13	7
4	0701NYT136	He slashed G:	28/01/2007	NEG	0	50	0	8	5	0	2	3	5	1	20	0.95	11	9
5	0701NYT136	Pressler slash	28/01/2007	POS	85	67	6	1	4	1	1	2	2	1	15	1.0	9	6
6	0701NYT148	Harris also ha	28/01/2007	NEG	0	11	0	2	0	0	1	0	8	1	14	0.929	8	6
7	0701NYT152	Those concer	28/01/2007	NEG	25	40	2	6	2	1	2	1	6	1	29	0.862	14	15
8	0701NYT157	Gordon Broth	28/01/2007	NEG	0	30	0	5	4	0	1	2	7	1	17	1.0	11	5
9	0701NYT157	Malden, base	28/01/2007	NEG	16	25	2	10	5	1	3	3	25	2	61	0.787	38	22
10	0701NYT165	With the assu	28/01/2007	POS	77	41	7	2	5	1	1	3	7	1	21	0.81	11	9

**Figura 14** Las diez primeras entradas del conjunto de datos de “debt” tras ser procesadas por *Lingmotif*

Tras obtener estos resultados, los datos fueron procesados automáticamente con las librerías *Pandas* (McKinney, 2010) y *Matplotlib* (Hunter, 2007) del lenguaje de programación *Python*. *Pandas* es una librería de software bajo licencia BSD diseñada para manipulación y análisis de datos. En particular, ofrece eficientemente estructuras de datos y operaciones para manipular tablas numéricas y series temporales. Por otro lado, *Matplotlib* es una versátil librería diseñada para la generación de gráficos a partir de datos estructurados. Con el fin de extraer las tendencias generales de sentimiento se suavizó la serie temporal, eliminando ruidos aleatorios y estacionalidad tras agregar los datos mediante una media anual de puntuaciones de sentimiento. Para facilitar la interpretación, se resumió esta información mediante tablas y gráficos.

Para cada uno de los términos estudiados se ofrece un análisis en tres partes. La primera observa la evolución del sentimiento mediante las cifras obtenidas a través de la media anual de las puntuaciones ofrecidas por el modelo de análisis de *Lingmotif*. Tras la observación de los datos estadísticos descriptivos, estos se dividieron en distintos intervalos temporales para su posterior análisis. Con vistas a cumplir con los objetivos específicos 3 y 4, la segunda parte corresponde a un análisis cuantitativo de las unidades léxicas (sustantivos comunes, verbos y adjetivos) que más cocurren con la unidad de estudio. Para ello se utilizó la utilidad de concordancia *The Sketch Engine* de manera que se extrajeron las colocaciones en un rango de cinco unidades. Los rankings para clasificar de mayor a menor significancia las unidades resultantes se ordenaron según el indicador estadístico *logDice* calculado por la herramienta. La tercera parte es una interpretación cualitativa con la intención de triangular los datos obtenidos. De esta manera, se pretende corroborar los hallazgos en las dos primeras partes y, además, trazar relaciones con los

*perfiles narrativos* históricos relatados en el Capítulo 2 relacionados con los eventos de la crisis.

### 7.1 Término 1: Credit

El término “credit” en el *Oxford English Dictionary* (2019) en su primera acepción como “The ability of a customer to obtain goods or services before payment, based on the trust that payment will be made in the future”. Como se revisó en el Capítulo 2, la congelación de un sistema financiero incapaz de ofrecer crédito fue el factor principal durante la primera etapa de la crisis global. De repente, la maquinaria financiera que durante los primeros años del siglo XXI ofrecía grandes facilidades para acceder al crédito, quedó paralizada tras la explosión de la burbuja de las hipotecas *subprime*. El optimismo que emanaban las enunciaciones relacionadas con el con el crédito a principios de siglo se invirtió bruscamente y, según los planteamientos de este trabajo, tras la explosión de la burbuja el contexto de este término pasó a enunciarse con una orientación semántica más negativa.

A través de la gráfica de medias de puntuación de sentimiento (Figura 15) y el resumen de descriptores estadísticos (Tabla 18), se pueden observar tres etapas claramente diferenciadas en las tendencias de la evolución del sentimiento de la palabra: (1) una primera etapa correspondiente a los albores de la crisis de crédito en el año 2007, (2) una segunda etapa de tres años (2008-2011) caracterizada por una puntuación de sentimiento negativa tras un descenso repentino, y, finalmente, (3) una etapa final con puntuaciones positivas en las puntuaciones medias de sentimiento con niveles similares a la primera etapa.

<i>Año</i>	<i>N</i>	<i>FQI1000</i>	<i>Sentimiento</i>	<i>pos_items</i>	<i>neg_items</i>	<i>neu_items</i>	<i>none_items</i>
2007	488	0,28	71,43	1,62	1,21	0,89	9,32
2008	1067	0,57	31,70	1,47	1,77	0,86	9,28
2009	884	0,44	31,61	1,47	1,74	0,92	9,57
2010	696	0,31	30,58	1,47	1,83	0,95	10,05
2011	787	0,29	31,25	1,43	1,67	0,98	9,84
2012	607	0,24	67,94	1,70	1,48	0,90	9,64
2013	688	0,19	72,50	1,93	1,31	0,99	10,23
2014	686	0,17	73,65	2,00	1,29	1,06	10,79
2015	861	0,20	68,87	1,82	1,56	1,03	10,80

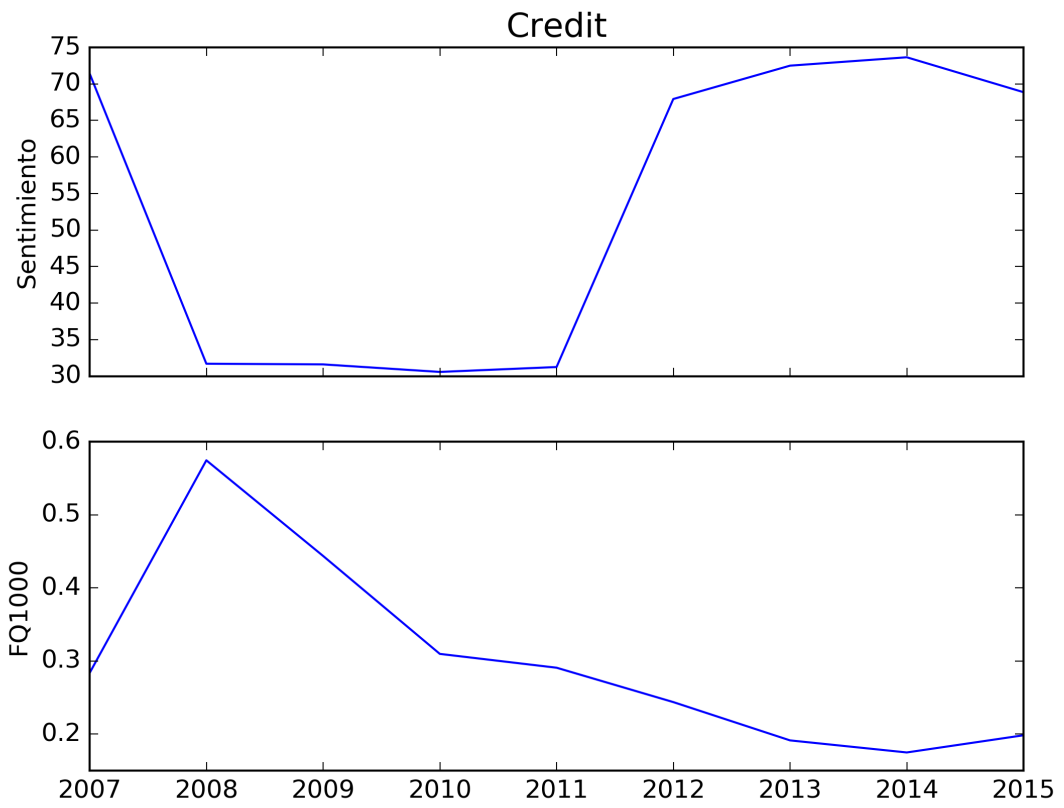
**Tabla 18** Resumen estadístico del análisis de las oraciones que incluyen “credit”

La primera fase corresponde a 2007, en los albores de la crisis, la palabra era utilizada en contextos con una orientación semántica positiva, alcanzando los 71,43 puntos. La segunda fase (2008-2011) corresponde claramente con la explosión de la



burbuja inmobiliaria y a los eventos que ocasionaron la congelación del sistema crediticio iconizados a través del terremoto provocado por la quiebra de Lehman Brothers. Esta etapa está manifiestamente marcada desde 2008 (con una puntuación negativa, 31,70). Estos valores medios negativos se mantuvieron estables durante los años más duros de la crisis crediticia para, finalmente, retornar a un contexto con valores medios positivos en 2012 (67,94 puntos). De hecho, la cifra de elementos positivos en las oraciones pasa de 1,62 en 2007 a 1,47 el año siguiente, mientras que las unidades negativas de 1,21 a 1,77 en el mismo periodo. Estas cifras se mantienen hasta el año 2012, comienzo del siguiente intervalo de estudio. La última fase (2012-2015) vuelve a contextualizar el término *credit* en un contexto positivo. Es pertinente recordar que en estas fechas la maquinaria periodística había puesto en marcha un discurso favorable a la reactivación del crédito desde los bancos centrales a la banca privada. Para entonces, la *Reserva Federal* estadounidense estaba consolidando su programa de flexibilización cuantitativa a través de la compra de activos bancarios y el nuevo paradigma discursivo tras el pronunciamiento “*Whatever it takes*” de Mario Draghi en 2012 estaba en marcha. En 2012, las puntuaciones de sentimiento tenían valores similares a los de 2007, con 67,94 puntos. No obstante, a pesar de la recuperación en los valores medios se observa una tendencia descendente al final del estudio de la muestra en el año 2015.

En cuanto a la frecuencia relativa de aparición de los términos a lo largo del corpus, se puede observar que la frecuencia base en 2007 (FQ1000 0,28) se duplica en 2008 (FQ1000 0,57). Tras este auge repentino, fruto patente de la frenética actualidad informativa, la frecuencia relativa descendiendo progresivamente aterrizando desde 2013 hacia el fin del estudio en 2015 (FQ1000 0,20). La frecuencia relativa es inversamente proporcional al sentimiento, es decir, cuanto más se ha publicado sobre *credit* en las noticias, se referenciaba el contexto de manera más negativa. Esto está motivado por las selecciones léxicas que aumentan lo noticiable de un hecho repentino y catastrófico como fue la paralización del sistema crediticio. A partir de este momento *credit* se convertía en una palabra evento y su mera mención se convierte en el reflejo de un momento histórico que afectó no solo a la economía y a las instituciones financieras sino en la cotidianidad de cientos de millones de personas alrededor del mundo.



**Figura 15** Gráfico longitudinal de las medias anuales de las puntuaciones de sentimiento para “credit” y la frecuencia relativa del término por cada 1000 tokens

### Primer intervalo de estudio

El siguiente paso corresponde, entonces, a profundizar con la observación de las colocaciones de “credit” en el corpus. Para ello se tomará como punto de observación las tres fases observadas en el paso anterior.

2007		
<i>Palabra</i>	<i>Cooc.</i>	<i>logDice</i>
card-n	102	10,741
carbon-n	36	9,245
market-n	69	9,133
<b>crunch-n</b>	28	9,069
tax-n	36	9,032
mortgage-n	20	8,219
line-n	19	8,146
global-j	18	8,025
<b>crisis-n</b>	14	7,963
bank-n	23	7,956
give-v	21	7,858
loan-n	15	7,816
new-j	27	7,729
<b>problem-n</b>	15	7,611
buy-v	19	7,601
home-n	15	7,546
consumer-n	15	7,482
use-v	19	7,445
fund-n	15	7,338

**Tabla 19** Principales colocaciones de “credit” durante el primer intervalo

La primera fase de análisis corresponde a la época inmediatamente previa a la explosión de la crisis (año 2007) y las oraciones analizadas priman elementos léxicos de carácter especializado con un sentimiento neutro, ya sean aquellos que forman unidades estables (como es el caso de “credit card” o “credit line”) u otros sustantivos como “carbon”, “market” o “tax” (Tabla 19). El análisis realizado mediante *The Sketch Engine* facilitó una serie de colocaciones que permiten observar de con facilidad las relaciones con las distintas categorías gramaticales que aparecen con el término objetivo, es el caso de:

- Sustantivos genéricos: *card, home*.
- Sustantivos específicos: *bank, carbon, consumer, crunch, fund, line, loan, market, mortgage, tax*.
- Adjetivos: *global, new*.
- Verbos: *buy, give, use*.

Entre los elementos de carácter neutro y pertenecientes al dominio de la economía, se destaca aquí el caso de “carbon credit”, el cual aparece 36 veces durante 2007 para salir del ranking en la siguiente etapa. El término hace referencia al crédito de carbono: un certificado comercializable que ofrece el derecho a emitir industrialmente una tonelada de dióxido de carbono, con el objetivo de reducir el crecimiento de los gases de efecto invernadero.

Gazprom, using Kyoto guidelines, plans to sell CARBON CREDITS to Europe.

In Europe, where there is a market for CARBON CREDITS , it is now worth only \$2 to offset a ton of carbon emissions.

To build a carbon market, its originators must create a currency of CARBON CREDITS that participants can trade.

CARBON CREDITS are advertised as a way for polluters to offset the greenhouse gases they produce.

Even some environmentalists oppose CARBON CREDITS because, they say, offsets tend to give the impression that global warming is being solved.

The world of CARBON CREDITS is filled with untested projects.

Durante este periodo, muchas de las palabras neutras, como es el caso de “tax”, suelen aparecer en contextos positivos o neutros. Una de sus realizaciones léxicas más frecuentes, aunque no la única, es la unidad “*tax credit*” (crédito fiscal), el cual hace referencia a un incentivo fiscal que permite a ciertos contribuyentes restar la cantidad del crédito que han acumulado del total que le deben al estado. Aquí se muestran algunos ejemplos extraídos aleatoriamente:

His business earns a TAX CREDIT for hiring former prisoners.

The TAX CREDITS have lapsed (and been restored) three times since then, and despite the stops and starts, the FPL Group has persevered.

Norman Lamont was told by senior civil servants in 1993 that he could raise £1bn a year if he reduced the TAX CREDIT on dividends to 20%.

A partir de julio de 2007, comenzó la relación de unidades negativas relativas al crédito, especialmente la expresión “credit crunch”. Esta última unidad (en castellano, contracción del crédito, una reducción severa y repentina del crédito) parte del ámbito especializado y terminó vulgarizándose y relexicalizándose de manera que en lengua inglesa se utiliza para referirse específicamente a la crisis de crédito que se mencionó en el apartado 2.1.1.

Now, with the dollar down more than 40 percent against the euro and the economy hamstrung by a deep CREDIT CRUNCH, a less assured America has become increasingly reliant on foreign capital

Chancellor Alistair Darling has said he wants a return to 'old-fashioned banking' from the complex, highly-leveraged business model that has left the banks vulnerable to the CREDIT CRUNCH that began in the US.

It is not surprising that in the growing CREDIT CRUNCH the market should start to become alarmed about its future viability.

The result was a CREDIT CRUNCH - the secondary banks found the financial environment much more bracing, with money both hard to come by and expensive.

The bank is not likely to be the only one affected by the CREDIT CRUNCH.

A CREDIT CRUNCH is going to hit industries like private equity and hedge funds very hard indeed.

La primera colocación con “*crisis*” ocurre también en el New York Times durante ese verano, específicamente a partir del mes de agosto de 2007 y hasta diciembre del mismo año aparece 14 veces en como colocación del término objeto. Todas estas relaciones con la crisis vienen relacionadas con el fenómeno del “credit crunch” y los problemas de la banca para financiarse.

The CRISIS in world CREDIT markets has cast a shadow over the financing of a number of big deals.

That loss, and Lehman's failure to communicate effectively with investors, catapulted the firm from one of many riding out the CREDIT CRISIS to one whose survival as an independent is in doubt.

He also knows that the late-summer turmoil in global bond markets because of the subprime housing problems and the ensuing CREDIT CRISIS could complicate his analysis.

## Segundo intervalo de estudio

Durante la segunda etapa de análisis (2008-2011), el contexto de *credit* se caracteriza por tener más colocaciones con unidades del dominio específico y sentimiento generalmente negativo como “default”, “squeeze” o “loss” y adjetivos como “tight”. Por otro lado, la puntuación de colocación de palabras neutras como p. ej., “rating”, “bank” o “tax” permanecen generalmente estables durante este periodo. En cuanto a las colocaciones, se encuentran las siguientes, junto al número de subcorpus anuales en los que aparece junto a “credit”:

- Sustantivos genéricos: *card* (4), *home* (2), *information*, *industry*, *image*, *month*, *office*, *percent*, *union*, *year*.
- Sustantivos específicos: *account*, *agency* (3), *analyst* (3), *bank* (4), *business* (2), *company* (3), *consumer* (2), *crisis* (4), *crunch* (4), *customer* (2), *debt* (3), *default* (3), *line* (3), *loan* (2), *loss*, *market* (2), *rating* (4), *risk*, *score* (2), *squeeze*, *swap* (2), *tax* (4).
- Adjetivos genéricos: *big*, *global*, *international*, *last* (2).
- Adjetivos específicos: *tight*.

- Verbos genéricos: *cut, give* (2).
- Verbos específicos: *buy, pay*.

Una simple revisión de la Tabla 20 observar a través de sus colocaciones el cambio de paradigma en el sentimiento. Hablar de “credit” en la prensa, y por extensión en las conversaciones a pie de calle, era hablar de crisis económica. Esto es claramente observable a través del rápido ascenso de “crunch” y “crisis” que pasan a tener índices de colocación muy similares a las de elementos tan establecidos léxicamente como “card”.

Como se puede ver en los dos primeros ejemplos a continuación, la evolución de las puntuaciones de *logDice* se justifican en que, mientras que en 2008 se hablaba de “credit crunch” como un fenómeno de actualidad. En su uso, se relaciona gramaticalmente con verbos con tiempos en *present perfect*. Este tiempo verbal se utiliza estilísticamente en prensa para referenciar eventos recientes en la línea temporal.

The arrival of the CREDIT CRUNCH last summer has made it extremely hard for indebted businesses to refinance while the commercial property market slowed dramatically with asset values falling.

However, the CREDIT CRUNCH has meant the banks that financed the mega deal have struggled to sell on the debt, dubbed the '£800m gorilla' in the City.

Este fenómeno, a partir de 2012, pasó a relacionarse progresivamente con tiempos en *past simple*. En el lenguaje periodístico en lengua inglesa, dicho tiempo es el más frecuente para referenciar eventos pasados sobre los que persisten sus consecuencias en la actualidad de la voz. Por ejemplo:

Coked-up bankers caused the CREDIT CRUNCH, according to the former drug tsar David Nutt.

That simply doesn't compare to the 150% bubbles we saw in some of the countries that were hit by the CREDIT CRUNCH.

El segundo colocado destacado es *crisis* y su uso y frecuencia son paralelos a los de *crunch*. Como se pudo observar, su coeficiente *logDice* varió en tan solo un año de 7,963 puntos a 11,002 y permaneciendo por encima de los 8 puntos hasta 2012, cuatro años después. Si bien hasta 2011, la crisis de crédito era considerada algo de actualidad, como evidencian el uso del *present simple* en gran parte de las oraciones, a partir de 2013 se empieza a referenciar en pasado, como se revisará en la tercera fase de estudio.

Alan Greenspan blamed much of the CREDIT CRUNCH and resulting financial crisis today on rampant demand from Europe for exotic financial derivatives based on the sale of US sub-prime mortgages.

Three years after the height of the CREDIT CRUNCH, bankers are still as reviled as ever, and it seems there are still too many of them.



2008			2009		
<i>Palabra</i>	<i>Cooc</i>	<i>logDice</i>	<i>Palabra</i>	<i>Cooc</i>	<i>logDice</i>
<b>crunch-n</b>	200	11,223	card-n	159	10,834
card-n	182	11,002	<b>crunch-n</b>	140	10,755
<b>crisis-n</b>	111	10,07	rating-n	66	9,588
market-n	156	9,737	<b>crisis-n</b>	57	9,136
rating-n	50	9,217	line-n	46	8,925
line-n	56	9,213	bank-n	69	8,636
score-n	37	8,849	market-n	58	8,519
<b>default-n</b>	33	8,663	tax-n	39	8,463
bank-n	67	8,595	consumer-n	34	8,234
<b>tight-j</b>	30	8,552	debt-n	31	8,232
swap-n	28	8,448	business-n	54	8,209
debt-n	33	8,435	<b>default-n</b>	23	8,173
loan-n	37	8,403	International-n	24	8,107
global-j	34	8,379	home-n	29	8,095
last-j	57	8,349	give-v	32	8,045
tax-n	30	8,263	loan-n	28	8,042
<b>squeeze-n</b>	24	8,257	agency-n	26	8,038
home-n	31	8,120	company-n	74	7,959
<b>loss-n</b>	27	8,007	analyst-n	25	7,885
customer-n	27	7,957	industry-n	26	7,612
2010			2011		
<i>Palabra</i>	<i>Cooc</i>	<i>logDice</i>	<i>Palabra</i>	<i>Cooc</i>	<i>logDice</i>
card-n	146	10,863	card-n	159	10,859
<b>crunch-n</b>	122	10,728	rating-n	111	10,323
rating-n	88	10,133	<b>crunch-n</b>	95	10,302
agency-n	47	8,961	<b>default-n</b>	34	8,758
tax-n	44	8,744	swap-n	29	8,607
<b>crisis-n</b>	38	8,62	<b>crisis-n</b>	41	8,532
bank-n	56	8,41	tax-n	43	8,528
analyst-n	33	8,374	agency-n	37	8,501
<b>default-n</b>	24	8,374	bank-n	57	8,308
score-n	21	8,214	<b>risk-n</b>	27	8,042
market-n	49	8,211	union-n	23	7,957
swap-n	19	8,069	analyst-n	24	7,773
union-n	21	7,857	<b>cut-v</b>	21	7,735
global-j	20	7,685	office-n	20	7,68
loan-n	19	7,684	business-n	46	7,587
include-v	29	7,622	pay-v	20	7,186
give-v	22	7,548	year-n	51	7,154
company-n	56	7,539	last-j	28	7,152
year-n	53	7,417	big-j	19	7,058
month-n	23	7,400	company-n	47	7,031

**Tabla 20** Principales colocaciones de “credit” durante el segundo intervalo

La tercera palabra de sentimiento negativo destacada sería “default” (incumplimiento crediticio), la cual destaca durante un lapso de cuatro años (2008-2010) y se sitúa en contextos claramente negativos. A continuación, pueden observarse algunos ejemplos:

For a student, a DEFAULT can destroy a CREDIT record, making it hard even to rent an apartment, let alone buy a home.

As the financial crisis deepened, hundreds of banks and thrifts closed and thousands more were saddled with bad loans and CREDIT card DEFAULTS , costing the industry billions of dollars.

Banks and card companies are bracing for a wave of DEFAULTS on CREDIT card debt in early 2009, and they are vying with each other to get paid first.

La colocación más frecuente de esta unidad con “credit” suele ser el término “credit default swap” (permuta de incumplimiento crediticio), un producto financiero que consiste en una operación financiera de cobertura de riesgos en el que el comprador de la permuta realiza una serie de pagos periódicos al vendedor y, a cambio, recibe de éste una cantidad de dinero en caso de que el título crediticio sea impagado. Dado que el incumplimiento era anteriormente una ocurrencia relativamente rara en la mayoría de los contratos de “credit default swaps” los pagos de primas del comprador al vendedor eran pequeñas y además no estaban reguladas por lo que eran altamente especulativas (Moreno, 2010). El mercado de los “credit default swaps” (CDS) atrajo una considerable preocupación de los reguladores tras el colapso de Bear Stearns en 2008 y sus consecuencias. La ampliación de los CDS fue responsable de la percepción de que dicha entidad financiera era vulnerable y, por lo tanto, restringió su acceso al capital mayorista y finalmente forzó su venta a JP Morgan. Esta colocación está reflejada ampliamente con un sentimiento negativo a lo largo del corpus y, como se puede observar en los ejemplos siguientes, estrechamente relacionada con la actualidad de la crisis bancaria.

After the sub-prime mortgage bubble burst in 2007, complex investments called CREDIT DEFAULT SWAPS, which insured against default of securities tied to the mortgages, collapsed.

In addition, institutions that had bought CREDIT DEFAULT swaps from Bear Stearns, insurance policies that protect against corporate bond defaults, were scrambling to undo those trades as the firm's ability to pay the claims looked dicier.

As the fund was being wound down, UBS said about 70 percent of its losses came from exposure to CREDIT DEFAULT SWAPS.

The banks have faced public criticism since the financial crisis for the opaque manner in which their traders bought and sold CREDIT DEFAULT SWAPS , a type of financial contract that allows investors to speculate and hedge against losses and that figured prominently in the crisis.

Buffett, known as the "Sage of Omaha" once warned that CREDIT DEFAULT SWAPS were "weapons of financial mass destruction".

Como enlace entre los dos intervalos de estudio, las palabras “rating” / “agency” con una orientación semántica neutra no estaban presentes en el ranking de 2007 y, sin embargo, aparecen en las primeras posiciones desde 2008 manteniéndose estables por encima de los 9 puntos hasta el último hito de estudio en 2015.

Las “rating agencies” (agencias de calificación crediticia) son firmas que califican a los productos de crédito según la capacidad del deudor para hacer frente a los pagos. Estas agencias jugaron un papel destacado en varias partes de la crisis partiendo de la crisis de las *subprime* desde el otoño de 2007. Los paquetes hipotecarios basura estaban calificados negligentemente por las *Big Three*: Standard & Poor’s, Moody’s, y Fitch Ratings.

A finales de 2009, más de la mitad de las obligaciones de deuda garantizadas por valor emitidas calificadas como AAA al final de la burbuja hipotecaria, pasaron a estar calificadas como basura. La complicidad manifiesta entre las agencias de calificación y las entidades financieras fue el desencadenante de la crisis crediticia, provocando finalmente el colapso de los tres bancos principales durante la burbuja: Lehman Brothers, Merrill Lynch y Bear Sterns.

A pesar de que “rating” es un elemento específico de carácter neutro en aislamiento, y como tal fue etiquetado en *LexiEcon*, el contexto en el que el uso de la palabra “credit” en colocación con “rating” es absolutamente negativo para los años 2008-2010. En este contexto, por lo general, el objeto de opinión hace referencia a compañías y entidades financieras. En la vecindad de término se encuentran unidades como “insufficient”, “cannot secure”, o “junk”. Una simple búsqueda de concordancias en el corpus pone de manifiesto la relación estrecha del término con el turbulento contexto económico:

At the moment most car firms, having insufficient CREDIT RATINGS , cannot secure a loan worth more than £200m from the EIB fund without government guarantees.

Standard & Poor's confirmed that situation after it cut its long-term CREDIT RATING on the company to BBB from A and said more downgrades were likely.

Buffett can't qualify for the best CREDIT RATING .

But these deals are only as good as the party on the other side, and when Security Capital's crucial CREDIT RATING was cut to junk this year, the potential that the company would pay out on the arrangement dimmed.

Just as in the mortgage markets, a sterling CREDIT RATING – the bond insurer's seal of approval – is no longer trusted.

Its CREDIT RATINGS were cut yesterday, adding to fears that it could become the next victim of the financial crisis.

The banks say they are worried about the CREDIT RATINGS of the countries that have been rescuing them.

Superada superficialmente la primera crisis a partir de 2010, las agencias de “rating” vuelven a tomar relevancia en el contexto de la crisis de la deuda soberana. Dichas agencias de calificación crediticia también emiten calificaciones para prestatarios estatales en función de su capacidad de devolver los préstamos. Además, las evaluaciones de crédito reflejan no solo el riesgo de incumplimiento percibido a largo plazo, sino también sirven para estudiar la evolución política y económica a corto plazo, lo cual se convierte en una medida de presión *de facto* del poder financiero hacia los estados. Entre los años 2010 y 2011 se produce un cambio en el campo semántico de los objetos de opinión relacionados con “credit rating” y “credit rating agencies”, pues estas pasan del campo de la banca al campo de las instituciones gubernamentales. En cuanto a la orientación semántica de los conjuntos y el puede observarse cualitativamente que las muestras observadas suelen ser de carácter negativo:

Elsewhere in Europe , CREDIT RATING agency Moody's downgraded Spain because of concerns over its weak growth prospects and considerable deterioration in the government's financial strength.

The CREDIT RATING agency Standard & Poor's raised its sovereign debt rating for Indonesia to BB+ last month, becoming the last of the three big agencies to rate the country one peg below investment grade.

The Observer Global leaders race to stem panic over US CREDIT RATING downgrade.

Iceland's CREDIT RATING outlook downgraded as Icesave row drags on

Finalmente, si bien en el periodo 2012-2015 el episodio de la crisis de la deuda soberana seguía estando de actualidad, en este contexto las compañías vuelven a aparecer con frecuencia como objetos de opinión. Al igual que en los tres periodos anteriores, “rating” sigue apareciendo en contextos negativos, como puede observarse en la muestra extraída del subcorpus que comprende los textos entre los años 2013 y 2015:

Lloyds Banking Group , revealed it too would set up legal entities in England to protect its CREDIT RATING.

CREDIT RATING AGENCIES will not consider Greece to be in default based on missing the I.M.F. payment, for the technical reason that the I.M.F. is not considered a commercial borrower.

Tesco's CREDIT RATING price has been cut by Standard & Poor's

CREDIT RATING AGENCY Moody's, which downgraded its outlook for the company's ability to repay its debts last week, said Phones 4u had £205m in notes due by 2019, £430m due in 2018, and a £125m revolving credit facility.

UK's CREDIT RATING downgraded from AAA to AA1 by Moody's

At the time, debt-laden Fortescue had looked to sell part of its rail and port assets after CREDIT RATINGS AGENCY Standard and Poor's said the company was the world's most vulnerable iron ore miner to lower prices.

On Monday, Standard & Poor's CREDIT RATING AGENCY downgraded RSA's debt to A- and put it on watch for further rating action within 90 days.

### Tercer intervalo de estudio

Por último, se analizarán las colocaciones del tercer y último periodo de análisis, relativo al intervalo entre los años 2012 y 2015. Aquí, el contexto de “*credit*” también aparece con unidades del dominio específico y sentimiento generalmente neutro. En relación con el intervalo anterior, la novedad reside en el ranking aquí presenta también unidades con sentimiento positivo (pasando de 31,25 a 67,94 entre 2011 y 2012) como es el caso de “*help*” en 2013 y, en 2014 el adjetivo “*cheap*” y los verbos “*expand*” y “*help*”. La ligera tendencia descendente que se visualiza entre las medias de 2014 y 2015 puede justificarse aquí al primar la terminología con sentimiento neutro entre sus colocaciones.

Los únicos elementos especializados de carácter negativo serían dos: el verbo “*cut*” y la ya mencionada reaparición de “*crisis*” aunque con un coeficiente *logDice* de 7,363, esto es, tres puntos por debajo de su pico en 2008. Además, de la clasificación de palabras con más colocados en este periodo que se muestra a continuación, en la Tabla 21 puede observarse con detalle el ranking de colocaciones.

- Sustantivos genéricos: *analyst* (2), *image* (2), *information*, *month*.
- Sustantivos específicos: *account* (2), *agency* (4), *bank* (4), *business*, *card* (4), *company* (3), *consumer* (3), *country*, *credit*, *crisis* (3), *crunch* (3), *customer* (2), *debt* (4), *executive*, *investment*, *line* (3), *loan*, *market*, *percent* (2), *price*, *rating* (4), *report*, *service*, *score*, *tax* (4), *union* (3), *year*.
- Adjetivos genéricos: *cheap*, *last*.
- Verbos genéricos: *give*, *help* (2), *include*.
- Verbos específicos: *buy*, *cut*, *expand*, *offer*, *sell*.

2012			2013		
<i>Palabra</i>	<i>Cooc</i>	<i>logDice</i>	<i>Palabra</i>	<i>Cooc.</i>	<i>logDice</i>
card-n	119	10,779	card-n	126	10,455
<b>crunch-n</b>	57	9,891	rating-n	60	9,363
rating-n	56	9,698	<b>crunch-n</b>	42	9,049
score-n	27	8,792	tax-n	46	8,427
tax-n	46	8,745	bank-n	66	8,358
agency-n	35	8,666	union-n	23	7,896
line-n	25	8,296	agency-n	27	7,863
image-n	20	8,255	<b>crisis-n</b>	25	7,832
<b>crisis-n</b>	29	8,247	<b>help-v</b>	30	7,424
information-n	19	7,891	consumer-n	23	7,302
analyst-n	21	7,873	last-j	34	7,287
account-n	17	7,789	include-v	28	7,117
bank-n	37	7,762	market-n	30	7,076
customer-n	24	7,739	price-n	21	6,922
debt-n	20	7,698	company-n	57	6,919
consumer-n	19	7,517	country-n	21	6,917
percent-n	24	7,491	year-n	50	6,868
buy-v	18	7,449	business-n	42	6,834
company-n	56	7,444	month-n	19	6,821
2014			2015		
<i>Palabra</i>	<i>Cooc</i>	<i>logDice</i>	<i>logDice</i>	<i>Cooc</i>	<i>logDice</i>
card-n	143	10,793	card-n	162	10,803
rating-n	59	9,602	rating-n	106	10,25
tax-n	70	9,154	tax-n	104	9,299
debit-n	30	8,771	agency-n	56	9,014
agency-n	29	8,125	union-n	46	8,846
<b>crunch-n</b>	19	8,125	debit-n	25	8,367
bank-n	54	8,094	bank-n	73	8,142
union-n	24	8,079	loan-n	28	8,042
image-n	20	7,993	line-n	23	7,759
<b>cheap-j</b>	19	7,843	analyst-n	28	7,677
account-n	22	7,816	debt-n	25	7,574
score-n	13	7,538	<b>cut-v</b>	24	7,484
offer-v	22	7,515	customer-n	34	7,451
<b>expand-v</b>	16	7,485	<b>crisis-n</b>	22	7,363
<b>help-v</b>	20	7,406	report-n	23	7,257
line-n	17	7,398	executive-n	28	7,177
give-v	25	7,278	consumer-n	22	7,140
company-n	70	7,237	investment-n	23	7,105
sell-v	18	7,193	service-n	23	7,081

**Tabla 21** Principales colocaciones de “credit” durante el tercer intervalo

En 2013 destaca la colocación con el verbo “help”. Sin embargo, el uso de esta unidad no parecer ser significativo al no hacer referencia a eventos históricos. De hecho,

entre sus 30 ocurrencias, existen múltiples ejemplos que no son del todo de carácter positivo, incluyendo expresiones irónicas o concierto significado modal (3º y 4º ejemplo):

Minister Li Keqiang, who took office in March, has said he plans changes that will promote sustainable growth, as opposed to relying on the easy CREDIT from state-controlled banks that HELPED the country rebound since the 2008 financial crisis.

But others doubt that unions would get such a carve-out because it would also encourage many nonunion workers to seek tax CREDITS to HELP with their employer-based plans.

The SEC has disputed accusations that it has not done enough to tackle the individuals and companies that HELPED cause the CREDIT crunch.

Legal to Censor, but Unwise Gabe Rottman, American Civil Liberties Union Pulling CREDIT card services would HELP the haters and hurt free expression.

It doesn't HELP your teen establish a CREDIT history or a relationship with a financial institution, so we recommend going the traditional route and opening up a checking account at your bank or credit union of choice.

American consumers riding out the last of the CREDIT boom HELPED lift global sales to a record 9.37 million cars in 2007.

De la misma manera, la presencia de “cheap credit” en el ranking de 2014 fue debido a hechos altamente noticiosos: el anuncio de la bajada de tipos de interés por parte del BCE como precuela del *Plan Draghi* de expansión cuantitativa. La bajada de tipos de interés fue acogida positivamente como se puede observar en los ejemplos más abajo. Sin embargo, es justo mencionar que existen ejemplos ponen de manifiesto visiones interpretativas y comentarios que ponen en cuestión la positividad del fenómeno. Estos ponen sobre la mesa la escasa demanda inicial de créditos a menor precio y las reticencias de los poderes centrales de la UE ante tales políticas. Algunos ejemplos encontrados en el subcorpus de 2014 son los siguientes:

Draghi has come under intense pressure over recent months to make CHEAPER CREDIT available to households and businesses to boost growth in the 18-member currency zone.

John Leech, head of automotive at consulting group KPMG, said that while CHEAP CREDIT had been great for driving sales

CHEAP CREDIT is essential when households and businesses are close to going bust.

Mario Draghi, the ECB president, is expected to tell his audience today that he is waiting to see how his previous attempts at offering cheap credit are faring before considering QE.

His "three arrows" of reform – higher government spending, CHEAPER CREDIT and shaking up old-fashioned business practices – appeared to be working.

ECB president Mario Draghi has responded to the lack of growth in the eurozone with CHEAPER CREDIT for banks in the hope they will increase their lending to businesses and consumers.

Financial markets were surprised to find limited demand for the ECB's CHEAP CREDIT scheme.

Germany wants the ECB to keep its powder dry, claiming CHEAP CREDIT will lead to rampant inflation.



El espíritu de recuperación de la positividad en el sentimiento del contexto de “credit” puede ejemplificarse con la aparición de “expand” como colocación en el subcorpus de 2014. Se hace referencia a la aparición en este trienio de algunas políticas expansivas relativas al crédito, como es el caso de la ampliación de las condiciones de los créditos fiscales que si bien fueron insuficientes supusieron algo de optimismo en el panorama depresivo observado en los años anteriores. Se observan aquí múltiples referencias a la expansión de los “tax credits” tanto en el contexto estadounidense como en el europeo. Un “tax credit” (crédito fiscal), es un incentivo fiscal que permite a ciertos contribuyentes restar la cantidad del crédito que han acumulado del total que le deben al estado. A continuación, pueden observarse algunas colocaciones seleccionadas de manera aleatoria:

Mr. Cowen adds other potentially useful policies, like EXPANDING the earned-income tax CREDIT or using urban policy to, say, make it easier for people who are not rich to live in San Francisco.

It's the give-and-take of the legislative process," explained Mr. Brown, who actually fought to EXPAND the tax CREDIT to ensure blighted manufacturing areas have access.

His threat sent negotiators back to the table to see if Republicans could add measures that would win liberal support, especially a permanently EXPANDED child tax CREDIT for the working poor.

Left off were the two tax breaks valued most by liberal Democrats: a permanently EXPANDED earned-income credit and a child tax CREDIT for the working poor.

Part of the additional resources helping new and smaller creditworthy companies obtain CREDIT , to EXPAND economic activity, jobs and productivity.

The ECB has already taken steps to EXPAND the supply of CREDIT in an effort to drive down borrowing costs and ease pressure on household budgets.

Se finalizará este análisis con la última observación de la colocación de “credit” y “crisis”. Si bien ya en 2012 y 2013, el nivel de cocurrencia con “crisis” descendió al tercio medio de la lista, en este último intervalo de estudio su uso varió significativamente pues este conjunto de unidades tendió a utilizarse para relatar eventos pasados. Como se puede ver en los tres últimos ejemplos, correspondientes a dichos años, la crisis de crédito se enuncia aquí mediante verbos en *past simple*, entendiéndose esto como un evento ya acaecido o en vías de superación. Tal es así que, durante 2014, esta palabra ni siquiera aparece en el ranking del año siguiente y vuelve a reaparecer en 2015. Existen múltiples ejemplos en términos similares a los que se observarán a continuación:

Its muscular response to the CREDIT CRISIS staved off economic disaster.

When the CREDIT CRISIS appeared, that financing – a significant part of which had come from selling short-term securities to United States money market funds – dried up.

The CREDIT CRISIS damaged that consensus as millions of homeowners lost their homes.



## 7.2 Término 2: *Debt*

El término *debt* es una unidad perteneciente al dominio económico que se define en el *Collins Cobuild* (2019) en su primera acepción como “something that is owed, such as money, goods, or services”, en su segunda acepción como “bad debt“ : “a sum of money that has been lent but is not likely to be repaid” y en su tercera definición como “an obligation to pay or perform something; liability”. Si bien su segunda acepción es negativa, en este estudio del dominio específico se toma en cuenta que el término cuando es usado de manera aislada indica un sentimiento neutro en su sentimiento y así queda establecido en el lexicón *plug-in*. No obstante, para la elaboración de *LexiEcon* se ha considerado un elemento potencialmente negativo, es decir, las expresiones que indiquen altas cantidades de deuda serán consideradas negativas, mientras que un descenso en sus niveles o su total cancelación será considerado generalmente como positivo (Figura 16).

<debt>	NN	neu	2
high_3_<debt>	NN	neg	3
low_3_<debt>	NN	pos	2

**Figura 16** Ejemplo de entradas relacionadas con “debt” en *LexiEcon*

La congelación de un sistema financiero incapaz de ofrecer crédito fue un factor muy relevante durante la primera etapa de la crisis global y los cambios en el uso de “debt” se reflejaron en sus contextos, cuya orientación semántica pasó a ser más negativa. Tras la paralización del acceso crédito millones de individuos, instituciones y empresas alrededor del mundo fueron incapaces de asumir pagos y aumentó el nivel de endeudamiento tras el colapso de la burbuja financiera acelerando el declive económico. De un día para otro, estas entidades fueron incapaces de afrontar pagos. Consecuencia de la crisis de crédito, la deuda pasó a convertirse en uno de los problemas principales para la población. Además de la quiebra y el rescate de las entidades bancarias privadas mediante la iniciativa pública alrededor de todo el globo, tras el colapso del sistema bancario, la crisis de la deuda se extendió a la Unión Europea desde 2009, extendiéndose principalmente a Portugal, Irlanda, Italia, Grecia y España.

A partir de entonces, los estados fueron incapaces de pagar o refinanciar su deuda pública ni de rescatar a los bancos con exceso de deuda sin la asistencia de la *Troika*: Comisión Europea, el Banco Central Europeo o el Fondo Monetario Internacional. Entre los múltiples factores estructurales que acrecentaron este problema, las causas de la crisis en la Zona Euro radican en la unión monetaria sin unión fiscal, unas diferencias muy

significativas en las condiciones sociales y materiales (p. ej., las pensiones públicas) entre estados miembros a lo que se sumó una limitadísima capacidad de respuesta de los líderes europeos. Esta asimetría de condiciones convirtió a la crisis económica en una crisis política y una emergencia social que permeó en todas las capas de la sociedad durante la mayor parte del periodo de estudio. Tras la crisis de la deuda, la población fue el sector que sufrió los reveses económicos más graves desde la Gran Depresión. Durante este periodo se produjo un aumento generalizado en el desempleo (España y Grecia superaron el umbral del 27%), decenas de miles de desahucios por impago y recortes en las coberturas de sanidad, educación y prestaciones sociales básicas.

El siguiente paso consiste en observar los factores que condicionaron el movimiento semántico de los valores de sentimiento del contexto del término “debt”. A través de los datos textométricos (Tabla 22) y la gráfica de medias de puntuación de sentimiento (Figura 17), se pueden observar tres etapas claramente diferenciadas en las tendencias de la evolución del sentimiento de la palabra. Teniendo en cuenta la evolución de las puntuaciones de sentimiento, se observa: (1) una primera etapa correspondiente a los albores de la crisis de crédito en el año 2007, (2) una segunda etapa de tres años que inicia con un pronunciado descenso a valores de sentimiento negativos (2008-2011) caracterizada por una puntuación de sentimiento negativa tras un descenso repentino y un posterior descenso más progresivo hasta alcanzar el mínimo histórico en 2011, (3) una etapa final (2012-2015) con tendencia ligeramente ascendiente con simetría en las puntuaciones, similares a las de la segunda etapa.

En el primer periodo (año 2007), la frecuencia relativa del término parte de una puntuación de 0,18, con 302 incidencias en el subcorpus. Según el modelo de análisis de sentimiento, la puntuación media de sentimiento parte de un sentimiento neutro en 2007 (50,56).



**Figura 17** Gráfico longitudinal de las medias anuales de las puntuaciones de sentimiento para “debt” y la frecuencia relativa del término por cada 1000 tokens

El segundo periodo (2008-2011) inicia con un descenso pronunciado de 9 puntos (41,01 en 2008) en el que continúa descendiendo de manera más amortiguada hasta llegar a los 39,61. Desde entonces, según los datos de agregación, la polaridad media de las oraciones que incluyen el término “debt” pasa a ser negativa. Asimismo, hay que destacar que esta caída en los valores de sentimiento ha ido acompañada de un ascenso contundente en la frecuencia relativa del término en el corpus subiendo 0,12 puntos en un año (FQ1000 0,31 en 2008) y 0,10 más el año siguiente (FQ1000 0,41 en 2009) llegando a su máximo histórico en 2011 (FQ1000 0,47). Se produce aquí el mismo fenómeno que en el análisis de “credit”: mientras más frecuente es un término en el discurso de la prensa, el contexto de este se valora de manera más negativa, respondiendo a los criterios clásicos de noticiabilidad, pues la prensa generalmente cumple la función de informar cuando un evento de relevancia tiene lugar especialmente cuando lo expresado es altamente evaluativo (emoción, expectativas, etc.) de manera que se aumente el hecho noticiable al compartir dichos sentimientos. En efecto, la media de elementos negativos se dispara de 1,21 en 2007 a 1,95 en 2011 (Bednarek y Caple, 2012).

El tercer periodo (2012-2015) refleja una leve tendencia al alza en las medias de sentimiento que de manera estable superan los 40 puntos (41,11, 42,91, 44,33 y 43,94 respectivamente). Tras un periodo inicial de continuidad en el nivel de frecuencia en el discurso (FQ1000, 0,46 en 2012), la frecuencia cae a la mitad (0,25 y 0,24 en 2013 y 2014), pero en 2015 vuelve a aumentar (0,38) tras los eventos relacionados, principalmente, con la crisis entre Grecia y la UE. En este intervalo la cifra de elementos positivos alcanza su máximo en 2014 con una media de 1,45.

Si se compara esta frecuencia relativa final, se puede considerar que el uso del término se consolida como un factor destacado y recurrente en el discurso de la prensa y, sin duda, uno de los factores más destacados de la crisis. A partir de 2008 “debt” se convertía en una palabra evento y un factor clave en el discurso de la población durante esta década.

<i>Año</i>	<i>N</i>	<i>FQ1000</i>	<i>Sentimiento</i>	<i>pos_items</i>	<i>neg_items</i>	<i>neu_items</i>	<i>none_items</i>
2007	302	0,18	50,56	1,11	1,21	1,64	8,68
2008	570	0,31	41,01	1,05	1,50	1,53	8,83
2009	809	0,41	42,46	1,07	1,49	1,57	9,12
2010	1038	0,46	39,61	1,15	1,90	1,59	9,59
2011	1263	0,47	38,47	1,20	1,95	1,60	9,45
2012	1148	0,46	42,11	1,24	1,76	1,57	9,30
2013	904	0,25	42,91	1,27	1,73	1,54	9,16
2014	945	0,24	44,33	1,45	1,74	1,64	9,72
2015	1657	0,38	43,94	1,41	1,81	1,67	9,75

**Tabla 22** Resumen estadístico del análisis de las oraciones que incluyen “debt”.

### Primer intervalo de estudio

La primera fase de análisis corresponde a la época inmediatamente previa a la explosión de la crisis (año 2007). Tras analizar las oraciones en busca de colocaciones se observó que priman elementos léxicos de carácter especializado con un sentimiento neutro, así como la presencia de adjetivos y verbos del dominio específico con un sentimiento de carácter tanto positivo como negativo. El análisis del listado de colocaciones (Tabla 23) clasificado según su categoría gramatical es el siguiente:

- Sustantivos genéricos: *amount, dollar, proportion*.
- Sustantivos específicos: *balance, collection, exposure, mortgage, obligation, security, sheet*.
- Adjetivos específicos: *cheap, collateralized, risky*.
- Verbos específicos: *assume, finance, pay, repay, reduce*.

2007		
<i>Palabra</i>	<i>Cooc.</i>	<i>logDice</i>
obligation-n	27	11,140
collateralized-j	15	10,497
finance-v	7	8,968
repay-v	5	8,778
amount-n	9	8,749
assume-v	5	8,594
<b>reduce-v</b>	8	8,542
proportion-n	4	8,540
dollar-n	9	8,453
<b>risky-j</b>	4	8,352
collection-n	4	8,320
sheet-n	4	8,320
<b>exposure-n</b>	4	8,310
security-n	7	8,301
mortgage-n	8	8,275
balance-n	4	8,205
pay-v	15	8,047
<b>cheap-j</b>	4	8,000
<b>deal-n</b>	12	7,987

**Tabla 23** Principales colocaciones de “debt” durante el primer intervalo

Una búsqueda aleatoria sencilla en el subcorpus de 2007 aporta una visión general del uso de la unidad en el contexto. Como se puede apreciar a continuación, el uso de “debt” aporta un carácter descriptivo de la situación económica sin una preponderancia clara de un sentimiento en específico en el contexto. En los ejemplos observados se alternan indistintamente segmentos factuales y evaluativos. Para profundizar es pertinente, por lo tanto, observar las evidencias que subyacen del uso de los términos encontrados en los listados y que son aparentemente de carácter neutro. Lo que sí queda patente es el uso de carácter especializado en los textos periodísticos ya que, dados los ejemplos, se usa discursivamente como un mero indicador macroeconómico:

With variations, the same trends have persisted since 2002, when the U.S. government again began adding to the national DEBT , after a brief period of paying debt down.

Critics say that the sale will lead to higher prices for users to pay off \$24.5 billion in new DEBT that the buyers put on TXU.

Consumption, of course, is not an ideal indicator of well-being; a high or steady level of purchases may reflect growing DEBT , and the ease of buying a big-screen TV does not reflect a comparable ease in buying good health care.

Skye contends that when Bandagro went bankrupt in 1981, Venezuela assured investors that it would assume and pay the bank's DEBTS .

We have dramatically reduced the amount of DEBT, and in various locations we've refinanced dollar debt with local currency debt so we don't have that same kind of foreign currency exposure.

Las primeras colocaciones de carácter específico en el ranking son “obligation” y “collateralized”. Principalmente, esto hace referencia a las “collateralized debt obligations” (CDO) ‘obligación de deuda garantizada’. Este término hace referencia a un tipo de seguridad respaldada por activos que originalmente se desarrolló como un instrumento para los mercados de deuda corporativa. A partir del año 2002 comenzaron a utilizarse como medio para refinanciar valores respaldados por hipotecas. Para este año de estudio, 2007, el mercado de CDO quedó dominado por “tranches” ‘tramos’ calificados como de alto riesgo cuyos activos eran principalmente hipotecas *subprime*. Su mención en el corpus está relacionada con su uso como producto que incentivase el flujo de préstamos de alto riesgo, principalmente hipotecario, que condujo a la crisis de las *subprime* de 2007-2009. En casi todos sus contextos, el uso de esta terminología tiene una polaridad negativa y forma parte de segmentos de carácter evaluativo. Por ejemplo, en su vecindad se presentan unidades como “inability”, “slowed”, “loss” o “disastrous”.

The biggest mistakes included Merrill's large investment in COLLATERALIZED DEBT OBLIGATIONS and subprime lending and the company's inability to sell when the markets slowed.

That news came as a shock to many of the firm's executives, who were distressed that Morgan Stanley, which was not a major underwriter of COLLATERALIZED DEBT OBLIGATIONS, would record such a large trading loss.

The strategy produced substantial profits for a time, but also resulted in a complex and ultimately disastrous trade in COLLATERALIZED DEBT OBLIGATIONS earlier this year that led to the \$3.7 billion write-down.

For example, at the end of September, the intermediate fund in which Ms. Ceaser-Hudson invested had almost 17 percent of its assets in mortgage-related securities – including COLLATERALIZED MORTGAGE OBLIGATIONS, home equity loans and pools of mortgages that were again combined into collateralized debt obligations.

Sean Egan, the co-founder of Egan-Jones, an independent bond rater, believes that MBIA and the other big bond insurers will be saddled with billions of dollars in losses as COLLATERALIZED DEBT OBLIGATIONS stuffed with subprime debt – so-called C.D.O.'s – they have insured continue to go south.

Dentro de las colocaciones más llamativas en este intervalo está el uso del “deal”, en cuanto a frecuencia y polaridad, como colocación de “debt”. Este término a priori se ha marcado como un elemento de sentimiento positivo ya que la existencia de acuerdos permite el flujo económico. Sin embargo, aquí se puede observar en los ejemplos que el carácter de las enunciaciones es eminentemente factual. “Debt” aquí, no es más que uno de los factores de compraventa de activos entre grandes compañías.

With the assumption of \$15.3 billion in DEBT , the DEAL would have a value of \$38.3 billion, Equity Office said in yesterday's announcement.

Including \$1.7 billion of assumed DEBT , the DEAL is worth \$6.8 billion.

Backed by Royal Bank of Scotland, Mr Davies' team had acquired the business from the trusts for £30m and £12m of DEBT in 2005.

Including DEBT the DEAL is worth some \$168m.

But the various DEALS left with huge DEBTS, and it was forced to turn to its shareholders for more money through a heavily discounted £800m rights issue in 2002.

Rio Tinto will substantially increase its DEBT with the DEAL.

### Segundo intervalo de estudio

Durante la segunda etapa de análisis (2008-2011), el contexto de “debt” se caracteriza por tener más coocurrencias con unidades del dominio específico y sentimiento generalmente negativo como “default”, “burden” o “crisis” y adjetivos como “bad”. Al contrario que en el caso de “credit” con p. ej., “card”, se encuentran escasas colocaciones estables que prevalezcan durante este intervalo. En este caso se habla de “level”, “pay” elementos con bajo nivel de especialización y con un carácter eminentemente neutro y mantienen niveles estables en el coeficiente *logDice*. El único elemento que se repite a lo largo de los cuatro años de enfoque es el término “obligation” que ya fue analizado previamente.

Este segundo periodo está marcado por la explosión del uso del término de estudio en el corpus. Además, en comparación con el intervalo anterior, este intervalo de estudio se caracteriza por el paso de referenciar la deuda en segmentos con un sentimiento indistinto a la presentación de una mayor cantidad de segmentos evaluativos de carácter negativo.

A partir de los eventos históricos que a partir de 2008 ponen en el centro del discurso económico a la deuda, emerge un nuevo conjunto léxico alrededor de “debt” que, por un lado, contextualiza de manera más negativa el uso de la palabra y, dada la situación, pasa a convertirse en una palabra de uso frecuente en el discurso de las noticias de carácter económico. Como se observa en la Tabla 24 las palabras del vecindario del término durante esta etapa variaron significativamente, relacionándose con unidades que ilustran de manera clara las principales preocupaciones del periodo. Se comienza entonces a hacer frecuente el uso de “bad debt” ‘deuda mala’ o ‘deuda incobrable’, “debt burden” y se habla del peligro de entrar en “default” y de la necesidad de refinanciar (“refinance”) y reestructurar (“restructure”) la deuda. A finales de este intervalo, tras los problemas de los estados para acceder al crédito, se empieza a hablar de los problemas de deuda soberana de los estados (“sovereign debt”, “national debt”). En términos

ilustrativos, en 2013<sup>56</sup> la deuda de Grecia en comparación con su PIB ascendía al 161%, Italia un 126%, Portugal un 120%, Francia un 90%, el Reino Unido un 89% y España un 85% en comparación con una media mundial del 64%. La última colocación relevante en aparecer aquí es el término “debt ceiling” ‘techo de deuda pública’, que tuvo una relevancia ideológica y política capital especialmente en Europa durante estos años. Se puede afirmar entonces que la deuda dejó de referenciarse de manera cuasi abstracta y en términos eminentemente macroeconómicos para materializarse en una problemática que afecta a las economías y por extensión a las relaciones humanas ya sean entre países o instituciones.

A continuación, se detallan las colocaciones clasificadas por categorías gramaticales entre 2008 y 2011, junto al número de subcorpus anuales en los que coocurre con “debt” entre paréntesis:

- Sustantivos genéricos: *amount, card* (2), *collector, country, crisis* (2), *euro, eurozone, government, level* (4), *problem, ratio, repayment*.
- Sustantivos específicos: *bank, burden* (3), *cash, collection* (2), *ceiling, credit* (2), *default* (2), *equity* (2), *GDP, load* (2), *loan, mortgage, obligation* (4), *payment* (2), *restructuring* (3).
- Adjetivos genéricos: *high*.
- Adjetivos específicos: *bad* (2), *collateralised* (2), *mortgage-backed, national* (3), *net* (2), *public, short-term, sovereign* (2).
- Verbos genéricos: *pay*, (4), *raise, reduce* (3).
- Verbos específicos: *default, downgrade, issue, refinance* (3), *repay, restructure* (2).

---

<sup>56</sup> Según el CIA World Factbook



2008			2009		
Palabra	Cooc	logDice	Palabra	Cooc	logDice
obligation-n	27	10,354	<b>bad-j</b>	29	9,251
collateralized-j	14	9,542	pay-v	51	9,237
<b>bad-j</b>	28	9,458	restructuring-n	17	9,033
card-n	20	9,385	<b>burden-n</b>	15	9,024
credit-n	31	9,161	restructure-v	15	9,022
pay-v	36	9,076	credit-n	28	8,942
<b>refinance-v</b>	11	9,058	obligation-n	14	8,891
short-term-j	10	8,899	payment-n	20	8,859
payment-n	14	8,864	load-n	12	8,755
loan-n	21	8,699	<b>reduce-v</b>	19	8,704
<b>default-v</b>	8	8,693	<b>repayment-n</b>	11	8,606
repay-v	8	8,642	<b>refinance-v</b>	11	8,603
collection-n	8	8,565	card-n	14	8,545
level-n	13	8,555	equity-n	15	8,534
amount-n	11	8,546	total-j	13	8,520
mortgage-n	18	8,521	cash-n	16	8,407
<b>load-n</b>	7	8,460	collection-n	10	8,364
issue-v	9	8,448	level-n	14	8,307
mortgage-backed-j	7	8,447	national-j	12	8,299
borrower-n	8	8,332	buy-v	21	8,162
2010			2011		
Palabra	Cooc	logDice	Palabra	Cooc	logDice
<b>crisis-n</b>	76	10,226	<b>crisis-n</b>	219	11,398
sovereign-j	43	10,17	sovereign-j	105	11,114
obligation-n	31	9,692	eurozone-n	48	9,668
level-n	45	9,621	<b>burden-n</b>	31	9,428
pay-v	62	9,329	ceiling-n	28	9,327
<b>reduce-v</b>	33	9,175	pay-v	58	9,072
<b>burden-n</b>	21	9,168	government-n	72	9,048
refinance-v	19	9,044	level-n	39	9,037
public-j	35	8,993	<b>reduce-v</b>	39	9,020
net-j	20	8,945	obligation-n	23	8,981
ratio-n	17	8,831	restructuring-n	23	8,966
restructure-v	16	8,805	national-j	24	8,696
collateralised-j	15	8,744	<b>default-v</b>	20	8,675
collector-n	15	8,697	country-n	44	8,624
equity-n	21	8,656	euro-n	22	8,604
national-j	19	8,633	<b>problem-n</b>	29	8,550
restructuring-n	15	8,619	bank-n	54	8,490
<b>default-v</b>	14	8,611	raise-v	25	8,387
<b>downgrade-v</b>	14	8,596	GDP-n	16	8,369
<b>high-j</b>	33	8,562	net-j	16	8,363

**Tabla 24** Principales colocaciones de “debt” durante el segundo intervalo

El adjetivo “bad” lidera el ranking de en las colocaciones de 2008 y 2009 (*logDice* 9,458 y 9,251 en 2009 y 2009). La “deuda mala” o “incobrable” es aquella que se debe a un acreedor existiendo retrasos o escasas probabilidades de cancelarse y en la cual los acreedores no disponen a tomar medidas para su cobro principalmente, debido a que el deudor no tiene el dinero para pagar por insolvencia.

Como puede observarse en el corpus, ya desde finales de 2007, los principales bancos europeos como el suizo UBS o los británicos Lloyds TSB y HSBC informan en sus cierres anuales que sufren de acumulaciones catastróficas de deuda procedente de los fondos *subprime* estadounidenses afectados por los impagos y retrasos masivos. Este clima de desconfianza y negatividad de las opiniones alrededor de “debt” se puede observar claramente en los ejemplos aleatorios del subcorpus de 2008-9 que se verá a continuación. En él se puede observar que “bad debt” se relaciona con previsiones económicas desastrosas. Estas están relatadas con verbos en tiempos gramaticales relacionados con el presente y el futuro, lo que indica la inmediatez del fenómeno. El léxico con el que el término de estudio coocurre suele tener también una polaridad muy negativa. A través de la observación se encuentran sustantivos que metaforizan incendios y enfermedades (“conflagration” o “epidemic”) o verbos de naturaleza semántica similar como “suffer”, “devastate” o “ignite”.

But Monument turned into the worst-performing unit within Barclaycard, as customers struggled to repay their balances and BAD DEBTS jumped.

Barclays and HSBC revealed substantial increases in BAD consumer DEBTS in the first half of the year.

An epidemic of unpaid BAD DEBTS would devastate lenders and ignite a conflagration that could pull down the economic and financial structure, ruining everybody.

HSBC has suffered another hit from the financial crisis after BAD DEBTS in the US soared by more than the City had feared.

The percentage of borrowers delinquent on payments and BAD DEBTS will rise as the year goes on, noted Keith.

It blamed the "rapidly deteriorating" economic climate for a rise in BAD DEBTS , a trend it expects to continue - especially on the property side as more developers are hit by falling house prices and a lack of buyers.

Banks also know that BAD DEBT is not usually at the hands of their own customers.

La siguiente colocación destaca en este periodo es “burden” (carga de deuda) y ganó notoriedad en los rankings de colocaciones en este periodo (con unos coeficientes estables *logDice* de 9,024 en 2009, 9,168 en 2010, 9,428 en 2011). “Debt burden” se define como “a large amount of money that one country or organization owes to another and which they find very difficult to repay” (Collins Cobuild, 2019). Este uso extensivo del

término “debt”, a diferencia de las referencias a la deuda estándar, se usa en contextos donde la suma de dinero adeudada es lo suficientemente considerable como para causar problemas de impago para el deudor. Este término refleja un uso de la terminología que dista de los planteamientos iniciales que proponen que los términos son meras relaciones factuales y asépticas entre concepto y denominación pues su propia definición implica una connotación negativa.

Tras observar la prosodia semántica, el término se relaciona con unidades léxicas con un carácter también negativo, como se puede observar en los ejemplos aleatorios ofrecidos por el corpus el término está relacionado léxicamente con unidades de sentimiento abiertamente negativo como es el caso de verbos como “struggle” o “hampered”, sustantivos como “load” y “difficulties”, verbos como “struggle”, adjetivos como “heavy”. Algunos ejemplos aleatorios en el subcorpus 2008-2011 pueden observarse a continuación:

It denied that the decision was influenced by the spectre of BAA's £9.3bn DEBT BURDEN , which the airport owner is struggling to finance.

The unusually heavy DEBT BURDEN means Tribune's bankruptcy is not a harbinger for the newspaper industry, said Rick Edmonds, a media business analyst at the Poynter Institute.

"Our country does not need just another spending bill, particularly not one that will load future generations with the BURDEN of massive DEBT ," he said.

Smaller lenders stand to lose £200m from move McCarthy & Stone, Britain's biggest retirement home builder, is facing possible administration because lenders are struggling to agree how to tackle its £900m DEBT BURDEN .

The company has been struggling with its DEBT BURDEN since June and its prospects have been hampered by the difficulties retirees are having in selling their homes in a plunging market in order to buy McCarthy & Stone properties.

La unidad léxica “default” (con un coeficiente *logDice* 8,693 en 2009, 8,6911 en 2011, 8,6975 en 2012) referencia al incumplimiento de las obligaciones (o condiciones) legales de un préstamo, por ejemplo, cuando un comprador de vivienda no realiza el pago de una hipoteca, o cuando una corporación o gobierno no paga un bono que ha alcanzado el vencimiento. Esta definición se entiende como negativa en su uso contextual.

Durante esta crisis se producen los mayores incumplimientos de la historia tras la quiebra de Lehman Brothers en 2008 como incumplimiento privado, y la bancarrota de Grecia, que supuso en marzo de 2012 el mayor incumplimiento soberano. Durante este periodo, no fueron pocas las empresas e instituciones que tras la sequía de crédito fueron víctimas del efecto dominó e incapaces de afrontar los problemas de liquidez tras el colapso del sistema crediticio.

Entre los elementos que constituyen su prosodia semántica se encuentran, en su gran mayoría, unidades léxicas de carácter negativo como los sustantivos “precipice”, adjetivos como “hard” o “subordinated” o sustantivos como “falls” o “risk”.

Mr. Fuld was referring to investors' panic after Russia DEFAULTED on its DEBT, pushing a big hedge fund to the precipice.

There was only light trading of credit-default swaps, which allow investors to protect against the risk of a company's DEFAULTING on its DEBT , and those trades have been a leading indicator about the health of banks.

Moreover, Mr. James said, about 60 percent of those companies have loans with relatively few restrictions, making it harder for these holdings to DEFAULT on their DEBT .

Tribune's papers were down 19 percent in the third quarter – and some major newspapers have DEFAULTED on DEBT or been put up for sale, with no takers.

Worldwide, there are about \$55 trillion of these contracts, which enable people to make side bets on whether borrowers will DEFAULT on their DEBTS .

They either DEFAULT on the DEBT or take bolder measures.

How can investors trust something issued by West Brom when the nationalised Bradford & Bingley is DEFAULTING on its subordinated DEBT?

After suffering sharp falls last week, world stock markets have now recovered their poise, with investors less concerned that Dubai World might DEFAULT on its DEBTS and spark global turmoil.

Asimismo, durante la última parte de este periodo de estudio “debt” comenzó a relacionarse de una manera íntima y recurrente con “crisis”, elemento que acabó convirtiéndose en su colocación principal en los años 2010 y 2011 (un coeficiente *logDice* de 10,226 en 2010 y 11,389 en 2011). Para este entonces ya había explotado una segunda crisis: la crisis de la deuda en la Zona Euro. En muchos de sus países miembros, las deudas privadas derivadas de la burbuja de la vivienda se transformaron, como se verá un poco más adelante, en deuda soberana como resultado de los rescates del sistema bancario y las políticas de austeridad de los gobiernos. Los efectos de la crisis de la deuda y el desconcierto que supuso su crecimiento en un corto espacio de tiempo se reflejan claramente en el discurso.

En una pequeña muestra como la extraída del subcorpus 2010-2011 se pueden observar metáforas como “under assault”, sustantivos como “victim”, “pressure”, “protests” o “cuts” en referencia a los países de la Zona Euro, principalmente Grecia. Finalmente, a través de la muestra puede observarse la sensación de ser un fenómeno fuera de control a través de verbos como “snowballing” o “sucked into”.

Share prices in Europe had slid on Wednesday when France was sucked into the eurozone DEBT CRISIS .

Greece's place in the list is perhaps the most assured as its DEBT CRISIS continues and it comes under assault from speculators.

A deal has been reached to help Greece tackle its DEBT CRISIS, after negotiations between Europe's leaders in Brussels this morning.

Athens could announce further austerity measures to tackle its DEBT CRISIS next week, under intense pressure from financial markets, despite widespread protests against existing spending cuts.

The multi-millionaire private equity boss at the centre of the DEBT CRISIS at music group EMI has spelled out the lengths to which he is going to avoid UK tax.

Greece's DEBT CRISIS has dominated headlines since the socialists revealed the country's public deficit would reach 12.7% of GDP – nearly twice that predicted by the former conservative government – within weeks of taking office last October.

Sitting in the sun-dappled square where Seixal's old town meets the River Tagus, Umberto Da Silva was clear about who would suffer as Portugal became the latest victim of Europe's snowballing sovereign DEBT CRISIS.

La unidad “debt ceiling” ‘techo de deuda’ (*logDice* 29,327 en 2011) es un término que referencia a los límites del déficit, generalmente constitucionales, cuyo objetivo es limitar la posibilidad de un gobierno gaste más de lo que recauda, limitando así el crecimiento de la deuda pública para garantizar la sostenibilidad de las finanzas públicas a largo plazo. Más allá de entrar a discutir sobre la adecuación de este tipo de medidas y los intereses a los que responden, la cuestión del techo de deuda fue una de las principales crisis políticas que tuvieron lugar en 2011 tanto en Estados Unidos como en Europa.

En el Congreso de los Estados Unidos, se planteó un debate sobre el aumento del techo de deuda, cuando lo común es una aprobación rutinaria libre de discusión partidista. El Partido Republicano, líder en la cámara, exigió la negociación de la reducción del déficit a cambio de un aumento en el techo de la deuda. Si el gobierno no aumentaba inmediatamente su techo de deuda, podría comenzar a realizar impago en los bonos. Es probable que esta situación condujese a una nueva oleada en la crisis financiera internacional. Finalmente, el 31 de julio, a dos días del vencimiento de los plazos se llegó a un acuerdo parcial que desbloqueó el techo de gasto. La crisis provocó la semana más volátil para los mercados desde la crisis de 2008 y la calificación de los bonos estadounidenses se desplomó. Esta crisis no se resolvió permanentemente y en 2013 tuvo lugar una nueva crisis, la cual queda reflejada en el ranking de colocaciones de ese año.

En marzo 2011, Europa propuso una nueva reforma del *Pacto de Estabilidad y Crecimiento* orientado a sanciones automáticas para los que violan los parámetros relativos al 3% de déficit / PIB al 60 %. Angela Merkel insistió en que la Comisión Europea y el Tribunal de Justicia jueguen un papel importante en el control de garantías de cumplimiento de las obligaciones soberanas. En 2011 de Alemania, el Francia y otros

países más pequeños de la Unión Europea dieron un paso más hacia la unión fiscal de zona euro con las normas presupuestarias muy estrictas y sanciones automáticas para aquellos que no las respeten. En diciembre de 2011, en el contexto del *Pacto Fiscal Europeo* promovido por Alemania, los miembros de la zona euro acordaron regular parámetros relativos al “techo de gasto” y la introducción de sanciones automáticas para quienes los violen. En el caso de España, esto supuso una reforma constitucional exprés en pleno mes de agosto. Este causó una crisis de legitimidad en un clima de pérdida de legitimidad de la clase política tras un ambiente generalizado de protestas en el país.

Observando el corpus, el sentimiento alrededor de la muestra obtenida en el corpus a partir de la búsqueda de “debt ceiling” es negativa tanto para los Estados Unidos como para Europa. En su entorno puede observarse que el término está rodeado de elementos negativos (p. ej., “battle”, “downgrade”, “risk”, “miscreants”, etc.). Algunos ejemplos a continuación:

The data comes at a time when the White House and Congress are locked in a battle over how to raise the DEBT CEILING to allow the government to borrow more money.

The Starbucks boss said he knew he had to speak out after watching Democrats and Republicans slug it out over raising the DEBT CEILING , an argument that contributed to ratings agency Standard & Poor's decision to downgrade US debt.

Obama accused the Republicans of using the raising of the DEBT CEILING , which he said was normally routine, as a lever to force the Democrats to agree to deep spending cuts that would hit the poorest in society while leaving the wealthiest unscathed.

There was widespread alarm in Washington this month when S & P, followed by Moody's and Fitch , other credit rating concerns, warned that the soaring federal debt and the political standoff over raising the DEBT CEILING had placed the nation's credit rating at risk.

In return for the German concession, Van Rompuy was instructed to come up with proposals by the end of the year to strengthen the stability pact which sets DEBT and deficit CEILINGS for the eurozone and introduce new sanctions for fiscal miscreants.

" The agreement, in the wake of a nasty and protracted battle over the DEBT CEILING , was worked out as both the White House and Congress were beginning to feel pressure from voters who said they have grown tired of political fights that hurt working Americans and the economy.

Como enlace con la siguiente etapa se analizará el contexto en el que aparecen las colocaciones “sovereign” y su cuasi equivalente “national”. Ambas hacen mención de la emergencia de la crisis de la deuda soberana que afectó principalmente a Europa. En el año 2010 se convierten en la segunda colocación más significativa de la lista (*logDice* 10,17 en 2010 y 11,114 en 2011) y se mantiene hasta el año 2015.

La explosión de la crisis soberana tras la crisis griega y su posterior contagio se convirtió en un nuevo episodio de la crisis global, el cual se revisó con detalle en el Capítulo 2. Como se puede observar en los ejemplos extraídos del subcorpus 2010-2011,

la colocación más inmediata de “sovereign debt” es “crisis”. A partir de este año, este término pasó de ser un elemento poco frecuente en el discurso que hacía referencia a un indicador macroeconómico a ser uno de los términos clave del dominio en esta época. En la gran mayoría de los casos observados, las referencias a la deuda soberana se realiza en un entorno léxico y contextual con sentimiento negativo.

With the euro zone facing a SOVEREIGN DEBT crisis that has much further to unravel, there were already signs in August that the UK's exports to key trading partners France, Germany and the Netherlands were suffering.

Tony Tyler, director general of Iata, said: "On the good news side, it appears that a worsening of Europe's SOVEREIGN DEBT crisis has been avoided for now.

Banco Santander, the largest bank in Europe, moved today to assuage fears of a looming SOVEREIGN DEBT crisis in Spain after Standard and Poor's (S&P) cut the country's credit rating.

Because the ratings agencies use slightly different methodologies, there is no single list of AAA-rated SOVEREIGN DEBT .

Gurría's comments came as interest rates on Italian SOVEREIGN DEBT rose back above 6% and the gap between Italian and German bond yields ballooned to the widest level since the single currency was created.

Quoting a document which it said was prepared by the German finance ministry, Der Spiegel said that a new Greek currency could lose as much 50% of its value if Greece pulls out of the eurozone, leading to an explosion in Greek NATIONAL DEBT and crippling its banking system.

Italy plans to sell off its national treasures; Greece is said to be keen to offload some surplus bits of islands ; now the rightwing Crawley-based Thomas Malthus Institute, under its president Stan Muccia, has drawn up a draconian austerity package aimed at wiping out Britain's £800bn NATIONAL DEBT .

In Europe, the various NATIONAL DEBT crises remain unresolved, with a continued monopoly of banker-friendly austerity programs, and their predictable consequences of rising unemployment and stagnation.

### Tercer intervalo de estudio

Finalmente, se analizarán las colocaciones del tercer y último periodo de análisis, relativo al intervalo entre los años 2012 y 2015. Si bien la tendencia de sentimiento es levemente ascendente dentro del umbral de la negatividad (pasando de 38,47 puntos en 2011 a 43,94 en 2015), el contexto de “debt” también coocurre con unidades del dominio específico y sentimiento generalmente negativos sin grandes variaciones con el periodo anterior. Este hecho es consecuente ya que, aunque surgen medidas de alivio, la situación de la deuda no mejora en Europa. Por ejemplo: las colocaciones con “crisis” o “burden” se mantienen estables durante este periodo. En todo caso, la novedad reside en el ranking aquí presenta también nuevas unidades con sentimiento positivo como es el caso de “relief” o “reduce” en 2014 y 2015. A continuación, se ofrece una lista de colocaciones clasificadas por categoría gramatical y la Tabla 25, clasificada por año:



- Sustantivos genéricos: *country* (3), *crisis* (4), *government* (3), *level*, (4), *pile*, *problem*, *student*.
- Sustantivos específicos: *bond*, *burden* (4), *ceiling*, *credit*, *default* (2), *euro*, *eurozone*, *growth*, *household* (2), *loan*, *mortgage*, *rate*, *ratio*, *relief* (2), *repayment*, *restructuring* (4).
- Adjetivos genéricos: *high* (4), *large*, *unsustainable*.
- Adjetivos específicos: *bad* (3), *national* (3), *net*, *public* (3), *sovereign* (4).
- Verbos genéricos: *buy*, *finance* (2), *owe*, *pay* (2), *raise*, *reduce* (4).
- Verbos específicos: *repay* (2), *restructure* (3)

La colocación “debt burden” se mantiene estable en este último periodo de estudio (*logDice* de 9,659 en 2012, 9,991 en 2013, 9,352 en 2014, 9,53 en 2015). Si bien algunos ejemplos son retrospectivos, en él se hace referencia a la inmediatez de los efectos macroeconómicos de una crisis continuada de la deuda que agrava la situación de la economía.

Entre los objetos de opinión más destacados se encuentra, de manera destacada, Grecia. El estado helénico fue un actor principal en la crisis de la zona euro, al entrar en una profunda recesión plegarse su economía una cuarta parte durante este periodo. Los debates sobre las condiciones de rescate a este estado se acrecentaron en 2015 en una situación de enfrentamiento con la UE tras el giro a la izquierda del nuevo gobierno y una solicitud de renegociación de las condiciones de rescate que culminó con el referéndum sobre la gestión económica y las condiciones de un nuevo rescate en julio de 2015. La carga de deuda de la Unión Europea fue considerada como inasumible.

El siguiente término de este periodo es “bad debt,” el cual vuelve a aparecer significativamente en el discurso en los años 2012, 2014 y 2015 con unas cifras similares a las observadas durante el bienio 2008-2009 (*logDice* de 8,444 en 2012, 8,891 en 2014 y 9,21 en 2015). Como se puede observar en los ejemplos, el uso de “bad debt” referencia a un grave problema enquistado desde el periodo de estudio anterior (p. ej., referencias a 2009) que impide el avance de la economía financiera.

Como se puede observar también en los ejemplos aleatorios extraídos del corpus, esta sensación de acumulación se puede observar a partir de colocaciones metafóricas como “mountain” o “sitting”. Las connotaciones negativas se pueden observar a partir del uso de colocaciones como “concerned”, “rising” o “starved”. Los ejemplos son los siguientes:



2012			2013		
Palabra	Cooc	logDice	Palabra	Cooc	logDice
<b>crisis-n</b>	122	10,644	ceiling-n	58	10,774
sovereign-j	58	10,424	<b>burden-n</b>	35	9,991
<b>burden-n</b>	34	9,659	<b>crisis-n</b>	57	9,835
level-n	41	9,250	sovereign-j	26	9,675
government-n	77	9,196	restructuring-n	25	9,489
pay-v	60	9,157	level-n	37	9,049
eurozone-n	31	9,099	raise-v	32	8,894
<b>reduce-v</b>	35	9,039	government-n	61	8,797
country-n	58	9,011	<b>reduce-v</b>	30	8,712
restructuring-n	21	8,942	public-j	33	8,708
national-j	25	8,896	<b>high-j</b>	37	8,327
pile-n	18	8,823	growth-n	29	8,187
euro-n	24	8,711	pay-v	35	8,167
problem-n	31	8,628	country-n	32	7,984
buy-v	32	8,576	restructure-v	6	7,544
finance-v	16	8,507	finance-v	7	7,541
<b>high-j</b>	37	8,502	mortgage-n	8	7,536
household-n	17	8,456	credit-n	10	7,534
<b>bad-j</b>	21	8,444	large-j	16	7,429
<b>default-n</b>	15	8,443	rate-n	15	7,402
2014			2015		
Palabra	Cooc	logDice	Palabra	Cooc	logDice
sovereign-j	27	9,600	<b>crisis-n</b>	136	10,100
restructuring-n	26	9,437	<b>relief-n</b>	121	10,009
<b>burden-n</b>	24	9,352	sovereign-j	103	10,093
<b>crisis-n</b>	43	9,242	restructuring-n	49	9,625
pay-v	70	9,096	<b>burden-n</b>	49	9,530
default-n	18	8,936	<b>reduce-v</b>	47	9,512
<b>bad-j</b>	29	8,891	net-j	66	9,263
household-n	24	8,887	repayment-n	45	9,232
level-n	37	8,842	<b>bad-j</b>	37	9,210
repay-v	16	8,823	government-n	46	8,901
student-n	22	8,808	billion-m	102	8,896
<b>reduce-v</b>	33	8,696	level-n	58	8,884
public-j	30	8,594	public-j	48	8,725
owe-v	14	8,566	ratio-n	47	8,706
restructure-v	13	8,56	restructure-v	26	8,686
national-j	20	8,554	repay-v	30	8,630
<b>relief-n</b>	13	8,498	national-j	32	8,577
loan-n	20	8,416	<b>high-j</b>	63	8,547
bond-n	16	8,340	<b>unsustainable-j</b>	22	8,503
high-j	40	8,274	country-n	61	8,450

**Tabla 25** Principales colocaciones de “debt” durante el tercer intervalo

Piers Haben, director of oversight at the EBA, said that while the resilience of the financial sector was improving because more capital was being accumulated in banks, he remained concerned about BAD DEBTS.

Regulators told the bank to strengthen its finances against rising BAD DEBTS , many inherited from the Britannia Building Society, which the bank acquired in 2009.

To be able to raise new capital, they might need to come clean about the size of their BAD DEBTS , the FPC said while acknowledging the difficulties associated with raising more capital and trying to lend to businesses and householders.

Cummings's division was the subject of an investigation by the Financial Services Authority after the enlarged bank – renamed Lloyds Banking Group – was forced to more than double its BAD DEBT provisions from £3.3bn to £7bn.

European banks are sitting on BAD DEBTS of €1tn – the equivalent to the GDP of Spain – which is holding back their profitability and ability to lend to high street customers and businesses.

He will begin with the merger with Britannia Building Society at the height of the financial crisis in 2009, which has left the enlarged bank with a mountain of BAD DEBTS and starved of capital.

European banks sitting on €1tn mountain of BAD DEBT , survey finds Analysis of 105 banks by European Banking Authority raises concerns over 'drag on profitability' from non-performing loans

La colocación “debt crisis” también permanece estable en este último periodo de estudio (*logDice* de 10,644 en 2012, 9,835 en 2013, 9,242 en 2014 y 10,100 en 2015). La permanencia de la colocación en las posiciones principales del ranking se justifica por la inmediatez de los efectos de la crisis de la deuda, devenidos en una crisis en la Zona Euro durante los años en los que transcurre este periodo.

Al igual que en el caso de “burden”, se observa que Grecia (“Greece”, “Greek”, “Athens”) y la Zona Euro están entre los principales objetos de opinión. En los ejemplos se observa la importancia que los efectos de esta crisis tienen no solo sobre los indicadores macroeconómicos, sino sobre sus efectos negativos que recaen inmediatamente sobre la vida cotidiana (“rising living standards”, “slump in demand”, “public spending cuts”...) de la ciudadanía. Al contrario que el uso de “credit crisis” durante este mismo intervalo, esta es mencionada a través del uso de tiempos verbales en presente pues es referenciada como un hecho inmediato de carácter prolongado. Por ejemplo:

Alongside fading worries about the Greek DEBT CRISIS and signs of rising living standards, any such recovery in headline growth could fan expectations that the Bank's policymakers are readying to raise interest rates, perhaps even before the end of the year.

Esprit, which generates about 80 percent of its sales in Europe, has been trying to restructure its business under the watch of Mr. Körber and Mr. van der Vis as it confronts a slump in demand because of the euro zone DEBT CRISIS.

Record world DEBT could trigger new financial CRISIS, Geneva report warns Concerted effort required to tackle economic woes as slow growth and low inflation cause global debts to balloon.

Poul Thomsen, a senior International Monetary Fund official who oversees the organisation's mission in Greece , also insists that, contrary to popular belief, Athens has achieved a lot since the eruption of the DEBT CRISIS in December 2009.

The chief economist role at the IMF is a crucial post and Obstfeld's appointment has been announced at a time when the Washington-based fund is heavily involved in the Greek DEBT CRISIS .

The sovereign DEBT CRISIS in the eurozone has seen a series of countries, including Greece and Portugal , accept swingeing public spending cuts, mass privatisations and drastic market reforms, as the price of receiving bailouts from the IMF and their European neighbours.

La coocurrencia del verbo “reduce” con “debt” durante esta época tiene continuidad con las frecuencias observadas en el intervalo anterior (*logDice* 9,039 en 2012, 8,712 en 2013, 8,646 en 2014 y 9,512 en 2015). Las coocurrencias de “reduce” y “debt” tienen, en general, connotaciones positivas. Sin embargo, como se puede observar en los ejemplos, en la mayoría de los casos se utiliza como un supuesto para la mejora de la economía (mediante el uso de la modalidad: “could”, “should”). Entre los ejemplos extraídos, se pueden hallar otras fórmulas como “to reduce the debt...” que se usan a modo de ingrediente en una fórmula que permita volver a la senda del crecimiento en una situación económica muy adversa:

Shaun Donovan, the housing secretary, has said the settlement will be a catalyst that proves that REDUCING mortgage DEBT is cost-effective for lenders.

These changes could also REDUCE household DEBT .

DEBTS should be " REDUCED organically through growth, or opportunistically when less distortionary sources of revenue are available", the IMF's researchers argue.

No dividends have been paid since the 2007 deal with spare cash used to REDUCE its DEBT , which at the last count stood at £7bn.

To REDUCE the DEBT to 124% by 2020, the ministers were putting together a package of steps including a debt buyback funded by a eurozone rescue fund, reducing the interest rate on loans and returning eurozone central bank 'profits' back to Greece.

"This year, I think, there are three priorities: to stabilize economic growth, which is not too big of a problem," to stabilize the prices of goods, "where already it looks like there could be some pressure," and to REDUCE the risk from hidden DEBT , like off-book wealth management products, said Zhao Xijun, deputy director of the Finance and Securities Institute at Renmin University in Beijing.

Como último elemento de estudio en este apartado, la emergencia del sustantivo “relief” comienza a destacar en 2014 y ascendió a la segunda posición en el ranking en 2015 (*logDice* 8,498 y 10,009 respectivamente). Este sustantivo es utilizado de manera metafórica como un conjunto indeterminado de medidas de “alivio” al “dolor” que supone la carga de la deuda sobre la economía, principalmente en el caso griego. “Relief” es utilizado en este intervalo como algo no ofrecido por los poderes económicos sino demandado por quienes sufren la crisis. Esto se evidencia en el uso de sintagmas verbales relacionados con la confrontación (“press”, “asking for” o “pushing for”) cuyos agentes (“cash-poor Ukraine”, “left politicians”...) suelen estar subordinados de manera antagónica a los grandes poderes (“International Monetary Fund”, p. ej.).

En plena crisis del euro todas las facciones (la *Troika*, el gobierno griego, las autoridades en economía) solo coincidían en una cosa: la deuda griega es insostenible para todo el bloque europeo y necesitaba algún tipo de medidas (“measures”, “agreement”, “concessions”) de alivio a las que se accedería mediante negociación. A pesar de la negativa de la mayoría del pueblo griego en referéndum a una nueva intervención, esto se tradujo en un acuerdo entre el gobierno de Tsipras y la *Troika* a un tercer rescate. En el caso presente se esperaría una prosodia semántica con connotaciones positivas en el sentimiento. Sin embargo, los ejemplos observados a través del corpus aquí a través de las coocurrencias en el léxico están relacionados con elementos neutros o negativos.

The International Monetary Fund has said it will refuse to take part in the bailout unless there is an "explicit and concrete" agreement on DEBT RELIEF for Greece.

But eurozone governments have so far resisted substantial DEBT RELIEF and are implacably opposed to any measure that could write off some Greek debts, otherwise known as a "haircut".

Politicians instructed by the radical left prime minister to fight Athens's corner, in what is being billed as the very last chance to cut a deal after five months of fruitless talks, will press for DEBT RELIEF in exchange for concessions demanded by international creditors.

It is asking for DEBT RELIEF under the terms of the November 2012 bailout agreement, and it wants to be able to take steps to deal with the humanitarian crisis caused by the 25% collapse in the size of the economy over the past five years.

After the power change, the cash-poor Ukraine started pushing for DEBT RELIEF .

Even excluding DEBT RELIEF , however, financial assistance to the least-developed countries fell in 2014, the OECD said.

As chairman of the International Institute of Finance – an industry group that represents the interests of large investment banks like Deutsche, Goldman Sachs and Morgan Stanley – he played a primary role in negotiating an agreement that provided crucial DEBT RELIEF for Greece.

### 7.3 Término 3: *Inflation*

El término “inflation” es una unidad perteneciente al dominio económico que se define como “a sustained increase in the general price level of goods and services in an economy over a period of time” (Barro, 1997, pg. 895)<sup>57</sup>. La tasa de inflación es un indicador macroeconómico que observa el cambio porcentual anualizado en un índice general de precios a lo largo de una serie temporal. Al igual que ocurre con “debt”, se habla aquí de

---

<sup>57</sup> En español: “Un aumento sostenido en el nivel general de precios de bienes y servicios en una economía durante un período de tiempo”.

un indicador *potencialmente negativo* pues las expresiones que indiquen el aumento de los valores de inflación se considerarán que tienen un sentimiento negativo y viceversa. Cuando el término es referenciado de manera aislada, este no indica un sentimiento claro y depende del contexto, por lo que está etiquetado como neutro en *LexiEcon* (Figura 18). Al ser un indicador potencialmente negativo, todas las entradas relacionadas con “inflation” están recogidas como tal.

inflation	ALL	neu	2
<decline>_3_inflation	VB	pos	1
<fuel>_3_inflation	VB	neg	3
<high>_3_inflation	JJ	neg	3

**Figura 18** Ejemplo de entradas relacionadas con “inflation” en *LexiEcon*

Como elemento de estudio, el análisis del contexto discursivo del término se justifica en que este fue un indicador especialmente destacado en el contexto europeo durante los años de la crisis del euro. Uno de los objetivos naturales del BCE es custodiar la estabilidad del euro y mantener la inflación bajo control. Paradójicamente esto se realiza a nivel global sin tener en cuenta las distintas particularidades de los estados miembros, entre sus competencias no está el mandato directo con respecto a ninguna economía de la Zona Euro, aunque la política económica de los países queda limitada al no poder emitir divisa para pagar a sus tenedores de deuda. A diferencia del mandato dual de la Reserva Federal, el Banco Central Europeo tiene como ordenanza únicamente el control de la inflación sin poder hacer regulaciones en el empleo.

El problema de la inflación se comenzó a hacer patente cuando la crisis se desarrolló en 2010 con el aumento en los déficits presupuestarios y con la proximidad de los vencimientos de deuda, los bonos emitidos se dispararon. Durante los años 2010-2014, dado que para muchos de los países del norte de Europa una inflación baja era beneficiosa, no se implementaron políticas que aumentasen la inflación, y por ende estimulasen los precios. Por ello, países del sur como Grecia se encontraron sin formas de ajustar su economía y aliviar la crisis.

Con un riesgo de caer en deflación, y por lo tanto el consumo, el BCE tomó medidas para estimular la economía y el crecimiento. En 2014 se implementaron medidas experimentales como una tasa de interés negativa a gran escala y así fomentar las exportaciones y los préstamos. Al asentamiento de las bases para la flexibilización cuantitativa, le acompañó la locuacidad discursiva del presidente Draghi (“Now we are in a completely different world”) que señaló que el BCE estaba dispuesto a hacer lo que fuera necesario para dar un giro a las economías de la Zona Euro: “Are we finished? The

answer is no. We aren't finished here. If need be, within our mandate, we aren't finished here.” (BBC News, 2014).

El siguiente paso consiste en observar los factores que condicionaron el movimiento semántico de los valores de sentimiento del contexto del término “inflation”. A través de los datos textométricos (Tabla 26) y la gráfica de medias de puntuación de sentimiento (Figura 19), se pueden observar tres etapas claramente diferenciadas en las tendencias de la evolución del sentimiento de la palabra.

Teniendo en cuenta ambos modelos de análisis, se observa: (1) una primera etapa en el año 2007, correspondiente a la etapa previa al estallido de la crisis, (2) una segunda etapa de seis años que inicia con un pronunciado descenso a valores de sentimiento negativos (2008) que se mantiene ligeramente ascendente hasta el año 2013. Finalmente, (3) una tercera etapa (2014-2015) de ascenso brusco a niveles positivos seguidos de una posterior caída al umbral de negatividad en el año 2015.

En el primer periodo (año 2007), la frecuencia relativa del término parte de una puntuación de 0,0790, con 136 incidencias en el subcorpus de ese año. Según el modelo de análisis de sentimiento, la puntuación media de sentimiento parte de un sentimiento positivo en 2007 (71,56) por lo que, en términos relativos, la mayor parte de las oraciones que hacen referencia a la inflación son de carácter positivo.

El segundo periodo (2008-2011) inicia con un descenso pronunciado de 47 puntos (24,09 en 2008) y tras él se estabiliza, ascendiendo levemente hasta los 32,30 en 2013. Al contrario que otros indicadores macroeconómicos como “credit” o “debt”, la caída en los valores de sentimiento no ha ido acompañada de un ascenso contundente en la frecuencia relativa del término en el corpus, sino que esta ha sido irregular en este periodo (FQ1000 0,17 en 2008, un pico de 0,23 en 2011 y 0,15 en 2013). El estadístico de unidades positivas cae de 1,21 en 2007 a 0,93 en 2008, mientras que el aumento de unidades negativas es muy considerable, pasando de 1,04 en 2007 a 1,64 en 2008.

El tercer periodo (2012-2015) engloba la época tras el cambio de políticas del BCE que refleja una pronunciada tendencia al alza en las medias de sentimiento que alcanza los 68,65 puntos en 2014, coincidiendo con las políticas de *Quantitative Easing* para posteriormente volver a caer en 2015 (32,90). Tras un periodo inicial de continuidad en el nivel de frecuencia en el discurso, la frecuencia aumenta en 2014 y 2015 (FQ1000 0,24 y 0,38). Este aumento de puntuaciones en 2014 se justifica al pasar de 1,31 unidades positivas en 2013 a 1,41 en 2014. La caída producida entre 2014 y 2015 se justifica por el aumento de unidades negativas pasando de 1,29 en 2014 a 1,44 en 2015.

**Figura 19** Gráfico longitudinal de las medias anuales de las puntuaciones de sentimiento para “inflation” y la frecuencia relativa del término por cada 1000 tokens

<i>Año</i>	<i>N</i>	<i>FQ1000</i>	<i>Sentimiento</i>	<i>pos_items</i>	<i>neg_items</i>	<i>neu_items</i>	<i>none_items</i>
2007	136	0,08	71,42	1,21	1,04	1,86	9,35
2008	308	0,17	24,09	0,93	1,64	1,90	8,51
2009	152	0,08	27,33	1,05	1,41	1,99	9,11
2010	251	0,11	28,26	1,16	1,44	1,92	10,01
2011	615	0,23	26,77	1,07	1,56	1,81	9,03
2012	300	0,12	30,44	1,38	1,69	1,81	9,83
2013	523	0,15	32,30	1,31	1,38	1,88	9,14
2014	922	0,24	68,85	1,41	1,29	1,95	9,64
2015	1637	0,38	32,90	1,43	1,44	1,96	9,88

**Tabla 26** Resumen estadístico del análisis de las oraciones que incluyen “inflation”

### Primer intervalo de estudio

Al igual que en los análisis anteriores, la primera fase de análisis corresponde a la época inmediatamente previa a la explosión plena de la crisis (año 2007). Tras analizar las oraciones en busca de colocaciones se observó que priman elementos léxicos de carácter especializado, así como la presencia de verbos de direccionalidad con un sentimiento de carácter tanto positivo como negativo sin un gran número de coocurrencias. De esta forma

se puede evidenciar que el valor positivo en la media de puntuaciones de sentimiento de este intervalo de análisis recae en los altos niveles de especialización de las oraciones analizadas.

Un análisis de las colocaciones de “inflation” en el subcorpus 2007 realizado mediante *The Sketch Engine* proporcionó un listado (Tabla 27) de palabras que permiten observar de manera sencilla las relaciones con las distintas categorías gramaticales que coocurren con el término de estudio:

- Sustantivos genéricos: *level, pressure, rise, threat*.
- Sustantivos específicos: *economy, expectation, interest, price, rate, risk, target*.
- Verbos específicos: *adjust, accelerate, combat, measure, rise, slow, hit, fall*.

<i>Palabra</i>	<i>2007</i>	
	<i>Cooc</i>	<i>logDice</i>
adjust-v	8	10,422
accelerate-v	5	9,729
rate-n	21	9,477
expectation-n	4	9,186
<b>combat-v</b>	3	9,175
target-n	6	9,113
measure-v	4	9,057
slow-v	4	9,022
<b>threat-n</b>	3	8,654
<b>risk-n</b>	7	8,588
interest-n	8	8,392
<b>pressure-n</b>	4	8,341
level-n	5	8,338
economy-n	6	8,278
<b>rise-v</b>	7	8,174
<b>rise-n</b>	3	8,121
hit-v	3	8,113
<b>low-j</b>	5	7,936
<b>fall-v</b>	5	7,901
price-n	14	7,837

**Tabla 27** Principales colocaciones de “inflation” durante el primer intervalo

Una búsqueda aleatoria de “inflation” en el subcorpus de 2007 ofrece oraciones con una connotación principalmente positiva. En ella, “inflation” se presenta como un indicador macroeconómico que sirve para conocer el estado general de la economía y el consumo. Entre los ejemplos no se observan indicios de continuidad en la narración que indiquen algún tipo de evento relacionado relevante de aquel año, sino que son más bien de carácter descriptivo genérico de la economía. A continuación, se puede observar una muestra de las oraciones estudiadas en el corpus y algunos ejemplos de casos:



These only pay 1.15% above the RPI, but investors can put in up to £15,000 tax free, an equivalent return of 9.25% gross for higher rate taxpayers at the current level of INFLATION .

Government economists recently credited those pricing plans with contributing to a slowdown of INFLATION in the Consumer Price Index for prescription drugs.

The Stern Review assumed that a dollar of economic damage prevented a century from now (adjusted for inflation ) is roughly as valuable as a dollar spent reducing emissions today.

China's burgeoning economic might means that its INFLATION and growth rates are likely to be felt across the globe.

Yet most economists, on a bipartisan basis, agree that relatively independent central banks have a better record of maintaining economic stability and keeping down price INFLATION .

Suddenly everyone is discussing the rate of inflation , the price of gas or flat screen TVs and the pay rise they might get as a result of INFLATION leaping to a 15-year high.

La colocación más significativa en este periodo es “adjust”, un término utilizado como participio “adjusted”, es decir las medidas de retorno que toman en cuenta la acumulación de las tasas de inflación para crear equivalencias entre precios históricos.

Prices are closing in on a record level, ADJUSTED for INFLATION , of \$101.70 a barrel in April 1980.

Otros ejemplos que sustentan la afirmación de que la positividad corresponde a “target” o el caso de “slow”, un elemento léxico que a priori puede suponerse como negativo y que, en todo caso, suele ser neutro alcanzando a ser positivo:

However, it expected consumer price INFLATION to SLOW towards its 2% TARGET as last year's rises in gas and electricity bills fall out of the figures.

"I don't think any one country, even as big as China or even the United States, would be able to stomach the adjustment" of raising interest rates far enough to SLOW global INFLATION , he said.

Con excepciones, los escasos usos de sustantivos considerados generalmente negativos como “pressure” suelen estar en contextos neutros o positivos. Como se puede observar a continuación, todos los ejemplos entran dentro de la narración de la cotidianeidad de los reportes macroeconómicos:

Analysts say they are not overly concerned about inflationary PRESSURE because the core INFLATION number was modest, with inflation rising just 1.5 percent in 2006.

That has put the onus on government leaders to try steering the economy, which is expected to expand by as much as 10 percent this year, away from a possible escalation of INFLATION PRESSURES that could derail some of their achievements and temper the country's climb in living standards.

"While we fully accept the need for today's rate rise, we see no reason for a further increase at present, as the impact of the 1% increase in rates since last August should be sufficient to keep INFLATION PRESSURES into 2008 under control," said Ian McCafferty, chief economist at the CBI.

So far the central bank has taken no action, and some critics say it is also under PRESSURE to keep INFLATION down.

### Segundo intervalo de estudio

La segunda etapa de análisis abarca el prolongado intervalo de tiempo que se expande a lo largo de los seis primeros años de la crisis (2008-2013). Aquí, el contexto de “inflation” se caracteriza por tener más coocurrencias con unidades del dominio específico (p. ej., “CPI” o “Consumer Price Index”, el indicador métrico de la inflación) y presencia de algunos elementos especializados con sentimiento negativo como “deflation” y otros genéricos como “threat” o “fear”, verbos como “erode” y adjetivos como “negative” o “double-digit”. Asimismo, se encuentran múltiples colocaciones estables presentes en las listas de cada uno de los años de estudio. En este caso hablamos de “rate”, “adjust” o “target” elementos utilizados para describir los datos macroeconómicos de la inflación y mantienen niveles de coocurrencia estables según indican los respectivos valores aportados por el coeficiente *logDice*.

El control de la inflación fue uno de los caballos de batalla de los bancos centrales durante este periodo. El Banco Central Europeo y otros bancos centrales mantenían los niveles de inflación a niveles artificialmente bajos, lo que reducía el consumo y frenaba la capacidad de crear empleo. Ya que es un indicador macroeconómico altamente abstracto, se puede observar que se utiliza metafóricamente una gran cantidad de verbos relacionados con la contención (“combat”, “control”, “curb”, “keep”, “prevent”...). El miedo infundido en el discurso a unos niveles de incertidumbre que podrían provocar una inflación descontrolada ya sea al alza (“double-digit”) como a la baja (“negative inflation”, “deflation”) es recurrente a lo largo de este periodo y se puede evidenciar fácilmente a través de la observación de ejemplos en el corpus.

A continuación, se puede observar una lista de las colocaciones más frecuentes del término “inflation” con el número de años en las que aparece entre paréntesis. Además, en la Tabla 28 se puede observar las veinte colocaciones más frecuentes de cada año.

- Sustantivos genéricos: *August, committee, consumer (2), drop (2), fall (2), fear, food (3), forecast (2), grocer, high, measure (3), pace, pressure, prospect, rise (4), statistics, threat, unemployment, view.*
- Sustantivos específicos: *adjustment, CPI (3), deflation, expectation (3), Fed (2), growth (2), interest, price (4), rate (5), sterling, target (5), wage (3).*
- Adjetivos genéricos: *annual, central, core, high (4), low, (4), moderate.*
- Adjetivos específicos: *double-digit (2), monetary, overall, real, negative.*

- Verbos genéricos: *control, curb, drive, ease* (3), *fall* (2), *influence, keep, outstrip, pick, prevent, remain, rise* (5), *worry*.
- Verbos específicos: *adjust* (6), *combat, erode* (3), *fight* (2), *fuel, hit* (2), *push* (2), *slow, target*.

Entre las colocaciones más destacadas en 2009 y 2010 están “negative” y “deflation” que claramente hacen alusión al miedo a la deflación, o inflación negativa. Esta ocurre cuando los precios caen porque la oferta de bienes es mayor que la demanda de esos bienes debido, en estos tiempos, a los efectos de la crisis: la reducción del crédito y el gasto del consumidor. Estos factores incluyen una escasez de dinero en circulación, lo que aumenta el valor de ese dinero y reduce los precios de manera que se producen más bienes de los demandados. Esto implica que las empresas disminuyen los precios y margen de beneficio, por otro lado, la escasez de circulación del dinero implica la retención de la divisa y la disminución del gasto. Como se puede observar en los ejemplos para ambos casos, el riesgo de deflación fue uno de los principales miedos en los mercados. Esta atmósfera de incertidumbre se puede observar mediante el uso de modalidad y los condicionales junto con la coocurrencia de unidades léxicas con sentimiento claramente negativo (p. ej., “worsening” o “contraction”).

This could soon become NEGATIVE INFLATION, with prices actually falling.

Experts are predicting a third consecutive quarter of contraction amid warnings that consumer price INFLATION risks slipping into the NEGATIVE next year, while unemployment is approaching levels not seen since the dark days of the 1990s.

It is already a consensus view that core consumer INFLATION will turn NEGATIVE soon, but we must watch if a worsening of the economy pushes Japan into a deflationary spiral, even though the Bank of Japan sees no signs of that happening right now.

It's unlikely we would have seen INFLATION drop to NEGATIVE levels for more than a year, given all the fiscal and monetary stimulus that's in the economy.

"If NEGATIVE INFLATION risks materialise, assistance from Greece's euro area partners may need to be considered," the OECD said.

2008			2009			2010		
Palabra	Cooc	logDice	Palabra	Cooc	logDice	Palabra	Cooc	logDice
rate-n	35	9,625	adjust-v	9	10,415	adjust-v	13	10,405
food-n	20	9,427	CPI-n	6	10,117	food-n	26	9,664
<b>high-n</b>	7	9,207	<b>negative-j</b>	5	9,316	rate-n	20	9,017
sharply-a	8	9,205	sterling-n	3	8,948	<b>rise-v</b>	18	8,959
<b>rise-v</b>	21	9,185	drop-n	6	8,916	combat-v	4	8,917
wage-n	8	9,099	influence-v	3	8,912	price-n	33	8,906
adjust-v	6	9,034	rate-n	16	8,791	<b>low-j</b>	16	8,864
Fed-n	9	8,967	<b>fall-n</b>	5	8,274	monetary-j	5	8,838
annual-j	10	8,874	prevent-v	3	8,240	<b>deflation-n</b>	4	8,830
adjustment-n	5	8,850	<b>push-v</b>	4	8,157	prospect-n	6	8,730
curb-v	5	8,838	central-j	3	8,154	<b>fear-n</b>	6	8,593
keep-v	16	8,827	<b>low-j</b>	8	8,096	expectation-n	5	8,573
expectation-n	6	8,820	August-n	3	8,093	<b>double-digit-j</b>	3	8,492
target-n	7	8,725	food-n	6	8,001	measure-n	7	8,454
<b>double-digit-j</b>	4	8,590	<b>fall-v</b>	10	7,982	hit-v	8	8,398
hit-v	9	8,586	target-n	3	7,962	<b>ease-v</b>	3	8,304
<b>fight-v</b>	5	8,540	consumer-n	10	7,927	<b>high-j</b>	19	8,295
core-j	4	8,453	view-n	3	7,857	<b>threat-n</b>	4	8,251
measure-n	6	8,387	committee-n	3	7,830	<b>control-v</b>	4	8,225
drive-v	7	8,372	rise-n	3	7,830	grocer-n	3	8,209
2011			2012			2013		
Palabra	Cooc	logDice	Palabra	Cooc	logDice	Palabra	Cooc	logDice
rate-n	73	10,101	adjust-v	14	10,211	target-n	30	9,666
<b>high-j</b>	76	9,607	<b>fall-v</b>	28	9,384	wage-n	21	9,385
<b>rise-v</b>	48	9,586	<b>low-j</b>	27	9,355	adjust-v	13	9,367
wage-n	18	9,290	rate-n	35	9,202	expectation-n	16	9,199
adjust-v	13	9,253	rise-v	18	8,841	rate-n	42	8,996
<b>rise-n</b>	22	9,170	<b>high-j</b>	33	8,825	<b>high-j</b>	46	8,825
price-n	64	9,074	<b>fall-n</b>	9	8,799	<b>low-j</b>	27	8,809
<b>growth-n</b>	37	9,053	unemp'yment-n	10	8,752	<b>rise-v</b>	27	8,759
hit-v	20	8,987	drop-n	6	8,660	price-n	48	8,743
measure-n	17	8,921	ease-v	5	8,618	<b>rise-n</b>	15	8,669
CPI-n	9	8,810	<b>price-n</b>	30	8,616	forecast-n	9	8,575
slow-v	11	8,790	statistics-n	4	8,441	remain-v	20	8,375
<b>fuel-v</b>	9	8,672	overall-j	6	8,412	<b>pressure-n</b>	11	8,242
target-n	14	8,548	CPI-n	3	8,285	target-v	7	8,226
forecast-n	9	8,428	consumer-n	17	8,264	<b>ease-v</b>	6	8,207
<b>erode-v</b>	7	8,425	rise-n	7	8,260	outstrip-v	5	8,182
push-v	12	8,423	<b>erode-v</b>	3	8,158	<b>moderate-j</b>	5	8,172
<b>fight-v</b>	9	8,416	pace-n	4	8,114	<b>worry-v</b>	7	8,134
interest-n	18	8,380	target-n	7	8,098	<b>growth-n</b>	24	8,130
real-j	13	8,357	pick-v	4	8,037	<b>erode-v</b>	5	8,097

Tabla 28 Principales colocaciones de “inflation” durante el segundo intervalo



En términos similares, se puede observar el uso de “deflation” en un contexto totalmente negativo. Si bien los objetivos de los bancos centrales era mantener unos niveles bajos de inflación, el discurso de la prensa evidenciaba los peligros que una eventual deflación pudiese causar. Al igual que en las oraciones en las que se incluye “negative inflation”, las formas verbales aquí suelen de tipo modal. Las unidades léxicas con las que más se relaciona la deflación están dentro del ámbito del declive del mercado laboral (“unemployment”, “fewer jobs”) que paralizaba el consumo y asfixiaba la economía, como se puede observar en los ejemplos aleatorios del corpus:

Sooner or later, policymakers will understand that the real threats are mass unemployment and DEFLATION, not INFLATION and the size of budget deficits.

In addition, there would be about 8.5 million fewer jobs, on top of the more than 8 million already lost; and the economy would be experiencing DEFLATION, instead of low INFLATION.

Bernanke said a prolonged period of high unemployment could choke off the US recovery and that the low level of INFLATION meant the risk of DEFLATION was greater than desirable.

But data on Tuesday provided new evidence that the biggest threat to prices right now was not INFLATION but DEFLATION.

In addition, there would be about 8.5 million fewer jobs, on top of the more than 8 million already lost; and the economy would be experiencing DEFLATION , instead of low INFLATION.

Durante los años 2011 y 2012, aparecieron distintos verbos que se utilizan metafóricamente para ilustrar el efecto de los niveles de inflación en la economía. El más significativo es el uso de “erode” utilizado para ilustrar el lento y progresivo daño que el control de la inflación y la falta de políticas expansivas hacían a los distintos aspectos de la economía. Los niveles de inflación se presentaban en la prensa como uno de los factores que impedían el crecimiento económico en todos los ámbitos: consumidores (p. ej., “older pensioners”), sector privado (p. ej., “Primark”), bonos del estado (“bond market”).

High INFLATION will ERODE the real value of the stream of payments; if inflation is low, the bond will hold its value better against other assets.

According to the research, the people who will experience the greatest levels of financial stress are young single people on limited incomes who rent small flats from local councils or housing associations, and older pensioners who have found their retirement incomes ERODED by INFLATION and are dependent on state pensions.

Nevertheless, the overall picture was one of cheap money leading to labor-saving capital investment, while INFLATION ERODED wages and restrictions and excessive government deficits dampened economic activity.

Trading updates from Primark , Premier Foods and Carpetright all point to households keeping a tight rein on expenditure as INFLATION ERODES their spending power.

As a basic-rate taxpayer, he would need to find an account paying 6.5% to stop his cash being ERODED by INFLATION.

Bad policies might lead to high INFLATION , which ERODES bond values, or a bond market might seize up at a time of financial stress.

Consumer spending in the economy has picked up in recent months, but analysts worry that it will peter out if the gap between pay rises and INFLATION continues to ERODE disposable incomes – a trend underlined by the latest data.

El fantasma de una inflación descontrolada se evidencia en el discurso a través de colocaciones adjetivales que rozan el uso fraseológico especializado como es el caso de “double-digit inflation”. Aunque técnicamente este haría referencia a una inflación que superase el 10%, el cual sería desastrosos según los economistas de tradición liberal, este se utiliza para infundir miedo al lector a través de dos estrategias principales. Por un lado, su uso se relaciona con alusiones a crisis anteriores (“Thatcher”, “oil crises of the 1970s”, “far cry”) o con la posibilidad de un aumento de los precios de la canasta básica (“less food”, “uncertainty”). El uso retórico de “double digit” en el discurso en detrimento de otras selecciones léxicas, entonces, aportaría una connotación negativa a las oraciones y así sería inferido por los lectores. Para ello, se muestran cinco ejemplos extraídos aleatoriamente del corpus:

Economists said DOUBLE-DIGIT INFLATION was a cause for concern because it meant less food would be eaten by the poor in a country with an annual per capita income of only \$835 but with pride in its having reduced poverty in the past decade.

That's a far cry from the DOUBLE-DIGIT INFLATION rates that battered the economy at times in the 1970s, but still worrisome.

A slump in investor confidence accelerated on Thursday, as stocks fell over 4% and the rupee dropped by 1.3% by midday due to fears of political uncertainty and dire economic challenges, including a DOUBLE-DIGIT INFLATION rate.

Prices have not risen at the speed they did during the oil crises of the 1970s, and financial policy makers have said they do not expect a repeat of the so-called wage-price spiral that led to DOUBLE-DIGIT INFLATION rates during that decade.

While Spain and Portugal are grappling with near DOUBLE-DIGIT INFLATION, Germany and other core countries are gripped with a deflationary shock, one that has also begun to infect the US which is flirting with a full blown recession.

Thatcher at least had the excuse of fighting DOUBLE-DIGIT INFLATION.

Aunque solo fue significativo durante el año 2010, el sustantivo “fear” simboliza el estado psicológico que unos niveles de inflación descontrolados infligían al discurso económico oficial. Como se puede observar en los ejemplos: los obstáculos tradicionales para la búsqueda de políticas expansivas son esencialmente dos: el miedo a la inflación, especialmente a lo que se conoce en el lenguaje especializado como “inflación acelerada” (“fueling inflation”, “high inflation”, “rising inflation”); y la preocupación por la tendencia de la balanza de pagos para el comercio exterior.

Faced with growing warnings of a double-dip recession and a eurozone crisis, the Bank is setting aside FEARS about high INFLATION to increase its programme of quantitative easing (QE).

During the last couple of months, rather than hope for growth and a need for an INFLATION hedge, we have FEAR of a double dip.

Delors criticised Germany for its insistence that the European Central Bank must not support debt-stricken members for FEAR of fuelling INFLATION .

Ms. Yellen and Professor Blinder were chosen to check the Fed's tendency to curtail growth for fear of inflation.

FEARS about high INFLATION should not prevent monetary authorities from pursuing highly accommodative monetary policy," the I.M.F. concluded.

Memories of unemployment during the Great Depression are uppermost in the minds of Federal Reserve staff, rather than FEARS of INFLATION .

Gold is used by investors as a hedging bet against rising INFLATION , but FEARS of a reduction in QE are damping concerns of higher prices because less central bank cash will ultimately flow into financial institutions.

Para concluir con este tramo, entre las colocaciones que evocan un sentimiento positivo se encuentra “ease”. Como se ha visto a lo largo de este intervalo de estudio, existe un alto factor emocional en el discurso relacionado con la inflación. La expresión de las emociones relacionadas con el miedo (“fears”, “anxiety”, “pressures”) está influida por los niveles descontrolados de inflación que desestabilizan la economía y que debía ser aliviada. El discurso periodístico de la época valoraba positivamente cualquier síntoma de mejora económica que mitigase los efectos de la crisis sobre la población, especialmente en relación con el consumo del hogar (“food price”, “commodity prices”, “consumer”). Algunos ejemplos se pueden revisar a continuación:

The rapid commodity decline has EASED fears of INFLATION, a reason central banks were able to lower interest rates around the world last week in an effort to salvage economic growth.

News of a smaller-than-expected fall in inflation last month EASED fears of DEFLATION but failed to help sterling.

The Sainsbury's figures are the lowest growth since the summer of 2005 and mark a year of steadily declining growth as food price INFLATION has EASED .

The government hopes the moves will EASE anxieties over INFLATION and help combat a widening income gap between the rich and the poor.

We expect household consumption to start increasing towards the end of the year mainly as INFLATION pressures EASE.

INFLATION EASED to 4.2% in June and even though rising utility bills could push it back above 5% soon, analysts believe any pick-up will be temporary.

UK INFLATION: Inflationary pressures should EASE this year as lower commodity prices take the heat off the consumer.



### Tercer intervalo de estudio

Finalmente, se analizarán las colocaciones del tercer y último periodo de análisis, relativo al los dos últimos años de estudio: 2014 y 2015. En 2014 se produce un pico repentino de positividad (68,85). Este máximo fue seguido de una vuelta al umbral de la negatividad (32,90) en el año 2015. En este periodo, el contexto de “inflation” coocurre con unidades con construcciones con un sentimiento gradualmente más positivo relacionadas con la direccionalidad de los indicadores (“low, y fall”) que copan la parte superior del ranking en ambos años y de manera destacada en 2014. A su vez, se observan menos elementos negativos en los ranking (Tabla 29) si se compara con el segmento anterior. A pesar de que una revisión aleatoria del subcorpus de 2014-2015 ofrece a simple vista resultados con una connotación más positiva, esto no coincide con los datos macroneconómicos. Aunque el giro en las políticas monetarias en Europa no se efectuó sino en junio de 2015 El giro discursivo del BCE previo al *Quantitative Easing* y el entusiasmo que anticipaban las inminentes reformas se puede observar p. ej., en el primer caso ofrecido a continuación:

Mario Draghi, the central bank president, said last week that the bank was determined to "do what we must" to stop insufficient INFLATION from taking hold.

It's above INFLATION , but it's definitely giving us the ability to have social impacts to measure.

But cheaper fuel will also lower the UK INFLATION rate, which at 1.3% is already well below the Bank of England's target rate of 2%.

In Japan, though QQE and short-term fiscal stimulus boosted growth and INFLATION in the short run, slow progress on the third arrow of structural reforms, along with the effects of the current fiscal consolidation, are now taking a toll on growth.

UK inflation fall to 1.6% lessens likelihood of interest rate rise CPI fall surprises forecasters who expected smaller dip as clothes discounting and drops in alcohol prices and bank charges fuel decrease Food prices have fallen, helping to cut CPI INFLATION in July to 1.6%.

In the three months to September 2014 wages increased by 1.3% when bonuses are excluded, compared to an inflation rate of 1.2% in September.

A continuación, se ofrece la lista de coocurrencias principales en este intervalo y su ranking anual en la Tabla 29:

- Sustantivos genéricos: *fall, figure, measure (2), price (2), rate (2), report, rise, unemployment.*
- Sustantivos específicos: *CPI (2), expectation, forecast, growth, target (2), wage (2).*
- Adjetivos genéricos: *annual, core, high, low (3).*
- Adjetivos específicos: *negative, zero.*



- Verbos: *fall* (2), *outpace*, *remain* (2), *rise* (2).
- Verbos específicos: *expect*, *adjust*.

<i>Palabra</i>	<i>2014</i>		<i>2015</i>		
	<i>Cooc</i>	<i>logDice</i>	<i>Palabra</i>	<i>Cooc</i>	<i>logDice</i>
<b>low-j</b>	109	10,261	<b>low-j</b>	206	10,462
rate-n	120	10,047	zero-m	75	10,309
<b>fall-v</b>	68	9,539	<b>negative-j</b>	79	10,260
wage-n	45	9,524	rate-n	155	9,621
CPI-n	18	9,255	target-n	57	9,369
annual-j	38	9,155	core-j	37	9,318
price-n	98	9,149	wage-n	63	9,298
<b>fall-n</b>	25	9,012	remain-v	72	9,205
unemployment-n	22	8,994	measure-n	46	9,153
target-n	28	8,991	expectation-n	39	9,083
adjust-v	17	8,969	<b>rise-v</b>	72	9,074
outpace-v	14	8,885	CPI-n	28	9,056
<b>growth-n</b>	56	8,697	expect-v	86	9,044
measure-n	21	8,688	figure-n	50	8,984
<b>high-j</b>	48	8,578	forecast-n	35	8,918
report-n	35	8,515	price-n	129	8,897
<b>rise-n</b>	22	8,491	<b>fall-v</b>	74	8,875
<b>rise-v</b>	33	8,489	<b>growth-n</b>	83	8,668
remain-v	29	8,476	consumer-n	51	8,595

**Tabla 29** Principales colocaciones de “inflation” durante el tercer intervalo

Si se compara con el intervalo anterior, “fall” se convierte en 2014 en la colocación principal, manteniéndose en cabeza también en 2015. Al contrario que en el segundo intervalo de estudio, la caída de la inflación aquí es considerada positiva. La observación del contexto de “inflation” en el corpus evidencia generalmente un sentimiento positivo (p. ej., “welcome”, “helping”, “brightening”).

INFLATION has FALLEN to its lowest level in over four years helping households across the country.

This FALL in the INFLATION rate is welcome, but the squeeze continues as prices are still rising faster than wages.

UK INFLATION has been FALLING in recent months, and fell to below the Bank of England's 2% target for the first time in more than four years in January, brightening the outlook for squeezed consumers and boosting hopes that wage growth will start to outpace price rises by the end of 2014.

Draghi was expected to take action to halt FALLING INFLATION across the eurozone.

FALLING INFLATION was the main contributing factor with the Kantar Worldpanel measure now standing at 1.9% – the lowest since July 2010.

Ken Mulkearn, head of pay and research at IDS, said: "The main factor in the comparison with the cost of living is the FALL in INFLATION, rather than any rapid acceleration in pay growth.

Sin embargo, cuando se observan las colocaciones de “low” en el año 2014, se observa una gran cantidad de ejemplos relacionados con el calentamiento de las nuevas políticas de expansión cualitativa siendo el Banco Central Europeo la entidad principal de la opinión. Al contrario que las valoraciones del mercado en el ejemplo anterior, aquí el contexto es generalmente negativo. Este giro discursivo se refleja especialmente en el último ejemplo en el que se habla de un nuevo paradigma económico en el que no solo prime el mantenimiento de bajas tasas de inflación y crecimiento.

It repeated a recommendation that the European Central Bank should be ready for quantitative easing should INFLATION stay too LOW.

INFLATION is too LOW in Europe It's not clear what the European Central Bank can do about the problem, though Mario Draghi is faced with a dilemma.

One should also not discount the effect on "animal spirits" – consumer, business, and investor confidence – that a credible commitment by the ECB to deal with slow growth and LOW INFLATION may trigger.

Archer said high consumer confidence, high employment and LOW INFLATION meant retailers could be in line to enjoy solid Christmas spending.

Mario Draghi, the bank's president, has so far resisted calls to tackle LOW INFLATION , but dropped a strong hint last month that he was about to relent, declaring that the governing council was "comfortable" with the idea of action in June.

Its new leanings suggest the emergence of perhaps a new consensus: Economic policy cannot be only about promoting LOW INFLATION and robust growth.

En cuanto al contexto del año 2015, las colocaciones relacionadas con “low inflation” tienen un carácter mixto de opiniones positivas y negativas. Una observación de una muestra aleatoria extraída del corpus, como se ve a continuación, evidencia que según la perspectiva o el emisor de opinión una situación de baja inflación es beneficiosa o si, por el contrario, esta es contraproducente. Por ejemplo, la opinión expresada en el corpus indica que los precios bajos pueden ser buenos para el consumo del producto final y adverso, por ejemplo, para las cifras macroeconómicas relacionadas con la deuda.

LOWER INFLATION means people have spent less on essentials and this, coupled with higher earnings and lower mortgage rates, means households in most parts of the country were left with more money with which to finance buying a home.

The figures bolstered expectations that consumer spending will continue to drive the recovery as households begin to feel some relief from LOW INFLATION and a gradual pick-up in wages.

Moreover, verbal intervention will be followed by policy action, because slower growth and LOW INFLATION – partly triggered by a strong dollar – will induce the Fed to exit zero policy rates later and more slowly than expected.

For example, what I have called the 'low-low, high-high' scenario: the risk of low growth-low inflation and high debt-high unemployment persists for a number of advanced economies.

Upbeat tone of the Nielsen survey underpinned by improving job prospects and LOW INFLATION

UK INFLATION predicted to FALL to lowest level in 25 years Forecasters claim falling cost of oil and supermarket price war will drag inflation down to 0.4% |The falling oil price is threatening to produce a zero inflation rate in the UK for the first time since 1960.

Una colocación interesante hallada en este año es “outpace”, la cual no indica un sentimiento a priori. No obstante, aquí incide positivamente en el contexto al utilizarse consistentemente para informar de que el por primera vez desde la explosión de la crisis los datos de aumento de salario han sobrepasado a los de inflación. En términos generales, la observación evidencia un sentimiento positivo reflejándose en la selección del léxico con el que estas unidades coocurren (p. ej., “hopes”, “growth”). A continuación, algunos ejemplos:

Wages growth accelerated slightly to 1% between October and December, compared with 0.9% a year earlier, adding to hopes that earnings growth will start to OUTPACE INFLATION by the end of 2014.

Britain's employers handed workers an average pay rise of 2.5% in the first quarter of the year, easily OUTPACING the official rate of INFLATION .

Wage growth has finally started to OUTPACE INFLATION after years of falling real earnings, and unemployment has fallen to its lowest level in five years, figures released on Wednesday show.

The trend marks a milestone for the economy, as wages have not consistently OUTPACED INFLATION since 2008.

The figures from the Office for National Statistics showed that INFLATION is once again OUTPACING wage growth, with the latest available data showing average pay rises of 1.7% in the three months to March.

All three groups have suffered a fall in real pay since 2009, as INFLATION OUTPACED wage growth during the financial crisis, driving living standards lower.

Household budgets have remained under pressure in the UK despite a backdrop of economic recovery, with real-terms wages falling for most of the past six years as price INFLATION has OUTPACED pay growth.

En 2015 los niveles de inflación atravesaron el umbral de negatividad y tanto “zero” se convirtió en una de las colocaciones principales. Anteriormente, se vio que la inflación negativa era vista como indeseable en el discurso. Sin embargo, como se puede observar en el discurso, cuando se habla de “zero inflation” la mayoría de los agentes portadores de opinión lo consideran como un factor deseable, especialmente en el contexto británico.

There are two reasons economists think ZERO INFLATION is a good thing.

The general assumption is that this is good news all round, even though some people do better out of ZERO INFLATION than others.

He said: "Our call on the CPI is that INFLATION will remain at ZERO in July, although we judge that there is a material chance that it will ease back into negative territory.

With INFLATION at ZERO across much of Europe, including the UK, debt levels remain stuck unless ministers can agree a budget surplus.

They believe that shop discounts – along with strong employment growth, high consumer confidence and INFLATION at ZERO – will keep households spending in coming months.

The TUC general secretary, Frances O'Grady, said: ZERO INFLATION is a mark of the weak condition the economy has been left in by a failed economic plan.

Se concluye este análisis con un ejemplo opuesto, pues cuando se observan los ejemplos relacionados con “negative” estos se presentan con una prosodia semántica más negativa. Si bien los conceptos describen situaciones idénticas, la relación aquí es diferente. Como se puede observar en los ejemplos de colocaciones extraídos aleatoriamente del corpus de 2015, estos se relacionan primordialmente con unidades portadoras de un sentimiento negativo como “stagnation”, “tumbling” o “fears”.

Last week Piketty poured cold water on the Spanish government's claims that the crisis was over and insisted the eurozone still faced the risk of stagnation, weak growth and NEGATIVE INFLATION .

Britain's brief flirtation with NEGATIVE INFLATION ended last month, with official figures showing that prices rose again in May helped by higher air fares and petrol prices.

The last time headline INFLATION was NEGATIVE in Britain was March 1960, according to the closest comparable data from the Office for National Statistics.

UK INFLATION forecast to remain NEGATIVE Consumer prices index (CPI) has been dragged down by tumbling global commodity prices and the effects of a strong pound.

Vicky Redwood, chief UK economist at Capital Economics , commented: The UK just avoided deflation in March, but INFLATION could yet dip into NEGATIVE territory at some point in the coming months ... The indirect impact of lower energy prices will take a while to come through and so the core rate could drop further in the coming months, tipping the headline rate below zero.

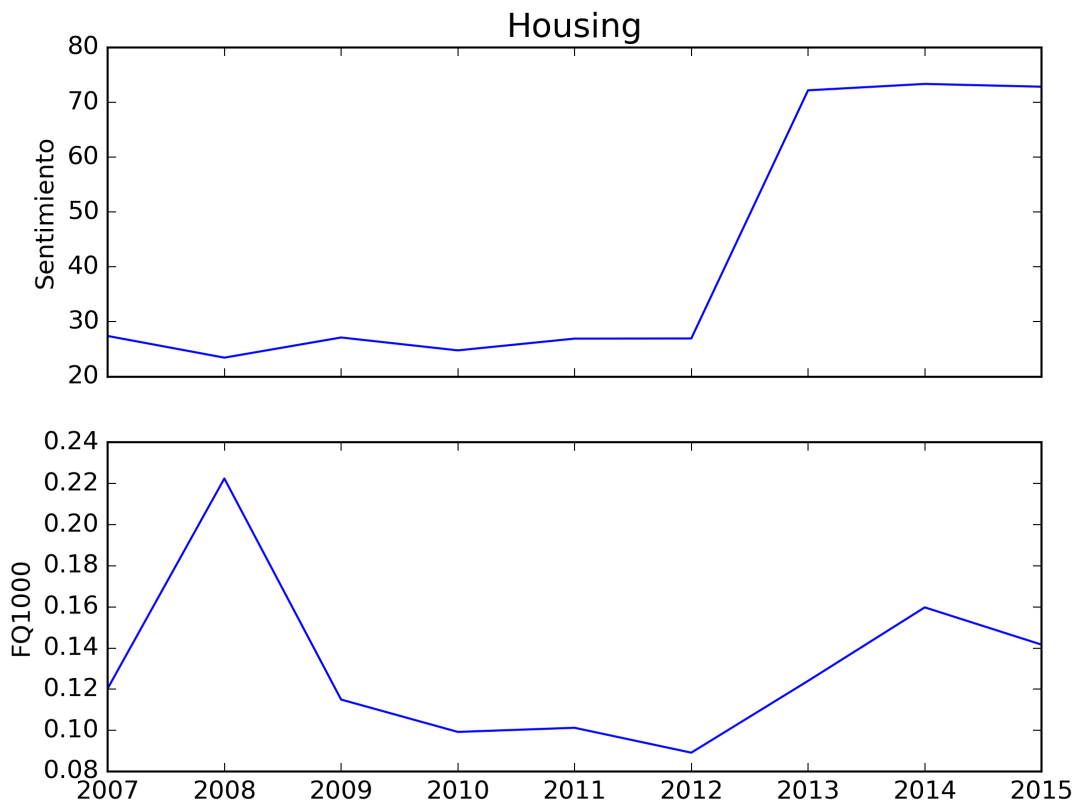
INFLATION turned NEGATIVE across the bloc late last year, raising fears of deflation, a self-reinforcing cycle of falling prices and poor growth that could have dire effects on banks and borrowers.

#### **7.4 Término 4: *Housing***

El cuarto término de estudio es “housing” (“vivienda” o “sector inmobiliario”) cuyo interés de estudio reside en los mismos cimientos de la crisis que se remontan a principios de la década de 2000. En este periodo, en las principales economías occidentales, las tasas de interés hipotecarias eran muy bajas y se relajaron los requisitos de los consumidores obtener préstamos relativamente grandes. Las personas, las empresas y los gobiernos tenían una buena liquidez y los bancos fomentaron las inversiones vinculadas a bienes raíces como una forma sencilla de obtener beneficios, considerándose el falso axioma de que la vivienda nunca bajaba. Se estaba formando una “bubble” (“burbuja”) económica y sus efectos se reflejaron inmediatamente en el discurso.

Como se revisó en el Capítulo 2, tras este mercado pujante se estaba urdiendo un complejo engranaje financiero que empezó a dar muestras de debilidad. Los prestatarios compraron viviendas sobrevaluadas y millares de personas fueron incapaces de cumplir

con el pago de las hipotecas. Como agravante, los pagos mensuales aumentaron en las hipotecas de interés variable. Ya para 2006, los precios de las viviendas dejaron de subir a una velocidad vertiginosa y en algunos casos comenzaron a estancarse o bajar. Poco después, explotó (“burst”) la crisis de las hipotecas “subprime” en 2007. La debilidad bancaria tras el congelamiento del crédito, unido a un discurso de shock y miedo causó las fallas bancarias. La ciudadanía vio a instituciones de alto nivel fallando y el discurso del pánico aumentó. En un evento histórico, se demostró que los fondos del mercado inmobiliario más allá de ser infalibles pueden hacer quebrar a la economía.



**Figura 20** Gráfico longitudinal de las medias anuales de las puntuaciones de sentimiento para “housing” y la frecuencia relativa del término por cada 1000 tokens.

El siguiente paso consiste en observar los factores que condicionaron el movimiento de la orientación semántica de los valores de sentimiento del contexto del término “housing”. A través de la gráfica de medias de puntuación de sentimiento (Figura 20) y los datos textométricos (Tabla 30) se pueden observar dos etapas claramente diferenciadas en las tendencias de la evolución del sentimiento del contexto del término: (1) una primera etapa negativa estable entre el 2007 y 2012, correspondiente a la etapa plena del estadillo de la burbuja inmobiliaria y sus consecuencias y (2) un ascenso

repentino que marca segunda etapa de tres años entre 2013 y 2015. Mientras que en la primera etapa la puntuación de sentimiento es muy negativa enteramente, la segunda etapa es positiva en su totalidad.

En el primer intervalo de estudio (2007-2012) los valores medios de sentimiento se mantienen estables por debajo de la franja de los 30 puntos (27,41 en 2007 y 26,95 en 2012). Tras un repentino ascenso de más de 45 puntos en las puntuaciones de sentimiento en 2013, el segundo intervalo se mantiene en valores positivos hasta el fin del estudio en 2015 (72,19 en 2013, 73,36 en 2014 y 72,85 en 2015). Igualmente, la cifra media de elementos positivos por segmento de análisis pasa de 1,08 en 2007 a ascender hasta los 1,49 en el año 2013, siendo este un cambio muy reseñable. De un año para otro, la cantidad de elementos positivos salta de 1,24 en 2012 a 1,49 en 2013, mientras que los negativos descienden pronunciadamente de 1,63 a 1,20 en el mismo intervalo.

En relación con las frecuencias relativas, al igual que ocurre con el análisis de “credit”, en 2008 el máximo de frecuencia (FQ1000 0,22 en 2008) coincide con la puntuación mínima en sentimiento (23,45). A partir de entonces la frecuencia de “housing” en el corpus desciende hasta su mínimo histórico (FQ1000 0,09) en 2012. A partir de entonces se produce de nuevo una tendencia ascendente en la frecuencia hasta alcanzar el pico en 2014 (FQ1000 0,16) e iniciar de nuevo una tendencia descendente con una caída al final del estudio en 2015 (FQ1000 0,14).

<i>Año</i>	<i>N</i>	<i>FQ1000</i>	<i>Sentimiento</i>	<i>pos_items</i>	<i>neg_items</i>	<i>neu_items</i>	<i>none_items</i>
2007	207	0,12	27,41	1,08	1,44	0,88	9,51
2008	413	0,22	23,45	0,97	1,72	0,87	9,58
2009	229	0,12	27,13	1,10	1,38	0,79	10,40
2010	223	0,10	24,78	1,09	1,62	1,04	11,14
2011	274	0,10	26,92	1,12	1,49	0,88	10,37
2012	222	0,09	26,95	1,24	1,63	0,88	11,51
2013	446	0,12	72,19	1,48	1,20	0,98	10,87
2014	627	0,16	73,36	1,55	1,20	1,13	11,79
2015	615	0,14	72,85	1,60	1,26	1,20	11,88

**Tabla 30** Resumen estadístico del análisis de las oraciones que incluyen “housing”.

### Primer intervalo de estudio

A diferencia de los análisis vistos anteriormente, la línea temporal del estudio comienza en plena crisis inmobiliaria y comienza con unos valores negativos que se mantendrán durante los años de la crisis del crédito. Tras analizar las oraciones en busca de colocaciones se observó que priman sobre el léxico económico, especialmente en comparación con otras unidades de estudio, elementos léxicos de carácter generalista

relacionados con el campo semántico inmobiliario (p. ej., “construction”, “stock”) y político (p. ej., “administration”, “department”). Si bien se trata de un factor económico clave y una palabra evento principal durante la crisis, el término “housing” conecta la economía con otras familias terminológicas. Analizando el sentimiento de sus colocaciones, se puede evidenciar que esta se relaciona con unidades léxicas de carácter negativo tanto específicas (“bubble”, “slowdown”) como genéricas (“excess”, “crash”). Esta amplia presencia de elementos negativos justifica que la media anual de puntuaciones de sentimiento tenga unos niveles tan bajos.

El análisis de las colocaciones de “housing” en el subcorpus 2007-2012 realizado mediante *The Sketch Engine* proporcionó un listado (Tabla 31) de palabras que permiten observar de manera sencilla las relaciones con las distintas categorías gramaticales que coocurren con el término de estudio:

- Sustantivos genéricos: *administration* (2), *advocate*, *agency*, *allowance*, *assistance*, *association*, *collapse* (6), *construction* (2), *council* (4), *crash* (5), *decline* (2), *department* (3), *development* (2), *downturn* (3), *estate* (4), *excess*, *federal* (3), *giant*, *government*, *group*, *maintenance*, *project*, *sector* (3), *space*, *stock*, *trouble*, *unemployment*, *unit*, *urban*, *year*.
- Sustantivos específicos: *boom* (6), *bubble* (5), *bust* (3), *inflation*, *finance* (4), *market* (6), *peak* (6), *recovery* (2), *shortage*, *slowdown*, *slump* (2).
- Adjetivos genéricos: *affordable* (5), *early*, *local*, *rental*.
- Adjetivos específicos: *low-income*, *social* (2).
- Verbos genéricos: *act*, *blame*, *build*, *fall* (2), *live*, *predict*, *warn*, *weaken*.
- Verbos específicos: *collapse*, *hit*, *improve*, *recover*, *slow*.



2007			2008			2009		
<i>Palabra</i>	<i>Cooc</i>	<i>logDice</i>	<i>Palabra</i>	<i>Cooc</i>	<i>logDice</i>	<i>Palabra</i>	<i>Cooc</i>	<i>logDice</i>
boom-n	13	10,149	<b>slump-n</b>	21	10,331	boom-n	13	10,045
<b>bust-n</b>	8	10,017	<b>bubble-n</b>	19	10,206	market-n	68	9,571
affordable-j	7	9,746	market-n	138	10,164	<b>bubble-n</b>	7	9,568
<b>decline-n</b>	10	9,736	<b>downturn-n</b>	22	9,912	allowance-n	6	9,529
market-n	67	9,718	boom-n	17	9,830	low-income-j	5	9,379
<b>downturn-n</b>	7	9,704	affordable-j	7	8,897	<b>downturn-n</b>	8	8,609
<b>slump-n</b>	6	9,648	federal-n	10	8,753	peak-n	4	8,528
<b>weaken-v</b>	5	9,403	<b>crisis-n</b>	17	8,752	<b>collapse-n</b>	5	8,333
<b>bubble-n</b>	5	9,208	<b>collapse-n</b>	9	8,747	<b>crash-n</b>	3	8,245
<b>slowdown-n</b>	4	9,005	price-n	37	8,478	price-n	19	8,121
<b>slow-v</b>	4	8,660	council-n	6	8,430	development-n	6	8,087
<b>trouble-n</b>	4	8,540	hit-v	9	8,386	stock-n	8	7,831
rental-j	3	8,524	<b>bust-n</b>	5	8,350	improve-v	4	7,783
project-n	7	8,499	sector-n	8	8,284	<b>crisis-n</b>	7	7,771
unit-n	6	8,405	finance-n	8	8,268	fall-v	9	7,744
council-n	4	8,370	<b>decline-n</b>	6	8,005	estate-n	4	7,669
<b>collapse-n</b>	3	8,346	warn-v	6	7,901	unemp'nt-n	3	7,432
<b>blame-v</b>	3	8,267	early-j	7	7,851	live-v	3	7,341
sector-n	4	8,231	depart'nt-n	6	7,848	local-j	4	7,318
price-n	19	8,202	fall-v	11	7,830	sector-n	3	7,105
2010			2011			2012		
<i>Palabra</i>	<i>Cooc</i>	<i>logDice</i>	<i>Palabra</i>	<i>Cooc</i>	<i>logDice</i>	<i>Palabra</i>	<i>Cooc</i>	<i>logDice</i>
market-n	65	9,265	<b>bubble-n</b>	24	10,981	boom-n	11	9,745
social-j	26	10,222	federal-n	12	9,669	urban-n	7	9,735
<b>bubble-n</b>	14	10,410	boom-n	11	9,537	<b>bust-n</b>	5	9,137
group-n	11	7,054	affordable-j	7	9,319	market-n	58	8,965
maintenance-n	10	10,102	finance-n	16	9,096	affordable-j	4	8,766
price-n	10	7,165	<b>excess-n</b>	5	8,973	assistance-n	4	8,737
estate-n	9	9,053	<b>collapse-n</b>	8	8,810	advocate-n	4	8,696
development-n	7	8,181	market-n	57	8,734	development-n	13	8,539
administ'ion-n	6	8,156	shortage-n	5	8,590	department-n	8	8,212
affordable-j	6	9,356	agency-n	16	8,508	recovery-n	5	8,174
<b>collapse-n</b>	6	8,601	con'ction-n	6	8,258	association-n	6	8,150
federal-n	6	8,833	giant-n	4	8,135	<b>crash-n</b>	3	8,133
year-n	6	4,583	act-v	5	8,020	<b>collapse-n</b>	4	8,042
country-n	5	6,199	social-j	11	7,995	finance-n	7	8,000
build-v	5	7,261	space-n	5	7,931	recover-v	3	7,923
<b>crisis-n</b>	5	7,095	recovery-n	5	7,923	construction-n	4	7,832
finance-n	5	7,681	estate-n	4	7,788	predict-v	3	7,752
government-n	5	5,752	council-n	4	7,624	price-n	16	7,726
boom-n	5	8,597	price-n	18	7,358	administ'ion-n	4	7,700
<b>collapse-v</b>	4	8,338	inflation-n	4	7,170	estate-n	3	7,690

Tabla 31 Principales colocaciones de “housing” durante el primer intervalo





“Bubble” es un término fundamental para el discurso económico de esta época. Su uso aquí hace referencia a la burbuja inmobiliaria que se dio en la primera década de los años 2000. El término, el cual está considerado como negativo en el dominio económico, es una de las colocaciones principales entre 2007 y 2011. Como se puede observar en los ejemplos extraídos del subcorpus, suele utilizarse para relatar historia reciente de manera que se justifique la causalidad de la crisis. Además, se puede observar en los ejemplos que, durante años, los analistas financieros fallaron en detectar y alertar de la crisis incipiente dado que los precios de la burbuja y los beneficios e intereses que les reportaban los créditos y la deuda les estaban siendo altamente rentables haciendo caso omiso a las voces críticas que alertaban de las consecuencias de la explosión de esta. En gran medida se destaca que ambos periódicos reparten la responsabilidad de esta burbuja a todos de una manera un tanto acrítica: entre los actores relacionados con la burbuja, se puede observar la gran diversidad de estos desde los bloques financieros (“German banks”) y los profesionales de las finanzas (“swaps”, “investors”) a los auténticos perjudicados, la ciudadanía (“young Greeks, Irish and Spaniards”) a la que se acusa de entrar en un ciclo de “borrowing binges” al intentar acceder a un bien básico como es la vivienda.

German banks helped to finance HOUSING BUBBLES in the periphery – usually not directly, but through loans to other banks.

Mr. Greenspan brushed aside worries about a potential BUBBLE , arguing that HOUSING prices had never endured a nationwide decline and that a bust was highly unlikely.

By contrast, the younger Greeks, Irish and Spaniards went on borrowing binges, driven in particular by rising demands for new homes and consumer goods that, in several cases , turned into HOUSING BUBBLES before going bust.

The Sage of Omaha has glibly defended ratings agencies, declaring that they "made the same mistake as 300 million other Americans" in failing to spot a HOUSING BUBBLE and a global financial crisis.

In the United States, during the HOUSING BUBBLE , the mythology was that we could build an economy out of endless debt, and no harm would come because the foundation upon which that debt was built – the price of housing – was indestructible.

The swaps enabled some investors to multiply their mortgage bets, while letting others wager that the HOUSING BUBBLE would pop.

El término “boom” hace referencia al periodo de bonanza en la construcción de vivienda con una connotación positiva. Está presente de manera estable entre las colocaciones principales entre 2007 y 2012. “Boom” sería la contraparte positiva del mismo fenómeno descrito por “bubble”. Al igual que en el caso anterior, este suele utilizarse para relatar historia reciente de manera que se evidencien los factores que causó

la crisis. A pesar de ser positiva, existe una diferencia primordial en la prosodia semántica de este término: cuando se utiliza “boom”, en su vecindario suelen aparecer unidades léxicas con sentimiento marcadamente negativo (“collapse”, “danger”, “loses”, “devastating”). A continuación, pueden observarse varios ejemplos:

Banks packaged thousands of home loans into bonds known as mortgage-backed securities during the HOUSING BOOM and sold them to investors around the world.

Much of the attention in the recent collapse of the HOUSING BOOM has focused on those in danger of losing their home or facing higher monthly payments in their adjustable mortgages.

Legal wrangling over Wall Street's behavior during the HOUSING BOOM has targeted virtually every step in the process, from making loans to borrowers with tarnished credit to the sale of securities engineered with the subprime loans.

At Wells Fargo, which was a major subprime mortgage lender during the HOUSING BOOM, regulators estimated that losses on first mortgages would hit almost 12 percent of its loan portfolio if the adverse situation proved correct.

"The financial shock that erupted in August 2007, as the US sub-prime mortgage market was derailed by the reversal of the HOUSING BOOM,, has spread quickly and unpredictably to inflict extensive damage on markets and institutions at the heart of the financial system," it said.

Spain, which has suffered a devastating HOUSING BOOM and bust, is also lagging behind, according to the consultancy, which says it could be a decade before either country gets its debts back to the "pre-bubble trend".

Tras el “boom” inmobiliario y su posterior burbuja, vino el estallido de esta (“bust”), una de las colocaciones más significativas en 2007, 2008 y 2013 (respectivamente, un *logDice* 10,017, 10,2016 y 9,137). El término “bust” está etiquetado en *LexiEcon* como un elemento negativo. La explosión de la burbuja de la vivienda fue un fenómeno global que afectó a los mercados financieros de todo el mundo (“Estonian”, “United States”, “California”, “Spain”...), cuyos efectos colaterales se sintieron mundialmente. En este caso no se relatan términos históricos, sino que se referencia a la misma como actualidad con tiempos verbales en presente. En las enunciaciones se observa que durante un intervalo temporal considerable no se conocían con certeza los efectos del estallido de la burbuja. Por eso, entre los elementos léxicos de su prosodia semántica se pueden observar unidades relacionadas con la incertidumbre, efecto siempre negativo en economía. Entre las unidades con las que coocurre “housing bust” están algunas tan expresivas y negativas como: “distress”, “havoc”, “casualties” o “loss-making”.

As experts debate whether we are headed for a HOUSING BUST , you would think that we should at least be able to define it.

The Estonian response to economic distress contrasts sharply to the United States, where Washington is borrowing freely and desperately trying to prevent the HOUSING BUST from wreaking havoc across the economy.

SunTrust Banks of Atlanta and Regions Financial of Alabama have been hard hit by the HOUSING BUST that has affected much of the Southeast, and their problems with commercial real estate and corporate loans have severely worsened.

Like California, some other states with outside problems from the HOUSING BUST are spending the money for something other than homeowner relief.

Then came the HOUSING BUST, which demonstrated that when lenders allowed people to buy houses they ultimately could not afford, it hurt the parties – while putting the economy itself in a tailspin.

Spain's central bank estimates some €180bn of potentially loss-making toxic real estate assets left over from the 2008 HOUSING BUST are held by banks.

A seven-member investor group including affiliates of the billionaire George Soros and of Michael Dell, the founder of Dell, have agreed to purchase IndyMac Bank, a failed lender and one of the largest casualties of the U.S. HOUSING BUST, for \$13.9 billion.

Otra forma de referenciar este evento histórico es la utilización de “collapse”, la cual está presente en el ranking de manera estable con una puntuación *logDice* entre los 8 y los 9 puntos en el intervalo 2007-2012. Claramente “collapse” se utiliza con un sentimiento negativo en todos sus contextos y como tal se refleja en sus asociaciones semánticas (“hardest”, “difficulties”, “danger”, “unemployed”...). Una búsqueda aleatoria muestra que siempre se hace referencia de manera narrativa al mismo evento ocurrido en el año 2008 y el uso de sus tiempos verbales indica simplemente proximidad (actualidad) o lejanía en la línea temporal. La gravedad de este evento y sus consecuencias provocaron un efecto negativo en los hablantes que duró, al menos, todo este intervalo pues su la mera mención de “housing” no tiene ningún uso contextual con una connotación neutra o positiva.

Once valued at \$7.1m (£3.52m) it is now thought to be worth \$5m ... due to Florida's much publicized HOUSING market COLLAPSE .

Construction workers in particular are in dire need of work, as they were among the hardest hit by the COLLAPSE of the housing MARKET.

While many high street retailers have been struggling to stay afloat, the homewares market has faced the brunt of the difficulties on the back of the COLLAPSE in the housing MARKET.

In the first stage of the HOUSING COLLAPSE , defaults and foreclosures were driven by subprime loans.

Much of the attention in the recent COLLAPSE of the HOUSING boom has focused on those in danger of losing their home or facing higher monthly payments in their adjustable mortgages.

The country is experiencing continuing mass unemployment (25 million Americans remain unemployed or underemployed), further COLLAPSE in the housing MARKET and an extremist political movement determined to slash all government spending directed at the people who are most likely to spend: the poor, the unemployed, and the middle classes.

Even prior to the SEC's prosecution, Goldman had been pilloried for being a key counter-party in swaps that led to the collapse of the insurer AIG and for taking "short" positions in advance of the COLLAPSE in America's HOUSING market.

Top, mortgage abuses contributed to the COLLAPSE of the HOUSING market; middle and bottom, after years of dismal job prospects, Americans are losing trust in many institutions, like Congress.

A partir del año 2011 y 2012, el término “recovery” apareció entre las colocaciones más frecuentes hasta el final de este estudio. Esta colocación porta un sentimiento positivo. Lejos de buscar nuevas formulas para arrancar la economía y paliar el grave problema de vivienda que gran parte de la población estaba sufriendo a raíz de la explosión de la burbuja, los poderes financieros buscaban reactivarse a partir de la recuperación de un mercado inmobiliario. De hecho, la mayoría de los ejemplos del discurso no hablan aquí de que el mercado inmobiliario esté recuperándose, sino que hablan del mismo en condicional de manera que se expresa casi de una forma más desiderativa que factual.

Una simple observación en el corpus evidencia la utilización de construcciones verbales relacionadas con la modalidad como “seems that”, “is expected to”, “would spark” o “may be”. En cierto modo, se podría decir que el discurso económico oficial, a través de sus voceros de la prensa, está invocando al crecimiento económico según el modelo las *finanzas conductuales*, de forma que aparentemente se quiere influir emocionalmente al lector influyéndole de una ventana de posibilidad optimista que distaba de la realidad económica. Durante este periodo se enunciaron términos positivos (“someway off”, “nascent recovery”, etc.) que pretendían calmar a los potenciales compradores de vivienda y animar a los inversores a reactivar el sector.

Without a HOUSING RECOVERY to revive sales of big items or major renovation supplies, Home Depot and its competitors are promoting smaller projects this spring, during what is the major selling season for home improvement stores.

With the prospect of forthcoming interest rate rises and continued shortage of mortgage funding, it seems that overall RECOVERY for the national HOUSING market is still some way off.

"Construction spending for the 2012 London Olympics is expected to help propel the economy and the HOUSING market RECOVERY ," the report said.

Reducing negative equity would spark a HOUSING RECOVERY .

It also points to one of the many problems that may be holding up the HOUSING RECOVERY .

The HOUSING sector's nascent RECOVERY foretells rising employment in the construction, real estate and mortgage finance sectors, as well as rising household wealth.

The problem is so widespread that this week the National Association of Realtors blamed faulty appraisals for holding back the HOUSING RECOVERY , saying its members had reported that more than a third of all deals were canceled, delayed or renegotiated to a lower price because of a low appraisal.

## Segundo intervalo de estudio

En esta etapa las puntuaciones medias pasan a ser positivas. Tras analizar las oraciones en busca de colocaciones se observó, al igual que en la etapa anterior, que preponderan elementos léxicos relacionados con el campo semántico inmobiliario (p. ej., “construction”, “development”, “infrastructure”) y económico (p. ej., “cost”, “crash”, “demand”). Analizando el sentimiento de sus colocaciones, se puede evidenciar que también en esta época esta se relaciona con unidades léxicas de carácter negativo específicas (“bubble”, “crash”) pero, sobre todo, destaca el aumento de elementos léxicos positivos (“affordable”, “boom”, “recovery”, “cool”) en el ranking si se compara con los segmentos anteriores. Esta amplia presencia de elementos positivos justifica que la media anual de puntuaciones de sentimiento tenga unos niveles altos. Sin embargo, una observación más detallada de los ejemplos en el corpus evidencia que existen muchos matices en el discurso periodístico que quizá no quedan del todo recogidos en el análisis de sentimiento. De hecho, se encuentran muchas oraciones, difícilmente mensurables con el clasificador, que critican el afán de crecimiento a toda costa de sector inmobiliario auspiciado por gobiernos y banca que tuvo lugar en esta época.

El análisis de las colocaciones de “housing” en el subcorpus 2012-2015 realizado con *The Sketch Engine* proporcionó un listado anual (Tabla 32) de palabras que permiten observar las coocurrencias más significativas con las distintas categorías gramaticales que coocurren con el término de estudio:

- Sustantivos genéricos: *activity*, *affordability* (2), *allowance*, *association* (2), *building*, *collapse*, *construction* (2), *crisis* (2), *development* (3), *federal*, *house*, *infrastructure*, *ladder*, *market* (3), *mortgage*, *price* (2), *recovery* (2), *sector*, *shortage* (3), *start*, *supply* (2), *urban*.
- Sustantivos específicos: *boom* (3), *bubble* (3), *cost*, *crash*, *demand* (3), *finance*, (2), *stock*.
- Adjetivos específicos: *affordable* (3), *booming*, *low-income*, *social*.
- Verbos genéricos: *build*, *rise*.
- Verbos específicos: *cool* (2).

2013			2014			2015		
Palabra	Cooc	logDice	Palabra	Cooc	logDice	Palabra	Cooc	logDice
<b>bubble-n</b>	39	10,795	<b>bubble-n</b>	38	10,496	affordable-j	28	10,059
market-n	137	9,724	market-n	214	9,985	boom-n	28	9,847
affordable-j	12	9,322	affordable-j	22	9,679	market-n	176	9,341
boom-n	13	9,245	<b>cool-v</b>	12	8,932	construction-n	25	9,280
recovery-n	15	8,921	boom-n	14	8,903	<b>bubble-n</b>	15	9,248
allowance-n	8	8,875	affordability-n	9	8,652	start-n	18	8,914
<b>shortage-n</b>	7	8,552	low-income-j	10	8,644	<b>shortage-n</b>	13	8,906
urban-n	6	8,473	construction-n	16	8,615	ladder-n	10	8,838
<b>collapse-n</b>	7	8,239	booming-j	9	8,527	affordability-n	9	8,70
<b>crash-n</b>	6	8,171	<b>shortage-n</b>	10	8,503	<b>cool-v</b>	9	8,617
development-n	16	8,149	association-n	11	8,359	association-n	13	8,470
federal-n	6	7,860	house-n	13	8,144	mortgage-n	13	8,401
activity-n	8	7,800	development-n	15	7,879	demand-n	22	8,274
stock-n	10	7,649	finance-n	12	7,838	supply-n	15	8,259
social-j	14	7,508	price-n	38	7,838	finance-n	17	8,132
price-n	20	7,471	demand-n	14	7,826	building-n	11	8,054
<b>crisis-n</b>	8	7,392	build-v	17	7,821	development-n	16	8,042
supply-n	8	7,345	rise-v	18	7,791	sector-n	22	7,994
demand-n	7	7,274	recovery-n	9	7,711	infrastructure-n	9	7,970
finance-n	6	7,264	cost-n	21	7,677	<b>crisis-n</b>	20	7,952

**Tabla 32** Principales colocaciones de “housing” durante el segundo intervalo.

El término “bubble” siguió siendo la colocación más significativa durante 2013 y 2014 (*logDice* 10,795 en 2013, 10,496 en 2014 y 9,248 en 2015), hasta que esta fue reemplazada por “affordable” en 2015. Tras una leve recuperación económica, las políticas económicas de choque implementadas por el gobierno estadounidense produjeron una mejora en el mercado inmobiliario que reavivó los miedos a una nueva burbuja inmobiliaria que esperan poder desinflar (“deflate”) a tiempo. Esto se puede evidenciar con las colocaciones negativas como “fear” o “concerned”.

Record number of estate agents raises fears of unsustainable HOUSING BUBBLE ONS data shows 562,000 people work in real estate sector – the largest number since records began in 1978 'We're no longer a nation of shopkeepers, we're becoming a nation of estate agents,' said one economist.

Britain now has a record number of estate agents, official figures have revealed, underlining fears that the fledgling economic recovery is based on inflating an unsustainable HOUSING BUBBLE.

MPs are also concerned that the monetary policy committee, which sets interest rates, will be forced to clamp down on a HOUSING BUBBLE by raising rates prematurely.

Land Registry Official data showing house prices have increased 3.4% in a year on average reignites fears of HOUSING BUBBLE.

Carney, the central bank governor, to wonder if there might be a HOUSING BUBBLE in the making and if so, what should be done about it.

Hoping to deflate a potential HOUSING BUBBLE , the Bank of England recently put a cap on lenders, saying that no more than 15 percent of their loan portfolios can consist of mortgages in which borrowers are lent amounts that exceed 4.5 times their income.

Al contrario que en la etapa anterior, aquí se habla en presente de un emergente “boom” de la vivienda (*logDice* 9,245 en 2013 y 9,847 en 2015). Si bien un “boom” es un término utilizado en términos positivos, este se usa tanto en contextos con un sentimiento positivo (“surge”, “taken off”) como negativo (“jeopardise”, “warning”). El corpus evidencia que la prensa no recibió con mucha positividad la mejora de los resultados económicos en un sector que en menos de un lustro había hecho saltar por los aires a la economía mundial y en los ejemplos se puede observar que se sacan a la luz los negocios relacionados con la corrupción (“black money”) o una nueva ola de gestión deficiente por el sector bancario (“exacerbate”).

The HOUSING BOOM "created a huge reserve of black money, especially in coastal areas", demonstrated by the "massive use" of €500 notes, which the report claims account for 70% of all the cash in circulation.

Bank of England warns housing MARKET BOOM may turn to crash.

HOUSING BOOM shifts focus to Scotland Price growth in the Scottish market is significantly outpacing London, while the average sale figure for a UK home has reached £273,000, official figures show.

HOUSING market Buy-to-let BOOM could jeopardise financial stability, says Bank of England Looser lending requirements for landlords could bolster house prices but also exacerbate falls if borrowers need to sell when interest rates rise, Bank says.

Since the Reserve Bank (RBA) began cutting interest rates in November 2011, the HOUSING BOOM has taken off.

In the HOUSING BOOM that followed, a massive surge of investors entered the market, raising their share to 52%.

El término económico “demand” se convirtió en una colocación clave en este tramo de estudio. Este término es un indicador económico con un sentimiento de naturaleza eminentemente neutra. No obstante, su uso dista de ser aséptico, ya que al igual que muchos indicadores econométricos pueden ser potencialmente positivos o negativos según las palabras en su contexto. En términos no extremos, un aumento de la demanda (en los ejemplos: “growing”, “new”, “picked up”) de la adquisición de vivienda implicaría movimiento económico y esto, por lo general, es considerado positivo. En los ejemplos se puede observar que tras la larga travesía en el desierto del mercado de vivienda el aumento de la demanda fue recogido de manera muy positiva. Esto se fundamenta en la coocurrencia con elementos léxicos que contienen un sentimiento altamente positivo (“economic growth”, “higher employment”). En los siguientes años, especialmente en el año 2015 (dos últimos ejemplos), se evidencian los efectos colaterales



con sentimiento negativo que implicaría este aumento de la demanda (“widening of the gap between rich and poor”, “yields are coming down”).

Oil-rich Saudi Arabia has growing HOUSING DEMANDS and is planning to spend about £260bn on infrastructure projects to meet its needs over the next five years.

In many markets, investors have rushed to meet the new DEMAND by building new multifamily HOUSING units or by buying up foreclosed homes and renting them out.

Then the government's Help to Buy scheme to kickstart the HOUSING market took hold and DEMAND picked up.

Company is third housebuilder to report booming business this week as Help to Buy scheme fuels HOUSING market DEMAND.

Economic growth, higher employment, increasing real earnings growth and very low mortgage rates are all supporting HOUSING DEMAND , with signs of a recent modest pick-up in demand.

Oxford Economics warned that affordability problems in the capital would intensify even further if the expected INCREASE in DEMAND for housing coincided with a widening of the gap between rich and poor.

As the DEMAND for investments in student HOUSING rises, yields are coming down.

La colocación final de estudio es “affordable housing” de carácter eminentemente positivo, el cual se convierte en el término más significativo en 2015 (*logDice* 10,059). Este es presentado de manera crítica, pues el transcurso de la crisis de vivienda dejó a millares de personas sin hogar en todo el mundo y la paralización de la construcción impidió el acceso a la vivienda a otras tantas. La demanda de vivienda asequible por la población se refleja en el corpus de manera crítica. Se observan ejemplos con unidades léxicas con un sentimiento positivos (“welcomed”, “commitment”). No obstante, la mayoría de los ejemplos presenta la problemática del acceso de la población a las mismas a pesar del amago de mejora en las condiciones de vivienda. Por poner un ejemplo de una problemática particular, como conclusión, una gran parte de los ejemplos expone la problemática que presentan las constructoras británicas que intentaban a toda costa encontrar atajos en la ley de manera que eviten cumplir con su responsabilidad legal y construir vivienda social (“opacity”, “show their sums”, “to be exposed”).

The Local Initiatives Support Corporation, already working with Morgan Stanley on AFFORDABLE HOUSING, welcomed their new commitment to funding such projects.

We urgently need to build more AFFORDABLE HOUSING in London , but unless all new homes are made available for Londoners to buy, they won't help solve the crisis.

"We wanted to increase the amount of AFFORDABLE HOUSING for students and create a sustainable, non-exploitative, community-led housing co-operative as an alternative to the private rental market.

A Conservative victory spelled an end to Labour's pledges to introduce a mansion tax on homes worth £2m or more, and to cap rent increases in the private sector to make housing more affordable for tenants.

The ratio of homes classified as AFFORDABLE HOUSING is 15%.



Real estate Councils to expose developers' 'murky' viability claims to increase cheap homes London's Islington council to publish developers' own assessments and demand evidence for profit claims to raise portion of AFFORDABLE HOUSING in new schemes Plans for the Mount Pleasant sorting office in Farringdon.

Secretive calculations used by property developers to avoid paying for AFFORDABLE HOUSING are to be exposed in a bid to increase the supply of cheap homes.

The opacity of developers' claims about how much profit they need to make if a scheme is to be viable, and therefore how much social housing they can afford to build, has been identified as a key reason for the fall in construction of AFFORDABLE HOUSING .

The Guardian view on AFFORDABLE HOUSING: the developers must be made to show their sums.

## 7.5 Término 5: *Markets*

El estudio del término “markets” (“mercados”) es especialmente relevante en el discurso de la crisis. Condensando lo revisado en el Capítulo 2, tras la explosión de la burbuja inmobiliaria y la congelación del crédito que provocó los cambios dramáticos sufridos durante la crisis actual motivaron que las grandes entidades financieras globales se convirtieran en el extremo de una cadena comunicativa dirigida a la población con la clase política como intermediaria.

El interés del estudio del sentimiento en el contexto de “markets” emerge durante los primeros años de la crisis y, especialmente en 2008, el público comenzó a escuchar con más frecuencia referencias al término en las noticias. El discurso de la clase dirigente durante los inicios de la crisis, así como su acción política, fue dirigido a la población con el fin de justificar medidas económicas de alto impacto para la ciudadanía y sus derechos (cambios en la legislación laboral, recortes en políticas públicas). De un día para otro comenzaron a hacerse frecuentes alusiones a los “mercados ” (“markets”) (enunciados como un sustantivo abstracto y plural) que eufemísticamente referenciaba a un compendio de corporaciones financieras transnacionales que, en gran parte de los casos, eran directamente responsables del colapso del sistema de finanzas (Alonso y Fernández Rodríguez, 2013). En estos tiempos, los líderes principales de Europa y los EE. UU. hablaban de la necesidad de calmar a los mercados y a realizar sacrificios mediante medidas económicas de calado, los cuales una década después evidenciaron ser una defensa radical de los intereses corporativos y la imposición de las tesis neoliberales. Este uso metafórico pone al dominio de las finanzas en relación con la mitología y los relatos religiosos: unos mercados/dioses invisibles e imprevisibles que mediados por la clase política/sacerdotal que necesitan sacrificios en forma de inyecciones de capital, recortes o privatizaciones (“provide liquidity”, “rescue package”) para ser calmados para recobrar la confianza (“confidence”). Obviamente, este discurso no pasó desapercibido para los

medios y quedó bien reflejado por la prensa: “markets” se había convertido en una palabra evento y su uso se duplicó entre 2007 y 2008. Una búsqueda en los subcorpus de 2008 y 2009 devuelve múltiples correlaciones con el verbo “calm”, el cual a priori se entiende como más propios de dominios específicos como la medicina o la meteorología que del campo de la economía y las finanzas:

The US Federal Reserve and European Central Bank tried to CALM the money MARKETS by promising to provide liquidity during the expected credit market freeze.

Economists were split on the merits of the move, with some calling it a mistake but others suggesting it would bring CALM to the MARKETS .

Other banks, he said, gambled incorrectly that CALM would return to the MARKETS after investors returned from summer vacations.

Business leaders hope that the rescue package will thaw out the frozen credit MARKETS and restore CONFIDENCE in struggling banks.

El siguiente paso consiste en observar los factores que condicionaron el movimiento de los valores de sentimiento del contexto del término “markets”. A través de la gráfica de medias de puntuación de sentimiento (Figura 21) y los datos textométricos (Tabla 33) se pueden observar varias etapas en las tendencias de la evolución del sentimiento del contexto del término en cuestión:

- (1) una primera etapa previa a la crisis correspondiente al año 2007 con sentimiento positivo (73,69).
- (2) una segunda etapa entre 2008 y 2011 con forma de uve doble (29,36, 69,37, 32,15, 31,69). Con la incertidumbre causada por la crisis la puntuación media de las oraciones que incluyen la palabra “markets” cayó 44 puntos hasta la zona negativa. En el periodo entre la crisis del crédito y la crisis de la deuda, el contexto asciende al umbral de positividad de manera puntual, para volver a caer y permanecer en el umbral negativo entre 2010 y 2011. Para ilustrar esto, el número de elementos negativos subió de 1,00 en 2007 a 1,53 en 2008.
- (3) Entre los años 2012 y 2014, los valores medios relacionados con el sentimiento de las oraciones que incluyen “markets” producen un ascenso repentino al umbral de la positividad (72,92, 72,40, 70,55). La media de unidades positivas pasa de 1,35 en 2011 a 1,51 en 2012, 1,56 en 2013 y 1,55 en 2014.
- (4) Finalmente, en el año 2015, las puntuaciones vuelven al umbral de la negatividad (29,77). Entre otros factores, es determinante observar que la

agravación de la crisis política del euro resultó ser uno de los factores principales del descenso de las puntuaciones de sentimiento. Tal es así que las puntuaciones positivas descienden de 1,55 en 2014 a 1,36 en 2015 y las negativas se disparan, pasando de 1,31 a 1,59 en el mismo intervalo.

**Figura 21** Gráfico longitudinal de las medias anuales de las puntuaciones de sentimiento para “markets” y la frecuencia relativa del término por cada 1000 tokens

En cuanto a las frecuencias relativas de uso del término de estudio, la palabra “markets” es mucho más frecuente que otros términos analizados anteriormente (p. ej., “inflation” o “housing”) pero al igual que estas, el aumento de la frecuencia supone generalmente un descenso brusco de las puntuaciones medias de sentimiento. Por ejemplo, el uso de “markets” se duplicó entre 2007 y 2008 (FQ1000 0,32 y 0,62, respectivamente) mientras que en ese mismo intervalo las puntuaciones bajaron de 73,69 a 29,36. A partir de entonces (2008-2014) la frecuencia de “markets” en el corpus se mantiene más estable (0,36, 0,41, 0,48, 0,39, 0,38, 0,44, respectivamente). Se produce un nuevo pico en la frecuencia en el año final del estudio, 2015 (FQ1000 0,61), coincidiendo también con el descenso de las puntuaciones medias del sentimiento (29,77).

<i>Año</i>	<i>N</i>	<i>FQ1000</i>	<i>Sentimiento</i>	<i>pos_items</i>	<i>neg_items</i>	<i>neu_items</i>	<i>none_items</i>
2007	556	0,32	73,69	1,30	1,01	1,01	10,12
2008	1155	0,62	29,36	1,23	1,53	1,01	9,57
2009	725	0,36	69,37	1,40	1,29	0,99	10,11
2010	920	0,41	32,15	1,36	1,36	1,05	10,05
2011	1306	0,48	31,69	1,35	1,38	1,11	10,17
2012	981	0,39	72,92	1,51	1,13	1,07	10,13
2013	1360	0,38	72,40	1,56	1,20	1,09	10,59
2014	1731	0,44	70,55	1,55	1,31	1,15	10,64
2015	2643	0,61	29,77	1,36	1,59	1,14	10,81

**Tabla 33** Resumen estadístico del análisis de las oraciones que incluyen “markets”

### Primer intervalo de estudio

Los efectos del pinchazo de la burbuja inmobiliaria no comienzan a notarse hasta el verano de 2007. Un análisis de las colocaciones de “markets” en este año permite observar la abundancia de sustantivos y adjetivos relacionados con el contexto económico (p. ej., “capital”, “credit”). En cuanto a su sentimiento, priman unidades léxicas de carácter neutro, con la presencia de una unidad etiquetada como positiva (“open”) y otra negativa (“turmoil”). La amplia presencia de elementos neutros en el listado no permite observar a simple vista las razones por las que se ha obtenido una puntuación media anual tan alta, por lo que una observación cualitativa de oraciones aleatorias extraídas del corpus permitió conocer los factores con los que se relacionan las oraciones con un sentimiento generalmente positivo.

El análisis de las colocaciones de “markets” en el subcorpus 2007 realizado mediante *The Sketch Engine* proporcionó un listado (Tabla 34) de palabras que permiten observar de manera sencilla las relaciones con las distintas categorías gramaticales que coocurren con el término de estudio:

- Sustantivos genéricos: *analyst, country, turmoil, world*.
- Sustantivos específicos: *capital, credit, currency, export, growth, mortgage, stock*.
- Adjetivos específicos: *financical, foreign, global, international, major*.
- Verbos específicos: *develop, emerge, open*.

La primera colocación de estudio es “emerge”, el cual es reconocido como verbo por el lematizador de *The SketchEngine*, pero que en la mayoría de los casos hace referencia al adjetivo “emerging” en alusión a los “emerging markets” (“mercados emergentes”). En el listado se observan varias unidades con sentimiento neutro que hacen referencia a distintos tipos de mercado (“emerging”, “developed”, “global” y “overseas”). En este intervalo observaremos ejemplos en el uso y contexto de los dos primeros.

2007		
Palabra	Cooc	logDice
emerge-v	38	10,629
financial-j	43	9,982
capital-n	24	9,447
international-j	15	9,080
<b>turmoil-n</b>	10	9,064
world-n	29	8,816
export-n	10	8,765
credit-n	29	8,755
foreign-j	12	8,720
develop-v	14	8,692
<b>growth-n</b>	16	8,660
currency-n	9	8,596
<b>open-v</b>	12	8,526
global-j	13	8,514
stock-n	18	8,513
overseas-j	7	8,415
major-j	10	8,265
country-n	16	8,153
mortgage-n	9	8,041
bull-n	5	8,041

**Tabla 34** Principales colocaciones de “markets” durante el primer intervalo

Este término hace referencia a los países con gran crecimiento que tienen características de un mercado desarrollado, pero que no cumple con los estándares para ser denominado un mercado desarrollado. Esto incluye, por ejemplo, a los BRIC (Brasil, Rusia, India y China) países que están en proceso de convertirse en mercados desarrollados en el futuro. Esta colocación es estable a lo largo de los años del estudio y contiene oraciones con todo tipo de connotaciones, sin observar que destaquen en este tramo oraciones con un sentimiento positivo, neutro o negativo. A continuación, se pueden observar algunas oraciones extraídas aleatoriamente del subcorpus:

We are one of the few global power companies that not only is operating in developed countries but also in EMERGING MARKETS , which, with respect to electricity, have much higher growth rates.

Toyota prides itself on pampering customers, but analysts are reporting weak or uneven service at Toyota sales subsidiaries, particularly in EMERGING MARKETS like China and India.

Brought in from PepsiCo last year to turn Kraft around, Ms. Rosenfeld faced weak sales in EMERGING MARKETS , few innovative products and a portfolio of brands that are about as fresh as a "Brady Bunch" rerun.

Cellphone operators in EMERGING MARKETS , and particularly in India, are adding customers at breakneck speed, making them attractive acquisitions for more sluggish Western companies.

But it could well become a cautionary tale, whether for investors in EMERGING MARKETS or for Venezuela – which, Skye says, may be liable for up to \$6 billion if all the bondholders press their claims.

En términos similares está “develop”, que al igual que el primer ejemplo, también está etiquetado como verbo por el lematizador de la suite de gestión de corpus, pero en casi todos los casos hace referencia a “developed” en la colocación “developed countries” y, por lo tanto, funciona como adjetivo. Esta colocación no responde a ningún evento histórico y contiene oraciones con un sentimiento indistintamente positivo, negativo o neutro. A continuación, se pueden observar algunas oraciones extraídas aleatoriamente del subcorpus:

Meanwhile, the capital that the United States exports to China, Russia, India and the Persian Gulf is increasingly being used to DEVELOP local financial MARKETS .

Shares of Coca-Cola rose sharply after the company reported strong gains in emerging markets like Russia, China and Africa and moderate gains in DEVELOPED MARKETS like Japan and the European Union, offsetting an expected decline in North America.

Tony Blair believes the best approach is to DEVELOP national MARKETS to cap and trade carbon, which could then be linked.

But Japan is one of the last standing DEVELOPED MARKETS that really hasn't hit its peak, and there's an opportunity there.

Since becoming chief executive in 2003, John McAdam has pursued a strategy of cutting back ICI's operations and focusing on DEVELOPING MARKETS .

La segunda colocación más significativa de este periodo es “financial”. El interés de observar esta colocación reside en la estabilidad de su significancia a lo largo del periodo de estudio y al ser la colocación con mayor número de coocurrencias. Entre los seis ejemplos seleccionados, ver más abajo, se observa que se trata de una narración de eventos financieros comunes y no se encuentra correlación en la descripción de eventos históricos. En cuanto a su sentimiento, se puede observar que cuatro de ellos son positivos. En su vecindario se pueden observar unidades con sentimiento positivo como “develop”, “confident”, “progress” o “earning double the pay”.

Meanwhile, the capital that the United States exports to China, Russia, India and the Persian Gulf is increasingly being used to develop local FINANCIAL MARKETS.

Subject to movements of FINANCIAL MARKETS and to external events, the directors remain confident that 2007 will be a further year of strong growth and progress for New Star," he said.

The person who knows how to play the FINANCIAL MARKETS, the middleman looking for someone who has something to sell and can anticipate who will buy it," he said – they will make a real difference.

A growing chorus of voices from Wall Street to Washington is concerned that the rules have eroded U.S. leadership in global FINANCIAL MARKETS .

FINANCIAL MARKETS have pencilled in a further quarter-point rise in coming months and analysts said Mr King's remarks reinforced the impression that borrowing costs had yet to peak.

Not far from the tomb of Vladimir Lenin, who vowed to destroy capitalism, investment bankers in Moscow are now earning double the pay of their counterparts in major FINANCIAL MARKETS.

A partir del mes de julio de 2007 comienza a coocurrir la primera unidad léxica de carácter marcadamente negativo: “turmoil”, el cual siguió apareciendo a lo largo del resto del año. En este caso sí se puede hablar de una colocación que describe un evento histórico: se habla aquí de la explosión de la burbuja inmobiliaria y la incertidumbre que produjo en los mercados. Como aquí se puede observar, esta crisis irrumpe en los mercados y en su vecindario se pueden observar más elementos negativos y relacionados con el shock y la expansión de la crisis (“unremitting”, “uproar”, “creep in”, “rocked” o “problems”).

On Tuesday, after a month of almost unremitting TURMOIL in financial MARKETS , Ben S.

The reason was the TURMOIL in the mortgage MARKETS and the resulting gyrations in share values on the stock exchanges.

But a cautiousness has begun to creep in, brought on by recent TURMOIL in the MARKETS , uproar over a lavish birthday party for the private equity executive Stephen A.

Sly Bailey, the chief executive, said TURMOIL in the financial MARKETS was a key factor.

Mr. Schott said that the scrutiny, which was first reported yesterday in Tax Analysts, a prominent trade publication, was not related to the recent TURMOIL in the mortgage MARKETS .

No financial institution has escaped damage from the TURMOIL that rocked the credit MARKETS in July and August as problems in the subprime mortgage market widened.

El siguiente elemento es “opening” con sentimiento generalmente positivo en este dominio específico. La apertura de mercados en el discurso periodístico liberal es considerada un hecho positivo. En los ejemplos, se puede observar en los ejemplos contextualizados que todos ellos tienen una valencia positiva o neutra. En las oraciones no se puede establecer relación con ningún evento especialmente relevante en el mundo de la economía:

Some of Wild Oats' smaller stores might eventually be closed while others might become larger Whole Food stores, he said, noting that such a plan would carry less risk than OPENING large stores in new MARKETS.

But social entrepreneurs may have their biggest impact by being disruptive innovators, OPENING up MARKETS that bigger organisations cannot see.

Unexceptionally he exhorted business "to get on our bikes and improve our share of world MARKETS" and pressed Japan to OPEN its markets in the face of a rising tide of Japanese imports to Britain.

That success has called into question the international economic orthodoxy that countries should keep their MARKETS as open as POSSIBLE at all times.

Von Braun called on Europe to reconsider its biofuel policies, to provide more aid to poor nations, to keep MARKETS OPEN and to boost production.

So manufacturers are aiming to provide casino operators with the money machines they need to justify the many billions of dollars they have spent, and continue to spend, expanding existing properties and OPENING facilities in new MARKETS .



Otro término del dominio económico con polaridad positiva que se encuentra en la lista de colocaciones es “growth”, dado que el crecimiento es un factor deseable en las principales escuelas de economía y como tal consta en *LexiEcon*. Una observación de la prosodia semántica de “markets” y “growth” en el corpus nos permite observar que se relaciona con otras unidades con valencia positiva como “boost” o “bolster”.

The trend is helping to boost economic GROWTH in emerging MARKETS .

The move into the US is viewed as a key component of Aviva's growth over the next few years, coupled with GROWTH in continental MARKETS such as Spain.

We consciously got into this area because there is a high need to get power into these MARKETS for economic GROWTH , and it allows us as a company to grow much faster.

In the telecommunications sector, low-income MARKETS are driving GROWTH for mobile phone operators, which are rolling out services throughout the emerging world.

By buying Danone's biscuits unit, Kraft is hoping that a larger presence in the snack business, as well as exposure to MARKETS in Europe and emerging MARKETS , will bolster growth .

The Hallmark Channels' geographic breadth is of particular importance to us, as it will bolster our position in existing MARKETS and boost our GROWTH in Asia and central and eastern Europe.

Finalmente, “credit” es la última palabra de análisis y ganó prominencia en el verano de 2007 y es claramente una palabra evento relacionada con el estallido de la crisis. La parálisis de la economía se puede observar con la colocación con el adjetivo “tight” en casi todos los ejemplos. En cuanto al sentimiento, se puede observar que todas las oraciones analizadas tienen una valencia negativa en relación con los eventos históricos que estaban aconteciendo y que se revisaron en 7.1. Además de adjetivos como “tight” o “tightening”, entre las palabras que se encuentran en el vecindario se encuentran “inability”, “mess”, “sour” o “crisis”.

The plunge came a day after the private equity firm buying Chrysler from DaimlerChrysler said it would complete the transaction for the automaker despite an inability to borrow the money in CREDIT MARKETS, as had been planned.

Perhaps as a result of tightening CREDIT MARKETS, U.S. Bank has asked for early redemption of \$400 million in bonds due in 2014 because Landry's, which is based in Houston, has not filed quarterly reports this year nor its annual report from last year, in violation of its credit agreement.

In recent months, Wall Street's mortgage mess, tight CREDIT MARKETS and the sour economy have combined to slow the torrent of private money flowing into the movie capital.

But its fate became increasingly tenuous amid the tighter CREDIT MARKETS, which have helped derail several other proposed private equity deals.

IndyMac's biggest problem was a \$10 billion portfolio of loans that it had been unable to sell last summer when CREDIT MARKETS froze up.

The crisis in world CREDIT MARKETS has cast a shadow over the financing of a number of big deals.



## Segundo intervalo de estudio

Durante la segunda etapa de análisis (2008-2011), la media de las puntuaciones de sentimiento de “markets” se agrupa en tres subfases: (a) una caída en el año 2008 al umbral de negatividad correspondiendo con la explosión de la burbuja crediticia, (b) una subida en 2009 al umbral de positividad de la parcialmente motivada por una recuperación momentánea la intervención del gobierno estadounidense y (c) una recaída tras el inicio de la crisis de la deuda que mantuvo la media de puntuaciones en el umbral de negatividad entre 2010 y 2011. Las medias de sentimiento responden claramente a los eventos históricos y se puede observar aquí con claridad la separación de las dos crisis. Mientras que en 2008 se escenificaba la explosión de la crisis crediticia y la posterior intervención del gobierno estadounidense, en el año 2010, se dan los movimientos que marcan el inicio de la crisis de la Zona Euro. En este año, Grecia es rescatada por primera vez y comienza una crisis multinivel los niveles en la Unión Europea.

Como es natural según la definición de colocación, en todos los años se mantienen coocurrencias fijas y en el caso de las relacionadas con “markets” estas suelen tener un sentimiento neutro (“financial”, “emerge”, “capital”, “international”...). Sin embargo, se destaca aquí que mientras que en 2009 y en el periodo 2010-2011 se observan colocaciones de carácter negativo (“turmoil”, “volatile”, “crisis”), en el año 2009 solo se encuentran colocaciones con sentimiento positivo o neutro como “recover” o “growth”. A continuación, se podrá observar la lista de las colocaciones más significativas cada año (Tabla 35) junto al número de subcorpus anuales en los que aparece:

- Sustantivos genéricos: *access*, *analyst* (2), *trading*, *turmoil* (2), *world* (4).
- Sustantivos específicos: *bond* (3), *capital* (4), *commodity* (3), *confidence*, *credit* (2), *currency* (2), *debt* (2), *economy*, *equity* (2), *growth*, *housing*, *investor*, *money* (3), *stock* (4).
- Adjetivos específicos: *finacial* (4), *foreign* (3), *global* (4), *international* (4), *local*, *major*, *new*, *overseas* (3), *volatile* (2), *wholesale* (2).
- Verbos específicos: *borrow*, *calm*, *develop* (2), *emerge* (4), *expand*, *fall*, *open* (2), *recover*, *turn*.

La primera colocación que se estudiará es “financial”, cuya puntuación *logDice* es consistente a lo largo del intervalo de estudio. Como se puede observar en los ejemplos extraídos de los años 2008 y 2010-2011, la mayoría de las oraciones están en un contexto con sentimiento negativo o neutro en que se encuentran muchas referencias a eventos de

la crisis económica, principalmente el colapso de Lehman Brothers. Para ilustrar esto, entre las palabras situadas en su vecindario se encuentran unidades con sentimiento negativo como “unsettled”, “unstable” o “pandemonium”.

And Mr. Matolcsy, an economist who once served as Hungary's representative at the European Bank for Reconstruction and Development in London, is not known for his diplomacy, and has sometimes unsettled FINANCIAL MARKETS .

The announcement left the FINANCIAL MARKETS convinced that it would be September at the earliest before the key US interest rate was raised from its current range of 0-0.25%.

Why can't countries borrow from the FINANCIAL MARKETS?

The 81% taxpayer-owned bank is reported to be among at least six major players in the FINANCIAL MARKETS, including Deutsche Bank in Germany and Citigroup in the US, caught up in the cartel investigation.

Following the 2008 financial crash and subsequent unstable economic and FINANCIAL MARKETS, Vigeo sought to support and strengthen the argument that all asset class investors, not only long equity (but also short equity, fixed income, sovereign, real estate and so on) must consider all risk factors - including ESG - when making investment decisions.

Claims that Barclays illegally took advantage of the pandemonium in FINANCIAL MARKETS in 2008 in order to acquire the US operations of bust investment bank Lehman Brothers at an unfair price have been thrown out by a New York bankruptcy court.

A showdown between the most powerful leaders in the eurozone and George Papandreou is under way amid increasing concern about the Greek prime minister's plan to hold a referendum and the impact it is having on FINANCIAL MARKETS.

En cuanto al año 2009, sí se puede apreciar una diferencia en el sentimiento de “financial markets”, ya que además de oraciones con un sentimiento negativo o neutro también es frecuente encontrar oraciones con sentimiento positivo. Si bien en este año hubo previsiones de mejora en los resultados económicos y como tal quedó reflejado en el discurso, estas previsiones de normalización fueron momentáneas al cernirse una crisis mayor relacionada con la deuda. Entre las unidades léxicas con sentimiento positivo se pueden encontrar “normalise”, “recovers”, “stenghtened”, entre otras.

2008			2009		
<i>Palabra</i>	<i>Cooc</i>	<i>logDice</i>	<i>Palabra</i>	<i>Cooc</i>	<i>logDice</i>
financial-j	191	10,979	financial-j	85	10,057
emerge-v	64	10,471	emerge-v	57	10,748
credit-n	114	9,986	capital-n	37	9,826
<b>turmoil-n</b>	36	9,757	stock-n	34	9,459
global-j	48	9,601	credit-n	33	9,284
capital-n	48	9,551	equity-n	22	9,236
stock-n	47	9,273	global-j	22	8,871
world-n	41	8,882	wholesale-j	9	8,476
<b>volatile-j</b>	17	8,785	international-j	12	8,475
money-n	38	8,746	foreign-j	12	8,455
commodity-n	17	8,624	<b>recover-v</b>	10	8,403
<b>confidence-n</b>	17	8,593	world-n	21	8,385
analyst-n	26	8,583	investor-n	21	8,317
housing-n	19	8,572	bond-n	10	8,262
<b>open-v</b>	19	8,548	trading-n	11	8,171
equity-n	19	8,530	major-j	11	8,005
international-j	17	8,503	money-n	18	7,983
foreign-j	17	8,486	<b>growth-n</b>	11	7,881
overseas-j	12	8,228	local-j	9	7,784
debt-n	15	8,089	turn-v	9	7,691
2010			2011		
<i>Palabra</i>	<i>Cooc</i>	<i>logDice</i>	<i>Palabra</i>	<i>Cooc</i>	<i>logDice</i>
emerge-v	77	10,851	emerge-v	120	11,045
financial-j	96	10,121	financial-j	179	10,715
capital-n	48	9,726	stock-n	91	10,278
bond-n	35	9,624	capital-n	41	9,358
stock-n	42	9,569	commodity-n	33	9,293
global-j	32	9,087	world-n	57	9,105
<b>volatile-j</b>	15	8,939	global-j	45	9,097
money-n	33	8,849	bond-n	33	9,057
currency-n	18	8,786	equity-n	27	8,913
<b>calm-v</b>	12	8,645	international-j	26	8,746
world-n	35	8,583	<b>turmoil-n</b>	16	8,485
overseas-j	13	8,571	develop-v	24	8,433
commodity-n	13	8,549	overseas-j	15	8,333
develop-v	18	8,496	currency-n	17	8,327
international-j	16	8,492	analyst-n	25	8,290
debt-n	21	8,363	access-n	15	8,165
wholesale-j	10	8,312	<b>new-j</b>	49	8,047
<b>expand-v</b>	13	8,286	<b>open-v</b>	16	7,980
foreign-j	13	8,213	<b>fall-v</b>	19	7,939
borrow-v	10	8,208	economy-n	20	7,693

**Tabla 35** Principales colocaciones de “markets” durante el segundo intervalo

"It appears to take some more time before the FINANCIAL MARKETS in the US and Europe normalise and the global economy recovers," said Watanabe, who in June will be replaced by Akio Toyoda, grandson of the company's founder, in an attempt to turn its fortunes around.

Now, some of the decline in the FINANCIAL MARKETS is occurring for a very good reason: real concern about profits and coupon payments.

In his inaugural speech, Obama seemed to suggest that the FINANCIAL MARKETS should expect greater oversight.

FINANCIAL MARKETS in all three cities have crashed.

It will take more time before the FINANCIAL MARKETS in the U.S. and Europe normalize and the global economy recovers," Toyota's president, Katsuaki Watanabe, said Friday.

Such moves are possible now, experts say, because the global FINANCIAL MARKETS have strengthened in recent weeks and commodity prices have rebounded.

The US bank JP Morgan gave a glimpse of a widening rift in fortunes between Wall Street and Main Street as lucrative business on the FINANCIAL MARKETS pushed its profits up by 36%, despite worsening defaults on its credit cards and home loans by recession-hit grassroots customers.

Schapiro, is struggling to revive its reputation as an effective watchdog of Wall Street after presiding over a near-collapse of the FINANCIAL MARKETS and failing to catch the \$65 billion Ponzi scheme run by Bernard L.

El adjetivo “volatile” adquiere significancia en el listado de colocaciones en los años 2008 y 2010, coincidiendo en una época de gran inestabilidad en los mercados, especialmente en el flujo de a crédito. Si bien en el dominio económico “volatile” tiene siempre un sentimiento negativo y como tal queda recogido en *LexiEcon*, la observación de la prosodia semántica del adjetivo durante este intervalo permite observar que a su alrededor suelen aparecer otras unidades léxicas con un sentimiento altamente negativo, así como unidades propias de los eventos descritos (“housing”, “recession”). En los ejemplos obtenidos de los textos de estudio pueden encontrarse colocaciones con una polaridad negativa como “irrational”, “beast”, “slack”, “poor”, entre otras.

These firms have been curtailed slightly by recent regulations aimed at making the MARKETS less VOLATILE.

MARKETS are extremely VOLATILE and irrational – all of a sudden people are dealing with a different beast," Cass said.

The Petrobras sale is also expected to attract interest from energy investors, who have rushed into natural resource companies as they perceive them to be safer than the recently VOLATILE credit and equity MARKETS.

The dollar's descent accelerated, particularly against the euro, in the second half of 2007 as falling housing prices, VOLATILE equity MARKETS and slack consumer spending pointed to a possible recession in the United States.

The hedge fund community is battling against poor returns, given VOLATILE credit and equity MARKETS , and against forthcoming tougher regulation.

While oil markets look well supported by Opec, BP expects gas MARKETS to remain VOLATILE and refining margins to remain depressed for the foreseeable future.

En la introducción a este análisis hablamos de la centralidad de “confidence” (“confianza”) en el discurso relacionado con los mercados. Si bien “confidence” es una unidad léxica que tanto en el lenguaje general como en el económico es considerado positivo, aquí se puede observar que en la prensa las menciones referencian a una crisis de confianza de los inversores en los mercados, mediante un conjunto abierto de verbos (“restore”, “stabilise”, “maintaining”). Aquí se puede observar que la orientación semántica de las oraciones es generalmente de carácter negativo o neutro y que existen múltiples asociaciones semánticas con un sentimiento negativo, en varios casos relacionados con las emociones (“panic-stricken”, “nervousness”).

The American people can be assured we will continue to act to strengthen and stabilise our financial MARKETS and to reassure investor CONFIDENCE .

The \$25.1bn (£15.38bn) bail-out, which was much larger than expected, aims to help restore investors' CONFIDENCE in the country's financial MARKETS and its currency, the forint, and marks the first time in the history of the EU that a member state has been rescued by the IMF.

The UK government, however, still seems more worried about maintaining CONFIDENCE in financial MARKETS than reducing unemployment – despite the fact that no one would seriously expect the UK to default on its public debt.

With Friday's loosening, the bank appeared to decide that now was the time to send a stronger signal to restore CONFIDENCE in panic-stricken financial MARKETS.

The crisis of CONFIDENCE gnawing at world MARKETS is already feeding back into slower growth.

The 62-year-old Anglo-Cypriot, who is a professor at the London School of Economics , urged the coalition to be blunt and avoid spin, for fear of sparking a Greek-style loss of CONFIDENCE on the capital MARKETS.

While some commentators welcomed the fall in sterling, arguing it would help manufacturers break into new markets with lower priced exports, there was a degree of nervousness that MARKETS were close to losing CONFIDENCE in the chancellor's economic strategy.

La siguiente unidad de estudio es “calm”, la cual marca que “markets” se convierte en una palabra evento. Como se recuerda en la historia reciente, ya desde inicios de la crisis personalidades como las del presidente del gobierno español Zapatero o Jean-Claude Trichet, el entonces presidente del BCE, utilizaron frases en la línea de “calm the markets” para referirse, principalmente, a políticas experimentales de liberalización de mercados e inyección de capital público en la banca privada.

Mientras que en un dominio general “calm” suele ser un elemento con sentimiento positivo (p.e “A calm sea”), en el dominio financiero este no se suele utilizar si no hay un evento que haya alterado los mercados previamente. Es reseñable aquí que, durante estos años se daba una situación metafórica de nerviosismo (“stress”, “turmoil”) en los mercados por la falta de crédito y un modelo ampliamente basado en la deuda solicitaban políticas de alivio financiero que permitiera el flujo de liquidez. Si bien esta expresión no

es nueva pues la primera referencia en el corpus es de agosto de 2008, esta ganó relevancia en el estudio en el año 2010 con 12 coocurrencias y un coeficiente *logDice* 8,645. En este año se produjo el rescate griego que dio inicio a la crisis en la Zona Euro y en el mismo discurso se puede observar la historicidad de los eventos al observarse portadores y objetos de opinión como gobiernos e instituciones “Spain”, “Greece”, “eurozone” o privados “investors” que, aprovechando el rescate, intervienen en el discurso solicitando al Banco Central Europeo liquidez ilimitada y políticas beneficiosas para las élites financieras. A través de los ejemplos del corpus, se puede observar que entre los elementos de asociación semántica aparecen unidades léxicas con sentimiento negativo como “cuts” y otras relacionadas con el campo semántico de la medicina como metáfora de enfermedad: “victim”, “contagion” y “anxious”.

Spain's government will seek to reassure investors over the country's rising debts today, in a bid to CALM the turmoil in financial MARKETS that has seen a heavy sell-off in Spanish shares continue for a second day.

While the eurozone leaders hoped the package agreed last night would suffice to CALM the MARKETS and they would not need to supply tens of billions of euros to Athens, the statement agreed they were "ready to contribute to co-ordinated bilateral loans ...

But the amount of intervention so far is far smaller than many investors and economists think is necessary to CALM MARKETS.

Today, head of the Greek central bank George Provopoulos tried to CALM MARKETS by saying spending cuts to achieve "ambitious fiscal targets" was the best option as further increases in tax rates would hurt economic activity which was already slumping.

While the measures were clearly designed to CALM the MARKETS, which have been eyeing up Spain as a potential victim of contagion from the Greek debt crisis, it was unclear what they would do to Spain's timid recovery.

His remarks exposed differences with the Central Bank of Ireland governor Patrick Honohan, who had attempted to CALM anxious MARKETS by saying that he expected Ireland to accept "tens of billions" in loans from the EU and IMF.

Otra unidad léxica que se repite a lo largo del intervalo (2008 y 2010) es el adjetivo “volatile”. Se habla de mercados volátiles cuando el mercado de valores sube y baja más del 1% durante un período de tiempo determinado. En el dominio económico, la volatilidad es un indicador y como tal tiene un sentimiento potencialmente negativo pues, en la mayoría de los casos, cuanto mayor sea la volatilidad, mayor será el riesgo (“risk”) de la seguridad (“security”).

Durante estos años, hablar de mercados era hablar de falta de estabilidad y volatilidad. Especialmente en este intervalo, cualquier noticia en cuanto al desarrollo de los eventos en la economía se traducían en caídas en el mercado de valores. Una revisión del corpus para este intervalo devuelve oraciones con un sentimiento negativo

generalizado. En cuanto a su vecindario, se pueden observar unidades con sentimiento claramente negativo como “depressed”, “crisis”, “yo-yoing”, “damage”, “risky”, entre otras.

While oil markets look well supported by Opec, BP expects gas MARKETS to remain VOLATILE and refining margins to remain depressed for the foreseeable future.

The stream of junk bond sales seen earlier this year, which included the £500m bond issue by Manchester United, has slowed to a trickle over the past few weeks because of VOLATILE MARKETS and the Greek debt crisis.

Stock MARKETS were described as VOLATILE with many shares yo-yoing throughout the day as traders attempted to assess the damage from the decline in creditworthiness of Spain and Greece.

Financial experts have predicted that there could be more blockbuster sales to come as ultra-wealthy individuals look for alternative ways to invest their money away from the risky, VOLATILE financial MARKETS.

The two sides are arguing about whether this reflected a shift in VOLATILE MARKETS, impairing the value of assets held on Lehman's floor, or whether it was quietly inserted into transaction documents.

Fund managers are seeking these situations as VOLATILE equity and credit MARKETS have made it more difficult to gain returns from tracking indexes, such as the FTSE 100.

The hedge fund community is battling against poor returns, given VOLATILE credit and equity MARKETS , and against forthcoming tougher regulation.

### Tercer intervalo de estudio

Durante la tercera etapa de análisis (2012-2014), la media de las puntuaciones de sentimiento de “markets” se sitúa siempre por encima del umbral de los 70 puntos, correspondiendo con un valle en la frecuencia relativa de uso del término en el corpus. En este intervalo, “markets” suele aparecer en el discurso en relación con una apertura a nuevos horizontes de expansión internacional tras un casi un lustro de paralización del crédito. Las empresas privadas son las entidades principales de estudio de la opinión y, en comparación la etapa anterior, destaca el descenso de las referencias a la crisis económica e institucional que afrontaban muchos mercados nacionales. Más abajo se puede observar la lista de las colocaciones más significativas cada año (Tabla 36) junto al número de subcorpus anuales en los que aparece:

- Sustantivos genéricos: *access* (2), *analyst* (3), *farmer*, *target*, *world* (3).
- Sustantivos específicos: *bond*, *capital* (3), *currency* (3), *equity*, *exchange*, *export*, *growth* (2), *investor* (3), *stock* (3).
- Adjetivos específicos: *developed*, *financial* (3), *fast-growing*, *foreign* (3), *global* (3), *international* (3), *local*, *major* (2), *new* (3), *overseas* (2).
- Adjetivos genéricos: *key*.
- Verbos específicos: *develop* (3), *emerge* (3), *expand*, *fall*, *grow*, *open*, *rig*.



2012			2013			2014		
Palabra	Cooc	logDice	Palabra	Cooc	logDice	Palabra	Cooc	logDice
emerge-v	112	11,249	emerge-v	151	11,274	emerge-v	216	11,425
financial-j	93	10,021	financial-j	112	9,877	financial-j	183	10,307
stock-n	45	9,711	stock-n	65	9,760	foreign-j	61	9,659
capital-n	38	9,562	capital-n	38	8,992	stock-n	71	9,603
bond-n	29	9,195	foreign-j	30	8,950	exchange-n	45	9,266
export-n	19	8,795	currency-n	28	8,939	currency-n	43	9,147
global-j	28	8,676	overseas-j	22	8,805	capital-n	42	8,96
world-n	37	8,672	key-j	26	8,658	global-j	56	8,893
develop-v	24	8,640	global-j	36	8,557	<b>rig-v</b>	25	8,748
<b>new-j</b>	62	8,634	international-j	25	8,480	international-j	35	8,711
international-j	18	8,498	develop-v	30	8,416	access-n	22	8,203
foreign-j	16	8,471	world-n	45	8,405	farmer-n	23	8,176
<b>fast-growing-j</b>	11	8,416	equity-n	18	8,317	develop-v	27	8,087
target-v	12	8,384	<b>growth-n</b>	32	8,153	<b>new-j</b>	71	8,003
overseas-j	11	8,303	access-n	18	8,134	local-j	25	7,959
developed-j	9	8,133	<b>new-j</b>	61	8,005	world-n	36	7,935
analyst-n	17	8,082	major-j	19	7,888	major-j	23	7,896
<b>open-v</b>	12	7,967	<b>grow-v</b>	25	7,883	<b>growth-n</b>	38	7,875
foreign-j	16	8,471	investor-n	23	7,845	analyst-n	23	7,818
<b>expand-v</b>	11	7,960	<b>fall-v</b>	18	7,758	investor-n	26	7,773

**Tabla 36** Principales colocaciones de “markets” durante el tercer intervalo

Las referencias al crecimiento se dan a través de las unidades “fast-growing” (2012) y “growth” (2013-2014). Obviamente, el crecimiento es un factor económico deseable y como tal queda registrado en *LexiEcon*. En ambos casos, las oraciones tienen un sentimiento general positivo que responde a una mejor situación para las empresas multinacionales de Europa y Estados Unidos que, a pesar de la crisis en sus países de origen, habían buscado expandirse en unos mercados emergentes (“emergent”) menos azotados por la crisis y se pueden observar múltiples referencias a Asia o África. En términos generales, las oraciones en las que coocurren “growth”/“fast-growing” y “market” suelen presentar una prosodia semántica que incluye unidades léxicas especializadas de carácter positivo como “bigger dividends”, “profitable” o “spur”. A continuación, puede observarse una serie de ejemplos extraídos del corpus:

EADS chief Tom Enders confirms push into FAST-GROWING Asian MARKETS It would be a 'mistake' to ignore Asia-Pacific, which will account for a third of new aircraft deliveries by 2032, says EADS boss EADS chief executive Tom Enders also said the company's €34bn merger with BAE Systems was not being driven by cost savings.

HSBC warns on jobs as cost-cutting continues In a strategy update, bank shifts focus to FAST-GROWING MARKETS and considers bigger dividend payouts to shareholders HSBC is focusing on cost-cutting and FAST-GROWING MARKETS.



Tesco , which has agreed a memorandum of understanding with CRE, said on Friday morning that the proposed tie-up was consistent with its "strategy of focusing on profitable routes to GROWTH in FAST-GROWING but less mature MARKETS , with a disciplined approach to the allocation of capital".

The government has been keen to build links with China as part of its efforts to get British companies trading with FAST-GROWING emerging MARKETS .

The changes, effective July 1, come as Procter & Gamble seeks to spur GROWTH in emerging MARKETS and improve profit and revenue globally.

As the company establishes manufacturing sites in GROWTH MARKETS , it aims to source at least 40% of components locally, fostering local enterprise while reducing costs and environmental impacts resulting from transport.

This is becoming increasingly important as insurers look to these markets for new opportunities for GROWTH now that the UK MARKET is saturated – Williams' pun.

The airline said 85 aircraft would be used to replace older ones, and the rest deployed for "organic GROWTH in existing MARKETS ".

Thus, emerging economies with large twin deficits and other macroeconomic fragilities may experience further downward pressure on their financial MARKETS and GROWTH rates.

En la misma línea, entre los años 2012 y 2014 “new” se asienta como una colocación propia de “markets”. Como no puede ser de otra forma, la apertura de nuevos mercados se considera un hecho positivo en un mundo globalizado donde el liberalismo es la ideología dominante y como tal queda reflejado en el discurso de la prensa. Al igual que en el caso anterior se puede observar que estos nuevos mercados abiertos a los que se hace referencia están relacionados con las economías emergentes, por ejemplo, China. A través de los textos, puede observarse que la prosodia semántica de “new markets” se asocia con elementos positivos, por ejemplo: “profits”, “diversification”, “impetus” o “help”.

If there is one big advantage that UK businesses have over other foreign competitors when it comes to accessing NEW overseas MARKETS it is 'Brand Britain'.

For the last decade the UK business, which generates two-thirds of group profits, has been the cash cow that funded diversification into NEW countries and MARKETS such as banking.

In fact, working out how things can be done more efficiently and which customers should be prioritised can not only lead to a leaner SME, it can provide the impetus to enter NEW MARKETS .

Innovative companies in emerging countries that are still in their infancy have the potential to catalyse NEW MARKETS and help address the youth unemployment challenge.

Volkswagen is growing so quickly in China and other relatively NEW MARKETS it is close on the heels of its two longtime rivals.

Much of the optimism about industry has been based on the idea that the cheaper pound will make British goods more competitive and coax manufacturers into exploring NEW MARKETS .

### Cuarto intervalo de estudio

La etapa final de análisis corresponde al año 2015, en el que la media de las puntuaciones de sentimiento de las oraciones que incluyen “markets” descienden al umbral de la negatividad hasta los 29,77 de puntuación media, mientras que su frecuencia relativa se catapultaba en tan solo un año hasta una FQ1000 de 0,61. Como ocurrió en el intervalo segundo de estudio, “markets” vuelve a convertirse en una palabra evento que se pone en relación con unidades léxicas que reflejan la turbulencia de la actualidad de ese año. De hecho, en la clasificación de oraciones más significativas solo se encuentran unidades léxicas de carácter neutro junto con dos coocurrencias con sentimiento negativo representativas de los eventos que ocurrieron ese año: “rig” y “turmoil”. Más abajo se puede observar la lista de las colocaciones más significativas del año:

- Sustantivos específicos: *stock, exchange, capital, currency, world, commodity, equity, analyst, export, investor, bond.*
- Adjetivos específicos: *financial, global, foreign, international.*
- Verbos específicos: *emerge, send, fall.*

La primera unidad de estudio es “rig” y hace referencia a una serie de acciones fraudulentas relacionadas con dos casos de manipulación en el libre mercado: la del “Libor”<sup>58</sup> y el ajuste artificial de las tasas de cambio de divisas. Seis años después de la quiebra del sistema financiero y del rescate multimillonario de la banca con fondos públicos, se demostró que los grandes poderes bancarios rompían las leyes de libre competencia y manipulaban sistemáticamente los tipos de interés para obtener ganancias de las operaciones, o para dar la impresión de que eran más solventes de lo que eran.

En mayo de 2015 las autoridades de arbitraje estadounidenses y británicas multaron a seis de los mayores bancos del mundo con 6.357 millones de dólares por vulnerar las leyes antimonopolio. Entre los efectos hacia la población, la manipulación de las tasas de interés hacía que las mensualidades de las hipotecas *subprime* fueran más caras de lo que debieran ya que las tasas de interés se alteraban artificialmente cada primero de mes para inflar las ganancias.

---

<sup>58</sup> El *Libor* (*London Interbank Offered Rate*) es una tasa de interés calculada mediante la presentación de estimaciones emitidas por los principales bancos.

2015		
<i>Palabra</i>	<i>Cooc</i>	<i>logDice</i>
financial-j	301	10,545
stock-n	278	10,833
emerge-v	221	10,958
global-j	163	9,933
foreign-j	88	9,635
exchange-n	73	9,378
<b>rig-v</b>	46	8,987
capital-n	61	8,933
currency-n	57	8,860
world-n	77	8,744
<b>turmoil-n</b>	37	8,657
commodity-n	37	8,478
international-j	39	8,424
equity-n	36	8,402
analyst-n	49	8,371
export-n	37	8,356
investor-n	58	8,321
send-v	36	8,309
fall-v	61	8,305
bond-n	36	8,302

**Tabla 37** Principales colocaciones de “markets” durante el cuarto intervalo

Estos eventos quedan recogidos inmediatamente en la prosodia de “markets”, a pesar del aumento del uso del término en el corpus la negatividad no se diluye y la colocación del término de estudio con “rigged” adquirió gran significancia (*logDice* de 8,987). Si bien “rigging” tiene un sentimiento negativo que queda recogido por *Lingmotif*, se puede observar también que en su vecindario destacan unidades de carácter negativo como “fined”, “penalties” o “accusations”. En el año 2015, hablar de los mercados se relaciona claramente con manipulación y falta de libre competencia. A continuación se pueden observar algunos ejemplos aleatorios:

It was among six major banks hit with £2.6bn of penalties in November for RIGGING the foreign exchange MARKETS , where £3.5tn changes hands each day.

It has already been fined for RIGGING interest rate and currency MARKETS and faces nearly £2bn of penalties for the way it sold mortgage bonds in the runup to the 2008 banking crisis.

It also comes after a year when the industry has faced big fines for RIGGING financial MARKETS .

He was docked £500,000 for control failures at the bank, including accusations of RIGGING the foreign exchange MARKETS .

RBS is braced for scrutiny of its pay deals following last year's £400m fine for RIGGING foreign exchange MARKETS and because of outstanding payments that were promised to chief executive Stephen Hester when he left in 2013.

The payouts come as the bank awaits a fine for its role in RIGGING foreign exchange MARKETS for which it has set aside £1.25bn and amid reports that the banks is poised to a settle a civil case in the US, alongside Citi, for £550m.

Finalmente, al igual que en 2008 y 2011, el sustantivo con sentimiento negativo “turmoil” vuelve a ser una colocación significativa en el contexto periodístico de la unidad “markets”. El cambio en la orientación semántica del contexto de “markets” durante el año 2015, así como el aumento de su frecuencia tiene que ver con varios eventos relacionados con los mercados internacionales. Fueron cinco los factores: la turbulencia económica a partir de la elección de Alexis Tsipras en Grecia, el fin de la paridad euro-franco suizo, un gran número de adquisiciones y fusiones entre multinacionales y el desplome de las acciones en la bolsa de Shanghai. Los ejemplos extraídos del corpus reflejan la correlación del uso de “markets” con los eventos históricos expuestos y, a su vez, su polaridad queda reflejada tanto en el sentimiento de las oraciones como en las colocaciones. Por poner ejemplo, si bien en el segundo intervalo se observó una tendencia positiva en el uso de “markets” en relación con los buenos resultados económicos en los mercados asiáticos, en este intervalo ocurre lo opuesto. En el vecindario de “markets” y “turmoil” pueden observarse múltiples elementos negativos tanto de naturaleza específica como pertenecientes al dominio general: “casualties”, “surge”, “plunging” o “collapse”.

The People's Bank of China's last rate cut in August triggered TURMOIL in world MARKETS after Beijing combined the decision with a 2% reduction in the yuan's value.

West Ham FC sponsor Alpari UK has become one of the biggest casualties of Thursday's TURMOIL on currency MARKETS sparked by the Swiss central bank abandoning its attempt to peg the franc against the euro.

Economists warn the pickup in UK exports could prove short-lived given a strong pound and the recent TURMOIL in global MARKETS sparked by China's shares rout.

Janet Yellen, the Fed's chair, made clear that its recent decision to delay a long-planned increase in rates was a result of the TURMOIL in emerging MARKETS , in particular China.

he SNB's announcement on 15 January that it had scrapped its cap of SFr1.20 to the euro – the centrepiece of its monetary policy since September 2011 – unleashed a surge in the value of the currency and TURMOIL on the foreign exchange MARKETS .

Other losses were set off by a domino effect of macroeconomic events – including surprise currency moves by central banks in China, Switzerland and elsewhere, as well as plunging oil prices, the collapse of commodity prices and TURMOIL in emerging MARKETS .

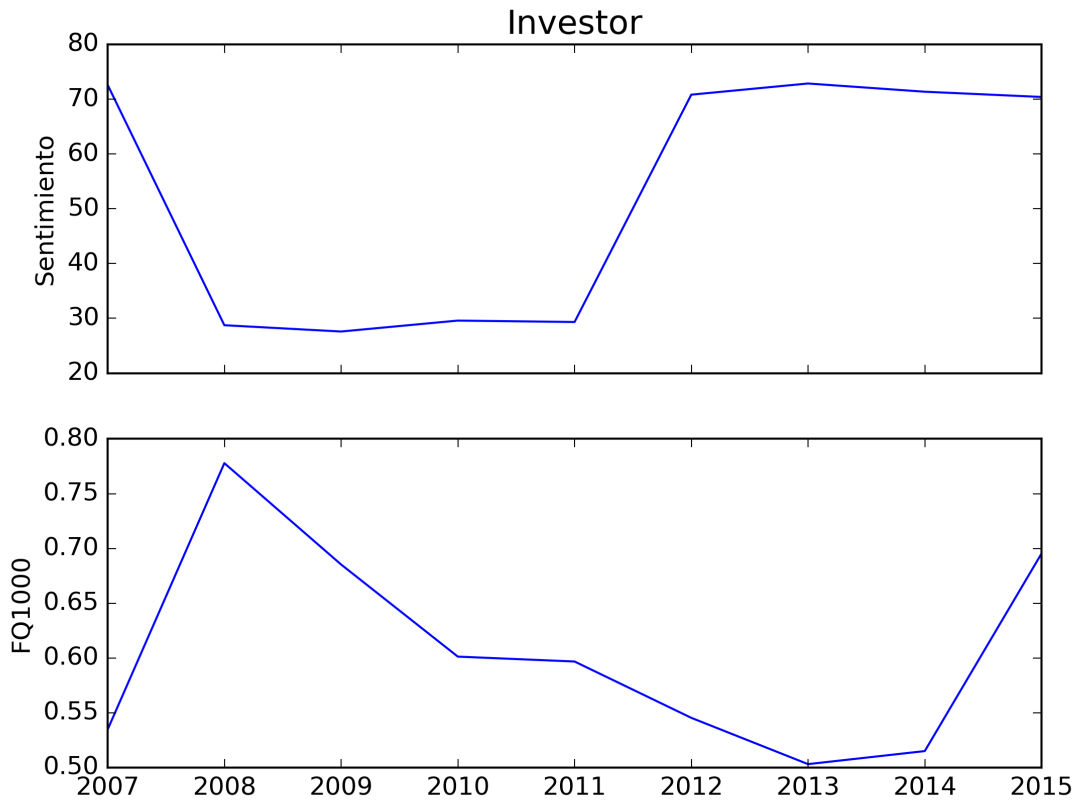
## 7.6 Término 6: *Investor*

El sexto término de estudio es “investor” (“inversor”), frecuente en el corpus y de uso común en la prensa económica. Tras el colapso de la banca, los inversores grandes y pequeños fueron protagonistas en el discurso relacionado con la crisis y su relación con la difusión de las opiniones en las noticias económicas. Por un lado, está la espantada de los grandes poderes capitalistas producida por el congelamiento del crédito a partir del año 2008 y, por otro, las dificultades para invertir de los pequeños profundizaron en la situación de pánico multiplicando la caída de los precios de inversiones antes de que estas siguiesen cayendo masivamente. Durante los años de la crisis, no fueron solo los inversores individuales sino también los inversores institucionales quienes estaban afectados por el desempeño de las noticias al no vender sus acciones con suficiente celeridad. El interés del estudio de la evolución del sentimiento y la prosodia semántica de la unidad “investor” recae exactamente ahí: los inversores eran observados por la prensa, los políticos y el público en general como un ente abstracto esencial para la estabilidad que debía mantenerse en calma y, en muchos casos, se tomaron determinadas medidas económicas condicionadas por el intento de evitar las posibilidades de huida de estos. Una revisión rápida de los sustantivos y verbos que modifican el término “investor” en *Word Sketch* muestra una gran cantidad de elementos léxicos relacionados con este sentimiento y la volatilidad de los inversores en esta época: “confidence”, “sentiment”, “appetite”, “revolt”, “rebellion”, “concern”, “fear” o “protection”.

El siguiente paso consiste en observar los factores que condicionaron el movimiento de la orientación semántica de los valores contextuales del término “investors”. A través de la gráfica de medias de puntuación de sentimiento (Figura 22) y los datos textométricos (Tabla 38) se pueden observar tres etapas claramente diferenciadas en las tendencias de la evolución del sentimiento del contexto del término: (1) una etapa inicial con una media de valores de sentimiento positivos en el 2007, correspondiente a la etapa previa al “credit crunch” y sus consecuencias, (2) un descenso repentino de las puntuaciones de sentimiento en el año 2008 que se mantiene hasta el año estable hasta el año 2011 y (3) una etapa final entre 2012 y 2015 que vuelve a medias positivas de sentimiento.

El primer intervalo de estudio (2007) tenía lugar en el escenario previo a la explosión de la crisis y la agregación de puntuaciones de sentimiento anuales mediante la media ofrecía una puntuación de sentimiento positiva (72,62) para las 962 oraciones que

incluyeron el término “investor” durante ese año. Durante ese periodo se produjo una caída repentina de las puntuaciones de sentimiento, manteniéndose las puntuaciones medias por debajo de la franja de los 30 puntos (28,74 en 2008, 27,59 en 2009 y 29,59 en 2010). En el tercer periodo se produjo un ascenso de alrededor de cuarenta puntos alcanzando unos niveles estables de positividad en el sentimiento medio de las oraciones (70,82 en 2011, 72,86 en 2012, 71,36 en 2013 y 70,41 en 2014). ¿Qué ocurre aquí? Mientras que en 2007 la cifra media de elementos negativos era de 0,9696, esta ascendió repentinamente a 1,2105 en el siguiente periodo de estudio, manteniéndose estable durante ese tramo. El ascenso de las puntuaciones en el año 2012 se justifica con un aumento de elementos positivos en las oraciones, pasando de una media de ítems positivos de 1,1622 en 2011 a 1,2695 en 2012 y un pico de 1,3886 en 2013.



**Figura 22** Gráfico longitudinal de las medias anuales de las puntuaciones de sentimiento para “investor” y la frecuencia relativa del término por cada 1000 tokens

En relación con las frecuencias relativas, al igual que ocurre con el análisis de muchos de los elementos estudiados aquí como “credit” o “housing”, en 2008 se produce un ascenso súbito en la frecuencia con el estallido de la crisis, alcanzando un pico de 0,78



en 2008, en comparación con una FQ1000 de 0,53 en 2007. Este ascenso corresponde, una vez más con un descenso de las puntuaciones de sentimiento. Desde entonces se produce una tendencia a la baja cayendo suavemente hasta el mínimo histórico en 2013 (FQ1000 0,50), cuando además se alcanza la puntuación máxima de sentimiento del estudio. A partir de entonces se produce de nuevo una tendencia ascendente en la frecuencia hasta acelerarse a final del estudio en 2015 (FQ1000 0,70).

<i>Año</i>	<i>N</i>	<i>FQ1000</i>	<i>Sentimiento</i>	<i>pos_items</i>	<i>neg_items</i>	<i>neu_items</i>	<i>none_items</i>
2007	920	0,53	72,62	1,09	0,97	0,94	9,89
2008	1444	0,78	28,74	1,07	1,21	0,95	9,54
2009	1365	0,69	27,59	1,04	1,25	1,02	9,77
2010	1351	0,60	29,59	1,20	1,28	1,13	10,46
2011	1615	0,60	29,33	1,16	1,30	1,13	10,02
2012	1358	0,55	70,82	1,27	1,19	1,07	10,74
2013	1809	0,50	72,86	1,39	1,10	1,13	10,51
2014	2021	0,52	71,36	1,34	1,18	1,18	10,62
2015	3018	0,70	70,41	1,25	1,18	1,08	10,38

**Tabla 38** Resumen estadístico del análisis de las oraciones que incluyen “investor”

### Primer intervalo de estudio

La primera fase de análisis corresponde a la época inmediatamente previa a la explosión de la crisis (año 2007) y las oraciones analizadas priman elementos léxicos de carácter especializado con un sentimiento neutro con los que forman unidades estables (como es el caso de “fund” o “institutional”) (Tabla 39). El análisis realizado mediante *The Sketch Engine* proporcionó un listado de colocaciones que disponen la observación a simple vista de las relaciones léxicas, así como con las distintas categorías gramaticales que coocurren con el término objetivo, como es el caso de las siguientes:

- Sustantivos genéricos: *group*.
- Sustantivos específicos: *fund, analyst, equity, stock, security, share, pension, money, return*.
- Adjetivos genéricos: *big*.
- Adjetivos específicos: *activist, foreign, institutional, individual, private*.
- Verbos genéricos: *attract, tell*.
- Verbos específicos: *sell*.

<i>palabra</i>	<i>2007</i>	
	<i>Cooc</i>	<i>logDice</i>
fund-n	51	9,503
institutional-j	20	9,380
foreign-j	24	9,244
private-j	31	9,140
individual-j	15	8,812
analyst-n	24	8,794
activist-j	13	8,784
equity-n	20	8,779
stock-n	25	8,764
security-n	16	8,606
share-n	28	8,602
tell-v	20	8,556
pension-n	13	8,484
big-j	24	8,340
group-n	23	8,318
<b>attract-v</b>	<b>11</b>	<b>8,298</b>
money-n	20	8,233
sell-v	25	8,177
among-i	15	8,172
return-n	11	8,134

**Tabla 39** Principales colocaciones de “investor” durante el primer intervalo

Una revisión de los resultados de la búsqueda aleatoria de “investor” en el corpus para el año 2007 devuelve oraciones de dominio específico con una connotación generalmente positiva sin que puedan encontrarse eventos ni tendencias históricas que justifiquen lo contrario. Para ilustrar esto, abajo se pueden observar algunos ejemplos extraídos del corpus:

One area where IGM's experience might be helpful for Putnam is INVESTOR retention.

INVESTOR forces ethics on to Tesco agenda

Barratt has secured irrevocable acceptances from INVESTORS holding 37.7% of the shares - Mr Wilson and his family, and Wilson Bowden's directors - and any counter-offer would have to be 10% higher than Barratt's agreed bid.

Lone Star remains a major INVESTOR in Japan, where it has acquired everything from banks to golf courses.

G.E.'s stock has been flat for years, but has recently moved higher, nearing the \$40 mark, partly because of INVESTOR enthusiasm for the company's move into markets like India.

El único término manifiestamente positivo en este dominio es “attract”. Este elemento hace referencia a las medidas que las empresas e instituciones realizan para fomentar el aumento de inversores externos. La prosodia semántica del término es positiva pues, como puede observarse en las oraciones aleatorias extraídas del corpus, en





su vecindario suelen encontrarse elementos positivos o neutros como “transparent”, “deals”, “high-paying”. Algunos de estos ejemplos son los siguientes:

Schwab, the founder of the company, says he wants to focus on lower, more transparent fees in order to ATTRACT younger INVESTORS .

At the same time, it serves to ATTRACT tax-exempt INVESTORS like pension funds and endowments, as well as foreign investors, two of the most active groups investing in hedge funds today.

Those deals ATTRACTED thousands of retail INVESTORS , not considered the smart money, like moths to a flame into Sirius stock.

MannKind, which has spent about more than \$700 million on its insulin, has ATTRACTED some prominent INVESTORS, like Legg Mason's marquee stock picker, Bill Miller.

Structured as a real estate investment trust, New Century ATTRACTED small INVESTORS who held the stock for its high-paying dividend.

He bought a house of his own to rent out and more recently has been thinking about how to ATTRACT INVESTORS to develop dormitory-style housing.

### Segundo intervalo de estudio

Este intervalo de estudio coincide con los años de crisis relativos a la explosión de la burbuja inmobiliaria y a la congelación del crédito. Durante estos cuatro años primaron las oraciones con una puntuación de sentimiento negativo. Aparte de las distintas colocaciones de carácter específico propias del término a lo largo de todo el intervalo de estudio como “money”, “institutional” o “fund”, aparecen algunas unidades léxicas de carácter no específico con una connotación negativa como es el caso de “mislead” o “concern”. Además, como ocurrió con “markets”, durante estos cuatro años aparece el sustantivo “confidence” el cual es recurrente hasta el fin de este y que se estudiará detenidamente más abajo. A continuación, se ofrece la lista de colocaciones clasificadas por categoría gramatical y por año (Tabla 40).

- Sustantivos genéricos: *concern, confidence* (4), *group* (3), *year*.
- Sustantivos específicos: *bank, bond* (2), *cash, company, equity* (2), *firm, fund* (3), *market* (3), *money* (3), *return* (3), *risk, security, share* (4), *stock* (3).
- Adjetivos genéricos: *big* (4).
- Adjetivos específicos: *financial* (2), *foreign* (2), *hedge, institutional* (4), *private* (2), *private, short-term*.
- Verbos genéricos: *attract, buy* (4), *include* (2), *look, lose, make, mislead, pay, put, raise, seek, sell* (3), *take, tell* (3).

La primera unidad léxica de estudio en el corpus es “confidence”. Al igual que ocurría con “markets”, los inversores eran interpelados como una amalgama indefinida

de capitales que no se atrevían a poner en marcha la maquinaria de circulación de capital. Los inversores se convirtieron en un actor principal del discurso público a partir de eventos como “*the collapse of Lehman Brothers last month*” y necesitaban acciones políticas que les permitieran calmar el estrés producido (“regain investors’ confidence”, “political uncertainty”, “dire economic challenges”) por el shock económico (p. ej., “downgraded it’s rating”). En todos los casos observados donde interviene la palabra “confidence”, esta es usada en contextos con una polaridad altamente negativa y en su vecindario se encuentran unidades como “problems”, “struggle”, “plummeted” o “waned”. Abajo se pueden observar algunos ejemplos aleatorios:

Analysts estimate that Morgan Stanley has more than \$100 billion in capital, but the firm has struggled to regain INVESTORS' CONFIDENCE since the collapse of Lehman Brothers last month.

Boeing announced on Wednesday a third delay in the delivery of its 787 Dreamliner, but the company's shares rallied, perhaps indicating INVESTOR CONFIDENCE that the company has a grasp on the problems that have plagued the plane, which is critical to its future.

The universal success of the combined value-momentum approach "greatly increases our statistical CONFIDENCE that INVESTORS would do well to focus on both strategies together rather than separately," Professor Moskowitz said.

A slump in INVESTOR CONFIDENCE accelerated on Thursday, as stocks fell over 4% and the rupee dropped by 1.3% by midday due to fears of political uncertainty and dire economic challenges, including a double-digit inflation rate.

HBOS shares plunge as INVESTOR CONFIDENCE wanes Shares in Britain's biggest mortgage lender plummeted today after Standard & Poor's downgraded its rating to AA- HBOS signs.

El verbo “mislead” es una unidad del dominio general con una polaridad negativa en todos los contextos y apareció en la lista de las colocaciones más significativas en el año 2009. Tras el colapso de las grandes corporaciones financieras se empezaron a destapar casos de maquillaje de rendimiento de cuentas llevados a cabo por estas corporaciones con el fin de atraer fraudulentamente a los potenciales inversores. Por poner uno de los múltiples ejemplos ocurridos ese año, el 31 de marzo de 2009 el New York Times abrió su sección de negocios con la noticia de que Grecia multó a Citigroup con 1,3 millones de dólares. Grecia acusó a Citigroup de engañar a los inversores en la venta de una serie de productos financieros de Lehman Brothers tras una demanda colectiva presentada por 92 inversionistas privados griegos. Los productos que habían sido ofrecidos a los inversionistas de Citibank exigían planes de inversión de capital garantizado de riesgo cero (Carassava, 2009).

<i>Palabra</i>	<i>2008</i>		<i>2009</i>		
	<i>Cooc.</i>	<i>logDice</i>	<i>Palabra</i>	<i>Cooc.</i>	<i>logDice</i>
money-n	59	9,251	fund-n	63	9,531
company-n	52	7,782	institutional-j	31	9,42
share-n	52	8,984	private-j	41	9,284
buy-v	41	8,733	money-n	48	9,032
sell-v	41	8,633	foreign-j	26	8,893
bank-n	39	8,102	share-n	41	8,832
institutional-j	37	9,598	buy-v	35	8,606
market-n	37	7,933	cash-n	24	8,559
stock-n	34	8,688	group-n	32	8,552
big-j	33	8,423	return-n	20	8,486
include-v	32	8,303	<b>mislead-v</b>	15	8,381
foreign-j	32	9,149	big-j	32	8,357
year-n	31	6,991	put-v	23	8,303
make-v	30	7,578	<b>lose-v</b>	22	8,294
firm-n	27	8,111	stock-n	21	8,254
financial-j	27	8,026	<b>attract-v</b>	15	8,232
take-v	26	7,68	financial-j	29	8,176
pay-v	26	8,11	confidence-n	14	8,140
tell-v	26	8,447	market-n	35	8,129
confidence-n	26	8,941	security-n	15	8,032
<i>Palabra</i>	<i>2010</i>		<i>2011</i>		
	<i>Cooc.</i>	<i>logDice</i>	<i>Palabra</i>	<i>Cooc.</i>	<i>logDice</i>
institutional-j	38	9,722	institutional-j	33	9,291
fund-n	64	9,601	fund-n	56	9,273
bond-n	33	9,219	foreign-j	33	8,98
private-j	34	8,89	buy-v	51	8,912
share-n	38	8,603	private-j	37	8,891
short-term-j	18	8,586	confidence-n	23	8,542
money-n	32	8,581	sell-v	41	8,514
include-v	46	8,554	tell-v	31	8,444
buy-v	34	8,528	share-n	39	8,408
tell-v	29	8,525	bond-n	24	8,406
foreign-j	21	8,517	group-n	40	8,387
raise-v	24	8,395	stock-n	25	8,251
sell-v	33	8,362	market-n	53	8,245
seek-v	19	8,283	big-j	33	8,085
big-j	32	8,266	equity-n	17	8,034
return-n	17	8,259	return-n	17	8,016
group-n	31	8,258	against-i	22	7,996
hedge-j	15	8,252	risk-n	20	7,942
confidence-n	15	8,231	look-v	24	7,926
equity-n	18	8,22	<b>concern-n</b>	17	7,904

**Tabla 40** Principales colocaciones de “investor” durante el segundo intervalo

Hablar de “investors” evocaba ese año la referencia del fraude a los mismos. La revisión del vecindario de “mislead” e “investors” evidencia rotundamente la negatividad del sentimiento en el contexto de estas. Entre otras unidades, en las oraciones extraídas del subcorpus 2009 pueden encontrarse sustantivos como “precariousness”, “dubious” o “falsified”. A continuación, algunos ejemplos:

The longtime leader of the brokerage firm is at the center of a federal investigation into whether Lehman executives MISLED INVESTORS about the state of the company.

The lawsuits, including one filed by New York State's comptroller, say Mr. Kurland was well aware of the risks, and even MISLED Countrywide's INVESTORS about the precariousness of the company's portfolio, which grew to \$463 billion in loans, from \$62 billion, three times faster than the market nationwide, during the final six years of his tenure.

Among those arrested were two former Bear Stearns hedge fund managers – Ralph Cioffi and Matthew Tannin – who were accused of contributing to the Wall Street firm's collapse by MISLEADING INVESTORS about a fund packed with dubious mortgage-backed securities.

His news hook: The Securities and Exchange Commission had fined G.E. \$50 million on charges of MISLEADING INVESTORS .

Federal regulators are seeking more than \$22 million from the former head of Kmart, who was found liable by a jury for MISLEADING INVESTORS about the company's finances before a bankruptcy filing in 2002.

Conaway, the former chief executive of Kmart, with MISLEADING INVESTORS and concealing information from Kmart directors.

The documents accuse DuPont of MISLEADING INVESTORS about certain product capabilities, as well as involvement in what Monsanto has said were several falsified letters to lawmakers and others who criticize Monsanto.

El término “risk” (“riesgo”) es la posibilidad de que el rendimiento obtenido de una inversión sea diferente de lo esperado teniendo en cuenta, también, el tamaño de la diferencia. Esto incluye la posibilidad de perder parte o la totalidad de la inversión original. Durante el año 2011, “risk” se convierte en una colocación significativa de “investor”. En relación con *LexiEcon*, este término es potencialmente negativo por lo que está marcado como neutro en el diccionario y las colocaciones con una direccionalidad ascendente están consideradas como negativas (p. ej., “increase the risk”). Durante estos años, el riesgo y sus indicadores (p.e. “risk premium”: “prima de riesgo”) estaban muy presentes en el léxico relacionado con los inversores. Si bien las inversiones de riesgo ofrecen mayores beneficios, durante estos años de turbulencia estos no eran especialmente deseables, como se puede observar en una muestra extraída del corpus. Entre los elementos léxicos que se encuentran en su contexto se encuentran muchas referencias a la crisis (“credit crunch”), verbos específicos como “devalue” o “disinvested” y sustantivos como “concerns”.

The credit crunch has also reduced the RISK appetite of many INVESTORS , who were now more willing to invest in funds closer to home, Friedman said.

High debt with the potential to devalue its currency present high RISKS for bond INVESTORS .

It started the practice in 2001 but really accelerated it in 2007 and early 2008, when INVESTORS belatedly discovered there were RISKS to high leverage ratios.

Yet unlike Treasuries, mortgage-backed securities guaranteed by all three of these agencies are exposed to many of the vagaries of the mortgage market, especially the RISK that INVESTORS will refinance their mortgages.

According to the Investment Company Institute, which surveys 4,000 households annually, the appetite for stock market RISK among American INVESTORS of all ages has been declining steadily since it peaked around 2001, and the change is most pronounced in the under-35 age group.

"Some [responsible investors] have disinvested because of concerns over stakeholder-related RISKS ," said ethical INVESTORS group Eiris in a document recently presented to the company.

En el año 2010 ganaron relevancia las “hedge funds” (“fondos de cobertura”), un producto financiero similar a un fondo de inversión con el objetivo de maximizar la rentabilidad a cualquier precio. Esto es, utilizando todas las posibilidades de inversión que ofrezca el mercado sin ningún límite, desligándose de los mercados. Un análisis introspectivo de “hedge funds” durante esta época prediría que colocaciones de este tipo tendrían una prosodia semántica negativa. Si bien era esperable que con la condena de fraude a Bernard Madoff a través de este tipo de productos el año anterior las oraciones relacionadas con “investor” y “hedge funds” se relacionasen con una polaridad negativa, esto no es así en todos los casos. Una muestra de las oraciones extraídas del corpus que incluyese esta colocación ofrece tanto oraciones neutras como oraciones negativas y ninguna de ellas hace referencia a un evento histórico o noticioso. Entre los elementos léxicos negativos encontrados en el vecindario están los siguientes “speculators”, “downturn” o “schemes”.

The circumstances of the defeat are likely to reopen a debate about the role of HEDGE FUNDS and other investors during takeovers.

The instruments, known as equity swaps, mimic ordinary shares and give investors like HEDGE FUNDS the benefits of stock ownership, including payments similar to dividends, without actually owning the shares.

All of your firms are involved in some form of proprietary trading, or using your own capital to make financial bets, not unlike HEDGE FUNDS and other private investors.

He said the Treasury had rigged the review process under pressure from backbench Labour MPs who refused to allow HEDGE FUND investors compensation.

But the watchdog said it would be too complicated to ban "short-termist' HEDGE FUND investors and other speculators from voting on whether a takeover should go through.

Alongside Wall Street banks and rich investors , HEDGE FUNDS that made huge profits betting on the downturn of the US housing market also benefited from the schemes.

### Tercer intervalo de estudio

El tercer y último intervalo de estudio coincide con una etapa diferente de la crisis. Durante estos cuatro años primaron las oraciones con una puntuación de sentimiento positivo. Las distintas colocaciones de carácter específico propias del término siguen presentes también a lo largo de todo el intervalo de estudio (“money”, “institutional” o “fund”). Aparecen en este intervalo algunas colocaciones no específicas de carácter positivo como “attract” o “safe”, al igual que vuelven a aparecer colocaciones con connotación negativa como es el caso de “concern” junto a nuevas como “worry”. A continuación, se ofrece la lista de colocaciones clasificadas por categoría gramatical y por año (Tabla 41).

- Sustantivos genéricos: *analyst*, *concern*, *confidence* (3), *group* (3), *meeting*, *money* (3), *pay* (2).
- Sustantivos específicos: *bank*, *bond* (3), *company*, *entrepreneur*, *fund* (4), *interest*, *market*, *return* (3), *risk*, *share* (4), *shareholder*, *stock* (3)
- Adjetivos genéricos: *big*, *foreign* (3), *individual* (2), *institutional* (4), *international*, *large*, *long-term*, *potential* (3), *private* (4), *retail* (2), *safe*.
- Adjetivos específicos: *activist*.
- Verbos genéricos: *attract*, *back* (2), *buy* (4), *hold*, *sell* (2), *tell* (4), *vote* (2), *worry* (2).

Es pertinente aquí realizar una revisión general de los textos del corpus que incluyen al término “investor” durante este periodo. Una búsqueda aleatoria en los textos devuelve una muestra que evidencia un uso del término con un relato propio del dominio específico en el que se relata la actualidad de las empresas, sin observarse referencias a la crisis. Se puede considerar entonces que, en mayor medida, para este intervalo la palabra “investor” no se trataría de una palabra evento en los términos propuestos por Moirand (2007, 2016). Asimismo, una observación de la prosodia semántica del término evidencia un conjunto de colocaciones positivas. Por ejemplo, pueden observarse adverbios como “positively”, adjetivos como “successful” o verbos específicos del dominio como “grow”. Se pueden observar algunos ejemplos a continuación:

INVESTORS reacted positively to the trading update, bidding Burberry stock up by 13 percent Thursday in London to close at £11.36, or \$18.17.

The challenge will be to uncover the asset classes, the models of investing, the distribution channels to offer investment products, and the technologies that will enable the next generation of INVESTORS to be successful.

But one thing every INVESTOR knows about car insurance is that very few companies grow at 22% for ever.

A source close to the company told Bloomberg on Monday that the INVESTORS included Will Smith, Len Blavatnik and American Express.

Pimco has more bond INVESTORS than anybody else.

Barclays Barclays finance chief investigated by FSA over fundraising In wake of Libor scandal, Chris Lucas is under FSA spotlight over £7.3bn funds raised from Middle East INVESTORS in 2008.

It is thought that most of Bumi's blue-chip institutional INVESTORS support Rothschild as co-chairman.

En 2012 aparece la colocación “safe” con un significativo coeficiente *logDice* de 8,226. “Safe” como adjetivo es considerado una unidad léxica con sentimiento positivo y como tal consta en varios ejemplos. Sin embargo, esta colocación en la mayoría de las ocasiones hace referencia a “safe havens” que hace referencia a una inversión que se espera que retenga o aumente su valor en tiempos de turbulencia en el mercado, p. ej., el oro, o un espacio físico con mejores condiciones económicas de manera que limiten su exposición a las pérdidas en caso de desaceleración del mercado. Esto es también considerado habitualmente positivo (p. ej., se puede observar “cut the cost”, “cheaper”) si la ubicación que se referencia es ese “safe haven” como es el caso de “UK” o “Germany” en los ejemplos. No obstante, si la entidad referenciada es un país con dificultades económicas en el vecindario aparecerán palabras negativas relativas a la situación (p. ej., “collapse” o “unstable”). A continuación, se ofrecen algunos ejemplos extraídos del corpus:

Most recently he brought together Glencore and Xstrata in a \$90bn (£56bn) deal, but Hannam also pursued JP Morgan's smaller mining activities in Afghanistan, working with the Pentagon to demonstrate the country is SAFE for foreign INVESTORS .

The UK's status as a SAFE haven for INVESTORS has cut the cost of borrowing by the Treasury to a 300-year low.

INVESTORS again turned to SAFE HAVENS, fearing political chaos and economic collapse in Greece having knock-on effects for the global economy.

The increase was driven partly by a flight to SAFE HAVENS by INVESTORS in politically unstable countries, and partly by inflows from UHNW families based in rapidly developing economies.

Switzerland has come under seige during the euro crisis from foreign INVESTORS seeking a safe HAVEN.

Government costs of borrowing: German government Bunds rally - making it cheaper for Germany to borrow as INVESTORS seek a safe HAVEN.

INVESTORS looking for SAFE places to put their money today drove the price of 10-year UK government bonds higher, pushing yields down to 1.47%.



2012			2013		
<i>Palabra</i>	<i>Cooc.</i>	<i>logDice</i>	<i>Palabra</i>	<i>Cooc.</i>	<i>logDice</i>
institutional-j	40	9,779	institutional-j	54	9,794
bond-n	39	9,313	fund-n	56	9,142
foreign-j	36	9,304	private-j	36	8,64
private-j	38	9,095	return-n	30	8,58
fund-n	38	8,952	confidence-n	24	8,434
buy-v	38	8,754	bond-n	25	8,419
individual-j	21	8,708	<b>attract-v</b>	24	8,39
vote-v	17	8,417	potential-j	24	8,274
money-n	30	8,343	money-n	36	8,225
back-v	16	8,320	share-n	37	8,204
group-n	32	8,280	tell-v	31	8,144
<b>safe-j</b>	15	8,226	sell-v	39	8,106
retail-j	18	8,184	buy-v	33	8,094
<b>worry-v</b>	14	8,164	risk-n	28	8,092
tell-v	23	8,117	stock-n	24	8,071
share-n	28	8,093	entrepreneur-n	19	8,036
pay-n	17	8,088	large-j	34	8,003
return-n	15	8,042	international-j	20	7,89
stock-n	17	8,019	shareholder-n	21	7,83
hold-v	17	7,945	bank-n	40	7,787
2014			2015		
<i>Palabra</i>	<i>Cooc.</i>	<i>logDice</i>	<i>Palabra</i>	<i>Cooc.</i>	<i>logDice</i>
institutional-j	57	9,72	institutional-j	67	9,387
private-j	51	9,044	fund-n	108	9,382
fund-n	52	8,963	activist-j	58	9,186
buy-v	58	8,747	money-n	83	8,954
foreign-j	36	8,742	foreign-j	55	8,822
confidence-n	27	8,471	tell-v	64	8,596
group-n	54	8,442	private-j	53	8,593
share-n	54	8,411	group-n	76	8,558
pay-n	34	8,406	share-n	86	8,507
potential-j	29	8,361	individual-j	38	8,49
analyst-n	36	8,347	bond-n	44	8,46
vote-v	23	8,305	retail-j	46	8,455
tell-v	35	8,211	stock-n	56	8,422
back-v	22	8,189	market-n	115	8,384
meeting-n	27	8,186	confidence-n	37	8,319
sell-v	39	8,047	<b>long-term-j</b>	37	8,283
company-n	119	8,011	buy-v	53	8,251
return-n	21	7,957	<b>worry-v</b>	33	8,247
interest-n	29	7,954	potential-j	37	8,21
<b>concern-n</b>	23	7,948	<b>big-j</b>	67	8,171

**Tabla 41** Principales colocaciones de “investor” durante el tercer intervalo

La siguiente unidad de estudio es “confidence” que vuelve a aparecer en los rankings de colocaciones en el intervalo 2013-2015. Mientras que en el en el intervalo anterior se hablaba de una pérdida de confianza de los inversores, ahora en muchos



ejemplos se puede analizar la prosodia semántica esta en términos positivos, principalmente. En la vecindad de los términos pueden encontrarse varios elementos positivos como, por ejemplo: “restore”, “reasonable”, “inspires”, “renewed” o “solidified”.

The perception of arbitrary outcomes, and the message it sends about Japan's tolerance for bad corporate governance, is hurting INVESTOR CONFIDENCE , analysts say.

One of the biggest current challenges is figuring out how to restore INVESTOR CONFIDENCE during the long period it will take for new treaties strengthening the euro bloc's foundations to be approved and put into effect.

A basic requirement of investment markets is for INVESTORS to have a reasonable CONFIDENCE that they are not being defrauded, and – if it turns out they are – that the government will seek to punish those responsible.

Whatever the approach, it inspires CONFIDENCE among INVESTORS when organisations demonstrate knowledge of the field and how their intervention complements, or contrasts with, existing practice.

The German and French stock markets jumped more than 4% while the main market in New York soared in early trading as renewed CONFIDENCE among INVESTORS erased several weeks of losses.

Foreign direct investment, likewise, rose 8 percent last year to \$2 billion, from \$1.9 billion in 2011, as INVESTOR CONFIDENCE in the country has solidified since Mr. Aquino took office nearly three years ago.

El último término de estudio de esta sección es “long-term”, el cual está presente en el año 2015 con un coeficiente *logDice* de 8,283. Al contrario que los inversores de capital riesgo o las “hedge funds”, un inversor calificado como “long-term” es calificado positivo por la estabilidad económica que ofrece a las empresas o instituciones donde se realizan las inversiones. En todos los casos examinados el sentimiento general de las oraciones extraídas del corpus es positivo. Algunos ejemplos son los siguientes:

Long-term incentives will build certainty for LONG-TERM INVESTORS and avoid putative measures.

Not surprisingly, the high-sustainability organisations have more LONG-TERM INVESTORS – every company's desire – than their low-sustainability counterparts.

Vittorio Colao, Vodafone chief executive, said: "We are a committed LONG-TERM INVESTOR in India and we have made clear all along that we have faith in the Indian judicial system.

Meysman is currently consulting with INVESTORS about new LONG-TERM incentive arrangements as he is in the final stages of recruiting a successor to Fontenla-Novoa.

The departure of LONG-TERM INVESTORS does not always lead to lower trading volumes.

Fourth, the novel one – an INVESTORS' forum to allow LONG-TERM fund managers to agree on how to confront errant companies.

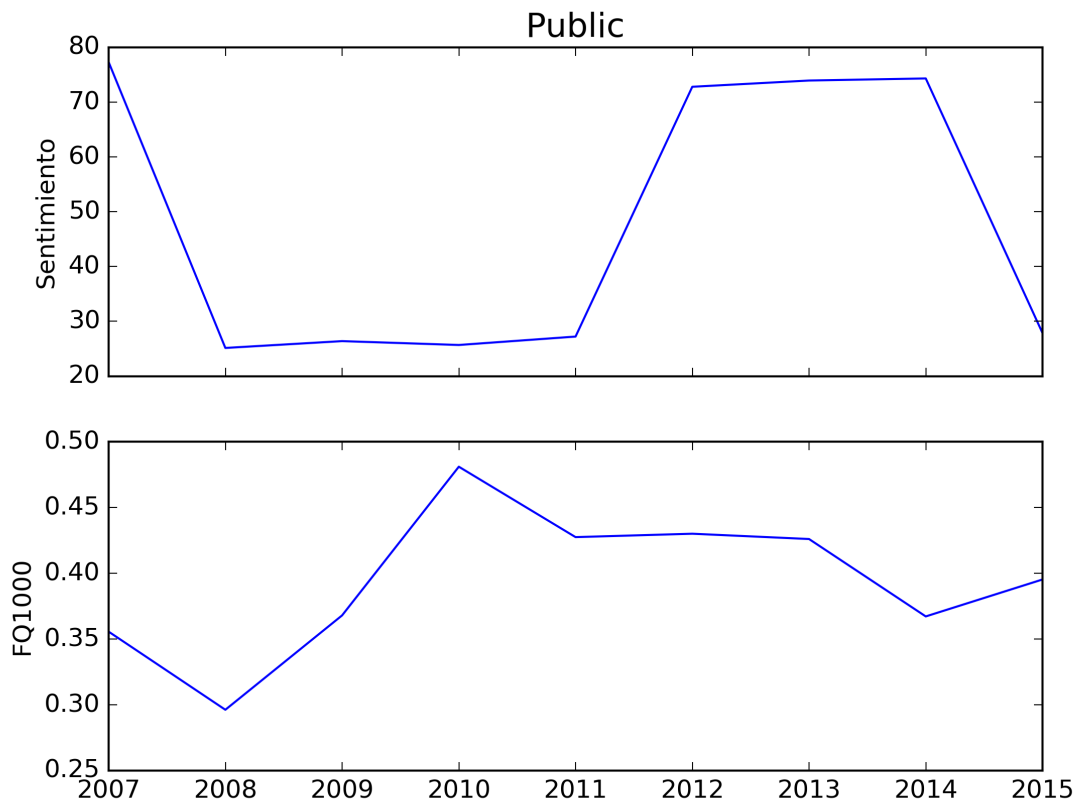
BP chief executive, Bob Dudley, said: "BP intends to be a LONG-TERM INVESTOR in Rosneft – an investment which I believe will deliver value for our shareholders over the next decade and beyond.

### 7.7 Término 7: *Public*

El séptimo término analizado en este trabajo es el adjetivo “public” (“público”). El interés de estudio de la evolución del sentimiento de este término recae en dos factores principales: (a) es el único término perteneciente a la categoría gramatical de los adjetivos y es de interés observar su comportamiento en, por ejemplo, comparación con los términos sustantivos además de, (b) la importancia de este término que referencia, a través de sus colocaciones, tanto a empresas del sector público como a los servicios públicos que estaban siendo recortados durante la crisis.

Con el inicio de la crisis, la deuda pública bruta aumentó de alrededor del 60% al 80% del PIB y los déficits públicos en la Unión Europea se triplicaron entre 2008 y 2010. A partir de 2009, los países de la UE comenzaron los planes (o directamente implementaron estrategias) para abordar el aumento de los déficits estatales. A pesar de que los gobiernos manifestaban una prioridad mantener el consumo y mitigar los efectos del desempleo, junto con la necesidad de estabilizar los mercados financieros, la respuesta de los gobiernos a las grandes dificultades provocadas por la crisis económica en muchos casos pasó por aplicar rígidas políticas de austeridad presupuestaria. También a partir de la crisis, muchos empleadores del sector público pasaron por alto los procedimientos tradicionales de negociación colectiva establecidos y los salarios de sus trabajadores se redujeron o congelaron. Esto provocó, por un lado, grandes bolsas de malestar ciudadano traducidas en protestas masivas y, a su vez, una bajada de la calidad de vida de la ciudadanía (Glasner, 2010).

El siguiente paso es la observación de la evolución de las puntuaciones de la orientación semántica del contexto del término en cuestión. A continuación, se observarán las gráficas de datos cuantitativos relativos al término, la serie temporal de medias de puntuación de sentimiento (Figura 23) y de los datos aportados por los indicadores estadísticos el análisis de los textos (Tabla 42). La media de las frecuencias relativas se mantiene estable durante todo el intervalo de estudios y oscila en un rango entre 0,30 en 2008 y 0,43 en 2012.



**Figura 23** Gráfico longitudinal de las medias anuales de las puntuaciones de sentimiento para el adjetivo “public” y la frecuencia relativa del término por cada 1000 tokens

Los gráficos de sentimiento permiten evidenciar claramente cuatro etapas en las tendencias de la evolución del sentimiento del contexto del término en cuestión:

- (1) Una primera etapa anterior a la crisis relativa al año 2007 con una media de sentimiento positivo (77,33).
- (2) Una segunda etapa entre 2008 y 2011 (25,14, 26,39, 25,69 y 27,22) con un descenso repentino y pronunciado, con una bajada de más de 50 puntos puntuaciones negativas. Para ilustrar la caída al umbral de la negatividad entre 2007 y 2008, el número medio de elementos negativos aumentó de manera muy significativa y pasó de 0,81 en 2007 a 1,06 en 2008.
- (3) Entre los años 2012 y 2014, los valores medios de sentimiento en las oraciones que incluyen “public” ascienden de manera repentina al umbral de la positividad (72,81, 73,95, y 74,33). La media de unidades positivas aumenta progresivamente, destacando el salto de 1,19 en 2013 a 1,31 en

2014. Por otro lado, las puntuaciones negativas pasaron de 1,27 en 2011 a 1,13 en 2012.

- (4) En último año de estudio, 2015, se produce una última bajada pronunciada del sentimiento de 74,33 en 2014 a 27,86 en 2015. Cuantitativamente, esto se justifica por un aumento de unidades negativas, pasando de 1,14 en 2014 a 1,35 en 2015.

<i>Año</i>	<i>N</i>	<i>FQ1000</i>	<i>Sentimiento</i>	<i>pos_items</i>	<i>neg_items</i>	<i>neu_items</i>	<i>none_items</i>
2007	612	0,36	77,33	0,90	0,81	13,47	12,43
2008	550	0,30	25,14	0,94	1,06	12,95	11,64
2009	733	0,37	26,39	1,08	1,20	13,83	12,15
2010	1081	0,48	25,69	1,16	1,44	14,41	12,76
2011	1157	0,43	27,22	1,19	1,27	14,63	12,76
2012	1071	0,43	72,81	1,15	1,13	14,55	12,49
2013	1532	0,43	73,95	1,19	1,09	14,98	12,70
2014	1441	0,37	74,33	1,31	1,14	15,43	13,25
2015	1716	0,40	27,86	1,28	1,35	15,21	13,17

**Tabla 42** Resumen estadístico del análisis de las oraciones que incluyen el adjetivo “public”

### Primer intervalo de estudio

La primera fase de análisis corresponde a la época en que la explosión de la crisis era aún incipiente (año 2007) y en las oraciones analizadas prevalecen elementos léxicos de carácter genérico con un sentimiento neutro con los que forman colocaciones estables (p. ej., “public relations” o “public offering”) (Tabla 43). El análisis realizado con *The Sketch Engine* devolvió una serie de colocaciones, junto con las distintas categorías gramaticales, que coocurren con “public”. Estas son las siguientes:

- Sustantivos genéricos: *affair, filing, health, interest, offering, pension, policy, relation, role, sector, service, space, statement.*
- Sustantivos específicos: *finance, firm, good, spending.*
- Adjetivos: *initial, private.*
- Verbos: *go.*

<i>Palabra</i>	<i>2007</i>	
	<i>Cooc.</i>	<i>logDice</i>
relation-n	54	11,202
offering-n	40	10,727
initial-j	29	10,305
go-v	61	9,477
policy-n	21	9,332
health-n	21	9,214
private-j	26	9,165
sector-n	13	8,979
spending-n	11	8,848
affair-n	9	8,674
filing-n	9	8,609
service-n	19	8,592
good-n	10	8,529
<b>interest-n</b>	13	8,401
role-n	9	8,236
statement-n	10	8,230
firm-n	20	8,218
space-n	7	8,100
finance-n	7	8,059
pension-n	7	8,037

**Tabla 43** Principales colocaciones de “public” durante el primer intervalo.

Dado el escaso número de colocaciones encontrado en el intervalo de estudio, es oportuno aquí hacer una revisión de la orientación semántica de una muestra de oraciones que incluyen el adjetivo “public” en el subcorpus del año 2007. Tras observar las muestras, se evidencia que una importante mayoría de las oraciones tienen una valencia neutra o positiva. Aquí, “public” coocurre con colocaciones consideradas neutras o, en algunos casos, positivas (p. ej., “vindicates”, “interest”, o “maverick”). Destaca aquí el uso de modificadores contextuales como como p. ej., “isn’t hostile”, “stem the bad press”, “no public outcry”, entre otras. Los ejemplos son los siguientes:

"We believe the verdict vindicates the serious PUBLIC interest in making sure that when insiders in a corporation deal with money entrusted to the shareholders, that they're not engaged in self-dealing, that they not break the law to benefit themselves instead of the shareholders," he added.

"The approach isn't hostile," said Olivier Labesse, the PUBLIC relations specialist hired by Mr. Arnault and Colony Capital to handle inquiries.

Opening the PUBLIC relations counterattack to stem the bad press that the revelations attracted internationally, the company's president said in an open letter to customers that the children who had been found to be embroidering decorations on toddlers' blouses for the company would be paid until they were of working age and then offered employment.

So far there has been no PUBLIC outcry from investors or internal revolts at Bear Stearns and Morgan over the fact that Mr. Cayne and Mr. Mack are staying on.

" He was similarly a maverick about his business dealings, giving Commerce the appearance of a piggy bank for insiders instead of a PUBLIC company.

In November, Hertz went PUBLIC at \$15 a share.

In some games we designed for workers in PUBLIC policy, we simulate the interactions between policy and daily life.

Como se verá también en el apartado 7.9, “public spending” (*logDice* 8,848) hace referencia al “gasto público”, es decir, los importes dedicados por las corporaciones gubernamentales al beneficio de la sociedad: desarrollo, educación, salud, obras públicas o seguridad. Se considera aquí que es este es un indicador económico potencialmente positivo y que, por ejemplo, una subida en el gasto público será positivo. Tras observar algunos ejemplos en el corpus, solo el segundo ejemplo del bloque tiene relación con los eventos históricos relacionados con la crisis y el resto de las oraciones tiene un sentimiento general positivo. La mayoría de las colocaciones observadas aquí tienen un sentimiento positivo, entre ellas se encuentran: “high”, “low interest rates”, “increase”, “doesn’t reduce”.

This is due in large part to high PUBLIC SPENDING and private consumption, fuelled by high oil prices and historically low interest rates.

However, he says Nicolas Sarkozy, or any other politician wanting to cut back public spending , will have a fight on their hands.

Ethnic diversity doesn't inevitably reduce SPENDING on PUBLIC goods.

PUBLIC SPENDING in law and order might not go down.

A Treasury spokesman said the government would be able to meet its fiscal rules while allowing SPENDING on PUBLIC services to increase by an average of 2% a year.

New York and California, for instance, are among the most highly taxed and most diverse states, although there is evidence that racial and ethnic tensions have whittled away at support for PUBLIC SPENDING .

En cuanto a la colocación con “sector” (*logDice* 8,979), principalmente el clúster “public sector” (“sector público”), esta se halla indistintamente en oraciones con sentimiento positivo (principalmente, las oraciones que no están en el contexto de los eventos de la incipiente crisis), mientras que en el contexto de la crisis “public sector” se presenta en un contexto negativo. Los ejemplos observados son los siguientes:

The corporate safety net of the 20th century is going away, and a fundamentally different private sector will require a fundamentally different PUBLIC SECTOR .

This business is very firmly positioned as an outsourced service provider for the PUBLIC SECTOR .

In truth, the deeper reality may be that command-and-control management lives on for the simple reason that, despite the rhetoric, 99 per cent of employees, whether in PUBLIC or private SECTOR , work in command-and-control organisations: in effect, centrally planned dictatorships that are set up to take orders from the CEO rather than the customer.

In Hungary, however, the unemployment rate rose, despite an increase in outsourcing jobs, to 8.2 percent from 5.8 percent, mostly because of government austerity measures to cut jobs in the PUBLIC SECTOR .

I had a great legal job in the PUBLIC SECTOR , but what I had learned about myself is that I liked being a lawyer and what I liked the most was negotiating and counseling," she said.

He said PUBLIC SECTOR IT costs £14bn a year, equivalent to 75 hospitals.

Las colocaciones con “service” (*logDice* 8,592) se encuadran en oraciones con sentimiento mayoritariamente negativo, junto a algunos ejemplos positivos. Sin embargo, ninguna de las oraciones parece estar contextualizada directamente con los eventos relacionados con la incipiente crisis económica. Entre las colocaciones que se observan están las siguientes, con sentimiento indistintamente positivo y negativo: “blocked”, “valued”, “increase”, o “effectively”. Algunos ejemplos revisados son los que se muestran a continuación:

Low- and middle-income families are not the only ones who have been harmed by our inability to provide valued PUBLIC SERVICES.

In the United States, trickle-down theory's insistence that a more progressive tax structure would compromise economic growth has long blocked attempts to provide valued PUBLIC SERVICES.

Many top earners would willingly pay higher taxes for PUBLIC SERVICES that promise high value.

A Treasury spokesman said the government would be able to meet its fiscal rules while allowing spending on PUBLIC SERVICES to increase by an average of 2% a year.

Most recently, Mr. Dusenberry worked as a volunteer on a team that created the "Generous Nation" PUBLIC SERVICE announcement that featured the line, "Don't almost give.

With its hard-to-fathom structure placed under the spotlight, the talk has been that potential financial problems could rebound not just on Macquarie shareholders but also citizens from dozens of countries which have key PUBLIC SERVICES effectively supplied by Macquarie.

### Segundo intervalo de estudio

El segundo intervalo de estudio es el que atañe a los cuatro años posteriores al estallido de la crisis (2008-2012) y, como consecuencia, las puntuaciones medias de las oraciones que incluyen el adjetivo “public” caen bajo el umbral de negatividad. Especialmente a partir de 2009, se pueden apreciar la aparición de colocaciones negativas, como son las referencias a los recortes del gasto en el sector público (“cut”) y, destacadamente, se observan las únicas colocaciones relacionadas con las emociones en todo este trabajo: “anger” y “outcry”.

El análisis realizado mediante *The Sketch Engine* (Tabla 42) ofreció una lista de colocaciones que permiten observar las relaciones con distintas palabras a lo largo de este

intervalo, así como las distintas categorías gramaticales que coocurren con el término objetivo. En este caso son las siguientes:

- Sustantivos genéricos: *affair* (3), *agency*, *anger* (3), *campaign*, *comment* (2), *disclosure*, *face*, *health* (3), *interest* (3), *job* (2), *offering* (4), *opinion*, *outcry*, *ownership*, *pension* (3), *policy* (3), *relation* (4), *safety*, *school* (2), *sector* (4), *service* (3), *support*, *test*, *transport* (3), *transportation*, *work*, *worker*.
- Sustantivos específicos: *cut* (3), *debt*, *deficit*, *finance* (3), *fund*, *pay*, *spending* (4).
- Adjetivos genéricos: *initial* (4).
- Adjetivos específicos: *private* (4).
- Verbos: *go*.

La primera colocación analizada será, al igual que en el intervalo anterior, la unidad “sector”. Tras extraer varias oraciones aleatorias del subcorpus correspondiente a este segmento temporal, se puede observar que la mayoría de las oraciones engloban un sentimiento negativo. Esta colocación coocurre principalmente con el sustantivo “cuts”, que hace referencia a los recortes presupuestarios que el sector sufrió durante esta época. Otras unidades encontradas en el vecindario también tienen un sentimiento negativo: “rig”, “conspiring” o “miserable”. Algunos ejemplos observados se pueden revisar a continuación:

More than 100 construction companies have been accused of conspiring to rig thousands of PUBLIC SECTOR contracts worth billions of pounds.

Unions have warned some staff with secure PUBLIC SECTOR status find their working conditions are made miserable and that they are deliberately rotated or demoted to spur them to resign.

"Foreign countries have strengthened national competitiveness with successful reforms in the PUBLIC SECTOR," the ministry said.

James Knightley of ING said: "The government will no doubt take this as a sign that the private sector can fill the gap created by PUBLIC SECTOR cuts, but with consumer confidence, hiring intentions surveys and housing activity data all softening, we remain cautious.

The PUBLIC SECTOR cuts, and their multiplier effects, taken in conjunction with other measures in the so-called "emergency" budget, would reportedly cost a total of 1.3m jobs.



2008			2009		
<i>Palabra</i>	<i>Cooc.</i>	<i>logDice</i>	<i>Palabra</i>	<i>Cooc.</i>	<i>logDice</i>
relation-n	46	11,112	relation-n	49	10,931
offering-n	22	10,062	sector-n	38	10,065
initial-j	19	9,798	spending-n	26	9,510
sector-n	22	9,531	private-j	28	9,340
health-n	17	9,203	<b>anger-n</b>	<b>13</b>	<b>9,145</b>
affair-n	11	9,097	offering-n	13	9,038
private-j	19	9,084	service-n	28	9,025
transport-n	9	8,891	initial-j	12	8,924
service-n	20	8,890	health-n	15	8,579
spending-n	11	8,707	finance-n	12	8,494
pension-n	10	8,703	policy-n	13	8,471
face-n	8	8,498	opinion-n	8	8,422
school-n	10	8,470	affair-n	8	8,374
interest-n	13	8,329	<b>outcry-n</b>	<b>7</b>	<b>8,368</b>
transportation-n	6	8,325	ownership-n	8	8,324
disclosure-n	6	8,251	<b>cut-n</b>	<b>10</b>	<b>8,214</b>
go-v	27	8,174	work-n	14	8,205
campaign-n	10	8,075	pension-n	10	8,160
comment-n	6	8,051	fund-n	17	8,085
agency-n	9	8,049	support-n	8	8,072
2010			2011		
<i>Palabra</i>	<i>Cooc.</i>	<i>logDice</i>	<i>Palabra</i>	<i>Cooc.</i>	<i>logDice</i>
sector-n	156	11,373	sector-n	150	11,146
relation-n	66	10,670	relation-n	55	10,437
spending-n	56	10,104	offering-n	47	10,221
offering-n	40	9,984	health-n	56	10,017
<b>cut-n</b>	44	9,775	initial-j	41	9,974
initial-j	35	9,770	private-j	45	9,521
finance-n	43	9,750	<b>cut-n</b>	36	9,492
interest-n	49	9,579	spending-n	32	9,227
private-j	42	9,380	service-n	47	9,120
service-n	45	9,310	pension-n	25	9,067
debt-n	34	8,956	transport-n	18	8,836
health-n	24	8,916	finance-n	22	8,712
<b>anger-n</b>	17	8,826	interest-n	27	8,614
safety-n	23	8,811	affair-n	14	8,501
policy-n	26	8,797	school-n	18	8,492
job-n	32	8,582	policy-n	20	8,256
comment-n	15	8,503	worker-n	20	8,253
test-n	15	8,485	job-n	27	8,233
deficit-n	17	8,480	<b>anger-n</b>	10	8,123
transport-n	13	8,371	pay-n	14	8,107

**Tabla 44** Principales colocaciones de “public” durante el segundo intervalo

La colocación con “service” (p. ej., *logDice* 9,120 en 2011) hace referencia a dos términos totalmente distintos: “public service announcement” (“comunicado público”) y “public services” (“servicios públicos”). Mientras que las colocaciones relacionadas con el primer término son neutras (como es el caso del primer ejemplo), las segundas suelen relacionarse con los recortes y, por tanto, con sentimiento negativo. Algunas de las colocaciones con sentimiento negativo son: “waste”, “distortion” o “poacher”. A continuación algunos ejemplos:

Visitors often see PUBLIC SERVICE ANNOUNCEMENTS in place of paid advertisements when they watch episodes and short clips.

It must be right that the WHO is transparent because there has been distortion of priorities of PUBLIC health SERVICES all over Europe, waste of huge sums of public money and provocation of unjustified fear.

Businesses benefit from PUBLIC SERVICES whether they make a profit or not – they should all make a contribution.

An increasing inefficiency of public-sector spending is one of the reasons why taxpayers in the United Kingdom feel that despite much higher taxes, they do not seem to have been rewarded with a commensurate improvement in PUBLIC SERVICES .

A new series of PUBLIC SERVICE ads, promoting greater involvement by fathers in their children's lives and aimed at American-Indian, Asian-American and Hispanic fathers, will be introduced today by the Advertising Council .

Sants joined the board of the FSA in 2004 when he gave up his lucrative City career to embark on " PUBLIC SERVICE " as head of the regulator's markets – a poacher turned gamekeeper.

Al contrario que en el intervalo anterior, durante los años 2008-2011 la asociación semántica de la colocación de “public” y “spending” tiene aquí un carácter negativo. Casi todos los ejemplos observados tienen un sentimiento negativo y hacen referencia a “cuts” (“recortes”) en el gasto público durante la crisis y esto se refleja en la aparición de dicho término durante el periodo 2009-2011. En el único ejemplo en el que no se observa un sentimiento negativo, la oración propone una receta para aliviar los efectos de los recortes diseñada por actores sociales (i.e. segundo ejemplo del bloque). Aparte de esto, el adjetivo “public” tiene asociaciones semánticas con un sentimiento manifiestamente negativo y entre sus colocaciones se encuentran: “revenues falling” o “ballooning”. Algunos ejemplos revisados son los siguientes:

But with unemployment rising, revenues falling and the need for PUBLIC SPENDING cuts, local authorities need to think how they can do more for less.

Debt levels in many countries are likely to explode in the years ahead, too, as governments spend billions of dollars on recapitalising their financial sectors, and boosting PUBLIC SPENDING to kick-start the economy.

OECD chief economist Pier Carlo Padoan said all governments facing ballooning deficits should seek efficiency gains from PUBLIC SPENDING , particularly in education and health, and avoid "harmful" labour and capital taxes.

McDonnell believes that the Irish government will wait to see what impact next month's budget, in which €6bn cuts in PUBLIC SPENDING will be announced, has.

However, a divergence in growth patterns between north and south, rising input prices and the prospect of PUBLIC SPENDING cuts across the region means the euphoria could be shortlived.

El inicio de la crisis de la deuda se hace notar inmediatamente en el discurso económico, y en el año 2010 aparecen dos términos económicos relevantes junto a las colocaciones principales de “public”: “debt” (*logDice* 8,956) y “deficit” (*logDice* 8,480). La colocación “public debt” (“deuda pública”) se había convertido en una *palabra evento* al ser uno de los mayores problemas de los países de la Zona Euro, pues la mayoría de las economías estaban profundamente endeudadas y los datos macroeconómicos anunciaban niveles alarmantes de déficit.

En cuanto al sentimiento de las unidades “debt” y “deficit”, estos no son términos positivos o negativos *per se* en este dominio específico. Sin embargo, en ambos casos la mayoría de las oraciones (en global) y sus colocaciones (en particular) tienen una orientación semántica negativa. Para ilustrar esto de manera sencilla, la colocación de “public” y *debt* se asocian con unidades léxicas negativas como p. ej.: “instability”, “default”, “weak” o “fears”. Algunos de los ejemplos de consultados para “debt” son los siguientes:

E.U. governments have realized that other metrics, like PUBLIC DEBT , private sector indebtedness, shifts in unit labor costs and real estate prices, may also cause instability.

The UK government, however, still seems more worried about maintaining confidence in financial markets than reducing unemployment – despite the fact that no one would seriously expect the UK to default on its PUBLIC DEBT.

Italy is pressing for private as well as PUBLIC DEBT to be taken into account.

The new rules, entailing possible financial penalties for fiscal sinners and big losses for banks in the event of orderly insolvency, will favour the strong dictating to the weak on budget, DEBT and PUBLIC finances discipline, while also entrenching eurozone exposure to the pressures of the bond and financial markets.

The rest would be largely DEBT FUNDED, resembling a public private partnership.

IMF bailout loan as austerity bites IMF team flies in to Athens amid fears over economic recovery and reforms as EU predicts PUBLIC DEBT at 160% of GDP by 2013.

Por otro lado, “deficit” se asocia con términos como “staggering”, “rein” o “sprawling”. Algunos de los ejemplos relacionados con “deficit” en el subcorpus del año 2010 son los siguientes:

The source of the profligacy that has pushed the country to near bankruptcy and at 13.6% earned it Europe's biggest PUBLIC DEFICIT and debt, is its all-pervasive state – what the embattled prime minister George Papandreou calls "Greece's big sick man.

It's not a promising starting point for agreeing on earlier, mutual fiscal supervision to ensure that much-flouted E.U. rules on PUBLIC DEFICITS and debts are actually respected.

The Greek government, which had desperately tried to avoid a bailout, caved in after soaring borrowing costs rendered funding of the country's staggering PUBLIC DEFICIT and debt virtually impossible.

The protracted fight over carried interest underscores the difficulty Congress faces in trying to close tax loopholes for businesses, even at a time of sprawling budget DEFICITS and widespread PUBLIC dismay about Wall Street's influence in Washington.

There is widespread speculation that the IMF will pour further fuel on the fire by demanding that the government steps up the unpopular policies to rein in a PUBLIC sector DEFICIT of 13.6% and €300bn debt over the next three years.

No solo la acepción de “public” en referencia a la “cosa pública” (i.e. “lo gubernamental”) se había convertido en una palabra evento, pues “public” en el sentido relacionado con “lo popular” o lo referente a “la ciudadanía” también se convirtió en un evento con amplia presencia en este intervalo. Como se revisó más en profundidad en el Capítulo 2, desde los inicios de los recortes públicos la ciudadanía auto organizada (los colectivos de beneficiarios y los empleados públicos afectados por las políticas de austeridad) inició una oleada de protestas que tuvo repercusiones sociales y políticas de calado no solo en Europa. La semilla de la indignación queda recogida en el discurso y se refleja en las selecciones léxicas. Tal es así, que ya a partir de 2009 aparecen dos emociones en los distintos rankings de colocaciones principales de “public”: “outcry” (“indignación”) con un coeficiente *logDice* de 8,368 entre las colocaciones, mientras que “anger” (“ira”) aparece en la lista de 2010 y 2011 con un coeficiente *logDice* de 8,826 y 8,123, respectivamente. Ambas colocaciones tienen un sentimiento negativo en la mayoría de los dominios específicos y en la mayoría de las oraciones donde aparecen tienen un sentimiento global negativo. Algunos ejemplos de la coocurrencia de “public” y “outcry” muestran asociaciones con unidades léxicas como “job cuts”, “Wall Street greed” o “crisis”.

Buffini became the public face of the private equity industry - and a hate figure for the unions - when several large buyouts followed by job cuts led to a PUBLIC OUTCRY about "asset strippers".

The irony is that despite PUBLIC perceptions, the OUTCRY over Wall Street greed is happening just as the firms are getting stingy.

One hedge fund manager on the list, Paul Touradji, said he understood the PUBLIC OUTCRY against people who are paid regardless of whether they earned money for their clients.

Such strong language aside, the episode here has not generated the broad PUBLIC OUTCRY of, say, the bonuses paid by the American International Group.

The resurgence of bonus guarantees underscores just how difficult it is to control Wall Street pay, despite the PUBLIC OUTCRY over how taxpayer money is being spent.

The plight of temporary workers let go en masse in the fallout from the global financial crisis has prompted a PUBLIC OUTCRY .

Los ejemplos de coocurrencia entre “public” y “anger” que se observan relacionan la indignación del público con el buen trato de los gobiernos en el rescate de los banqueros (“bankers”) como entidad colectiva que personaliza al sector financiero en quiebra. Mientras, como se ve en el bloque anterior, la clase política imponía medidas de austeridad y condenaba al desempleo a la ciudadanía de a pie. Algunos de los ejemplos son los siguientes:

They are the firm that came out of the crisis in the best shape financially, along with JP Morgan, so they are a natural target of legitimate PUBLIC ANGER and frustration.

Politicians said today's polling showed the true extent of PUBLIC ANGER and revealed support for "wholesale reform" of the financial sector.

Of course, a cynic might suspect that hiding the use of a corporate jet is probably driven as much by a desire to deflect PUBLIC ANGER as by the need to protect some company executive.

Last year a combination of PUBLIC ANGER and the poor state of bank finances depressed City payouts.

The RBS boss has been a lightning rod for much of the PUBLIC ANGER over bankers' pay as he braces himself to pay bonuses to 22,000 investment bankers

The bank's chief executive, Lloyd Blankfein, has spent months trying to soothe PUBLIC ANGER over Goldman's conduct.

### Tercer intervalo de estudio

Durante la tercera etapa de análisis (2012-2014), la media de las puntuaciones de sentimiento de “public” se sitúa siempre por encima del umbral de los 70 puntos, correspondiendo con un valle en la frecuencia relativa de uso del término en el corpus.

Si bien las puntuaciones son positivas, no se aprecian colocaciones con sentimiento positivo, aunque sí un menor número de colocaciones negativas. Estas colocaciones son, de nuevo relacionadas con los recortes en los servicios públicos (el sustantivo “cut” en 2012), y un sustantivo relacionado con las emociones: “outcry”. Como era esperado, no se observan verbos entre las colocaciones del adjetivo de estudio. Más abajo se puede observar la lista de las colocaciones más significativas cada año (Tabla 45) junto al número de subcorpus anuales en los que aparece:

- Sustantivos genéricos: *account* (2), *affair* (2), *body*, *campaign*, *chair*, *comment*, *committee* (2), *debate*, *domain*, *eye*, *good*, *health* (3), *institution*, *interest* (2), *offering* (3), *opinion* (3), *outcry*, *pension*, *perception*, *policy* (3), *relation* (3), *safety*, *scrutiny*, *sector* (3), *servant*, *service* (2), *transport* (3).

- Sustantivos específicos: *cut*, *debt*, *finance* (2), *spending* (3), *trust*.
- Adjetivos genéricos: *initial* (3).
- Adjetivos específicos: *private* (3).

2012			2013			2014		
Palabra	Cooc	logDice	Palabra	Cooc	logDice	Palabra	Cooc	logDice
sector-n	139	11,002	relation-n	55	10,531	relation-n	29	10,196
relation-n	69	10,755	sector-n	77	10,068	offering-n	24	9,857
private-j	55	9,830	private-j	51	9,675	initial-j	23	9,730
service-n	56	9,458	health-n	55	9,674	health-n	32	9,256
offering-n	24	9,311	offering-n	24	9,344	trust-n	14	9,054
health-n	33	9,282	initial-j	21	9,137	sector-n	35	8,884
spending-n	31	9,251	committee-n	23	9,014	transport-n	11	8,874
policy-n	35	9,026	account-n	23	8,700	private-j	23	8,859
interest-n	34	8,971	finance-n	20	8,636	committee-n	16	8,729
initial-j	18	8,860	service-n	43	8,614	affair-n	9	8,672
<b>cut-n</b>	22	8,755	transport-n	13	8,557	opinion-n	8	8,536
affair-n	16	8,714	opinion-n	13	8,523	account-n	16	8,384
servant-n	14	8,618	policy-n	29	8,518	scrutiny-n	6	8,135
finance-n	20	8,599	interest-n	23	8,276	spending-n	11	8,105
transport-n	15	8,590	eye-n	11	8,247	policy-n	21	8,091
body-n	15	8,463	chair-n	10	8,179	domain-n	5	8,059
opinion-n	13	8,413	campaign-n	21	8,160	institution-n	7	7,987
pension-n	16	8,412	good-n	13	8,124	<b>outcry-n</b>	4	7,840
comment-n	14	8,410	perception-n	9	8,073	debate-n	6	7,811
debt-n	21	8,182	spending-n	13	8,065	safety-n	8	7,752

**Tabla 45** Principales colocaciones de “public” durante el tercer intervalo

Dado que no hay un patrón mayoritario de colocaciones con sentimiento positivo o negativo este intervalo que permita intuir los motivos por los que el sentimiento de “public” ha ascendido al umbral de la positividad, es pertinente aquí observar una muestra de oraciones que incluyan al adjetivo de estudio. La observación evidencia que, aunque las puntuaciones medias asignadas por *Lingmotif* a las oraciones son positivas, la gran mayoría de las oraciones que incluyen el adjetivo “public” tienen un sentimiento neutro, propio de oraciones factuales. Sin embargo, tras un análisis detallado de las colocaciones, se observa una gran cantidad de oraciones con un sentimiento negativo. Observado esto de manera cualitativa, en las oraciones se encuentran múltiples alusiones al histórico de la crisis y a la dureza de sus efectos que continuaron durante este intervalo. A continuación, se pueden revisar algunas de las oraciones aleatorias revisadas:

Rival estate agent Countrywide went public in March , valuing the owner of Bairstow Eves, Gascoigne-Pees and Hamptons International at £750m.

Stock Spirits slumps as Polish vodka maker warns on profits Company says results could be below expectations due to duty rise and competition Nick Fletcher Stock Spirits should have been toasting just over a year as a PUBLIC company, but a profit warning has left a nasty taste for investors.

Toyota's first female senior executive, Julie Hamp, is also a foreigner – an American who oversees PUBLIC relations at the automaker's North American division.

Finally, management scholarship should not be excluded from the current sustainability debates and PUBLIC scrutiny; otherwise, it will continue to stand in the way of mainstreaming the sustainability agenda.

"Land needs to be seen again as a PUBLIC good.

This decision can be seen as at odds with a true conservative philosophy, which usually embraces the provision of PUBLIC goods like a strong military and general national security.

EU considers full PUBLIC scrutiny of companies' tax arrangements Commissioner unveils first steps to tackle tax-dodging practices, but critics call proposal 'feeble' EU commissioner Pierre Moscovici holds a press conference in Brussels on tax and financial transparency.

Stephen Brozak, president of WBB Securities, a boutique health care investment bank, said it would be difficult for a PUBLIC company to continue with Mr. Shkreli at the helm.

Durante este intervalo, la colocación “public sector ”(“sector público”) (*logDice* 11,002 en 2012, 10,068 en 2013 y 8,884 en 2014) encabeza la lista de colocaciones de 2012 y está presente en dos años consecutivos. El término está reflejado de una manera negativa en relación con unos tiempos adversos económicamente. Como muestra, a continuación, se pueden observar algunos ejemplos de oraciones negativas que incluyen unidades léxicas de la misma valencia en su vecindario. Este es el caso de “opportunistically”, “loss” o “fail”.

In the PUBLIC SECTOR , employees may opportunistically set up companies in their own or in a family member's name in order to tender for contracts.

Unions added about 282,000 new members in the private sector, but that was partly offset by the loss of 118,000 members in the PUBLIC SECTOR , where budget pressures have meant layoffs and hiring freezes.

This report is a wake-up call to both the PUBLIC and private SECTORS to come up with constructive ways to realign the expectations of an increasingly anxious global community," said John Drzik, chief executive of management consultants the Oliver Wyman Group.

"Central government has made good commitments to opening up PUBLIC SECTOR procurement contracts to smaller firms but it needs buy in from the wider PUBLIC SECTOR too.

For far too long, executive positions in both the private and PUBLIC SECTORS have been the preserve of white males.

It's very often PUBLIC SECTOR projects that fail badly.

El uso de “spending ”(*logDice* 9,251 en 2012, 8,065 en 2013 y 8,105 en 2014) en los ejemplos observados devuelve principalmente ejemplos negativos. Este uso, obviamente, corresponde con los efectos de la crisis en Europa. Aparte del primer ejemplo, que va en correlación con lo observado en el análisis de “housing”, la prosodia



semántica de los elementos observados tiene una connotación claramente negativa y nivel especializado en el uso del término. A continuación, se observan algunos ejemplos:

Now, PUBLIC construction SPENDING is on the rise again.

Photograph: Alamy Doctors, patients and pharmaceutical companies have often existed in silos, but with PUBLIC sector SPENDING on healthcare continually on the rise and longer-term conditions such as cancer and alzheimer's disease affecting more and more people, the system could very well use a boost of collaborative innovation.

The OECD , which backed Osborne's austerity measures in the last parliament, said: "Real PUBLIC SPENDING is to fall significantly in fiscal years 2016 and 2017, to stabilise in 2018 and rebound in 2019.

It also assumes China's annual growth rate will stabilise at 6.5 – 7%, thereby enabling the risks posed by pockets of excessive leverage in the shadow-banking system to be gradually defused, even as the economy's growth engines continue to shift from exports and PUBLIC capital SPENDING toward domestic consumption and private investment.

A more radical approach to debt: do nothing It also warned Osborne against heaping most of the pain from PUBLIC SPENDING cuts on the poorest, saying that cuts should be fair, adding: "Although the composition of measures is yet to be defined, it is important that they mitigate distributional effects.

As Marchel Alexandrovich of City broker Jefferies pointed out in a recent note, that will potentially have a long-term impact on these countries' economic viability – narrowing the pool of taxpayers to fund future PUBLIC SPENDING far beyond the immediate struggle of emerging from recession.

Todos los ejemplos observados en los que aparece “cut ”tienen un sentimiento negativo. Todos ellos hacen referencia a los recortes en los servicios públicos y sus efectos en distintos sectores. Las enunciaciones se realizan frecuentemente en tiempos presentes lo que evidencia la actualidad de la colocación.

That means disproportionate CUTS for PUBLIC health and research and development.

Some 77% of the budget deficit is being recouped by PUBLIC expenditure CUTS and benefit cuts, and only 23% is being repaid by tax increases.

Businesses, says Cridland, still want CUTS in PUBLIC expenditure to take precedence.

The deep CUTS in PUBLIC spending have spawned a new generation of social entrepreneurs whose survival depends on working with market forces to build businesses capable of making a profit.

It says PUBLIC sector CUTS hurt the economy and exaggerated the recession, but no more than expected

Additionally, state budget CUTS are forcing PUBLIC institutions to reduce aid for graduate students, who in some disciplines have traditionally been paid to attend postgraduate programs.

It has been bailed out twice by international agencies and remains racked by the region's highest unemployment and frequent PUBLIC protests over government CUTS .

El uso de “public outcry” (*logDice* 7,840 en 2014) está relacionado con las quejas del público en relación con quejas relacionadas con las empresas públicas (p.e “Public Accounts Commitee”), concesiones de servicios públicos (p. ej., “Centrica”) o privadas (p. ej., “FCA”) principalmente en el contexto de la crisis en la gestión de “lo público”. Obviamente, se trata de un término con sentimiento negativo, y la totalidad de las



oraciones en las que aparece contienen esa polaridad. La totalidad de los ejemplos es la siguiente:

In November 2012, the Public Accounts Committee began an enquiry into tax avoidance, following PUBLIC OUTCRY that companies had avoided nearly £900m of tax.

The rejection of the more competitive commercial generated a large amount of news coverage, but there was no widespread PUBLIC OUTCRY demanding that CBS accept it, as SodaStream executives had hoped.

Wonga chairman to step down as FCA prepares to rein in payday lenders Errol Damelin will reportedly step down in the coming months, but will retain a valuable share in the company he founded Errol Damelin is said to have grown weary of having to constantly defend Wonga from political attacks and PUBLIC OUTCRY .

The equation is simple: companies that are run for the long term will thrive in the long term, companies run for the short term will be found out by events or legislation or PUBLIC OUTCRY .

Centrica caused a PUBLIC OUTCRY in April when it handed Laidlaw shares worth up to £2m.

The film transfixed audiences and led to a PUBLIC OUTCRY over the treatment of the whales, which perform in some of the park operator's most popular live shows in Orlando, San Antonio and San Diego.

Burrows avoided prosecution after making an out-of-court-settlement with Southeastern railways, but had to resign from his high-flying job at Blackrock after a PUBLIC OUTCRY .

#### Cuarto intervalo de análisis

La última etapa de análisis (año 2015) corresponde a una caída en la media de las puntuaciones de sentimiento de “public” por debajo del umbral de negatividad, hecho que coincide con una tendencia ascendente ligera en cuanto a la frecuencia en el corpus.

En el estudio del ranking de colocaciones (Tabla 46) se vuelven a apreciar nuevos dos unidades negativas. El sustantivo “cut” vuelve a aparecer por primera vez desde 2012, y emerge otro sustantivo que expresa emoción, en este caso “outrage”. Además de la lista de las colocaciones, se ofrece a continuación la clasificación de las unidades según la categoría gramatical y el número de subcorpus anuales en los que aparece:

- Sustantivos genéricos: *comment, committee, domain, health, infrastructure, offering, opinion, outrage, relation, scrutiny, sector, service.*
- Sustantivos específicos: *account, cut, debt, finance, policy.*
- Adjetivos genéricos: *initial, private.*

<i>Palabra</i>	<i>2015</i>	
	<i>Cooc.</i>	<i>logDice</i>
offering-n	32	9,821
initial-j	32	9,817
sector-n	71	9,630
finance-n	36	9,217
relation-n	18	9,142
spending-n	34	9,126
health-n	27	8,968
committee-n	25	8,861
account-n	26	8,753
opinion-n	12	8,606
private-j	23	8,464
service-n	37	8,270
policy-n	26	8,093
comment-n	9	7,898
<b>outrage-n</b>	6	7,887
<b>cut-n</b>	14	7,881
domain-n	6	7,872
debt-n	19	7,872
infrastructure-n	9	7,862
scrutiny-n	7	7,844

**Tabla 46** Principales colocaciones de “public” durante el cuarto intervalo.

Los datos cuantitativos ofrecidos por Lingmotif evidencian una negativización del contexto de uso de “public” durante esta época, por lo que es pertinente investigar los ejemplos y establecer las causas de la reaparición de unidades léxicas con sentimiento negativo como “outrage”, “cut” y “debt”. Se comenzará con una exploración aleatoria de las oraciones que contienen “public” en el subcorpus de este intervalo permite observar múltiples oraciones neutras, acompañadas de algunas oraciones de carácter negativo. Por ejemplo, se observan coocurrencias con unidades como “cuts”, “freezes”, “tax avoidance”, entre otras:

A new round of severe PUBLIC service cuts and pay freezes will keep the UK in the slow lane," said O'Grady.

House members, who face re-election every two years, tend to be more sensitive to PUBLIC opinion.

There is no Edward Clarkin living in Connecticut, according to PUBLIC records.

Unlike high-profile US activists such as Carl Icahn and Bill Ackman, ValueAct usually exerts influence behind the scenes instead of using high-pressure PUBLIC campaigns to get what it wants.

If our political plan needs to be endorsed by PUBLIC opinion, it will be.

PUBLIC scrutiny of tax avoidance had intensified as President Franklin D.

Far from being a good sign, that was actually the lowest quarterly growth in PUBLIC sector wages ever seen: The private sector experienced better wages growth if you include bonuses (0.8% in the quarter) but the growth including bonuses is pretty erratic – in the June quarter it was zero.

" Quite what will now replace it is now unclear: with their customarily elegant use of English, Sedgemoor council will only say that "further consultation will be had with the PUBLIC and key stakeholders, with such consultation having due regard to the need to deliver any new project at no cost, apart from feasibility funds".

The company should be severely punished for it's immoral, devious and deceitful approach to PUBLIC health risks as well as mis-selling to car drivers.

La colocación "public outrage "hace alusión principalmente a la reacción popular a las prácticas del capitalismo depredador que prima el beneficio ante el bien común. Estas prácticas fueron señaladas públicamente durante todo este periodo por los movimientos antiausteridad, que reclamaban a los gobiernos mayor vigilancia y arbitraje. En este caso, no se habla de un uso de "gestión de la cosa pública", sino en uso de "public" en referencia a "lo popular" en un contexto negativo. Un ejemplo paradigmático sería el del empresario estadounidense Martin Shkreli. El empresario es considerado uno de los mayores exponentes contemporáneos: dueño de fondos de cobertura y de la farmacéutica "Turing". En septiembre de 2015, obtuvo la licencia de fabricación del medicamento antiparasitario "Daraprim" y multiplicó 56 veces el precio de este (de 13,5\$ a 750\$ por pastilla), y fue calificado en la prensa como "el hombre más odiado en América". Shkreli fue acusado en un tribunal federal, y posteriormente fue condenado a siete años de prisión por conspiración y fraude de valores (BBC Mundo, 2018). Algunos ejemplos que evidencian la indignación popular son los siguientes:

The plans for the Chagas disease drug called to mind what made Mr. Shkreli the target of PUBLIC OUTRAGE to begin with.

When PUBLIC OUTRAGE at the failures of the financial system was at its height during the Occupy protests, the challenge posed was "what is your alternative?"

Its findings came as the Labour leader, Ed Miliband, insisted in his interview with the comedian Russell Brand that Labour had to respond to the PUBLIC OUTRAGE at multinationals using complex arrangements to pay less tax.

Just a year ago, the PUBLIC OUTRAGE was over Sovaldi, a new hepatitis C drug being sold by Gilead Sciences for \$1,000 a pill, or \$84,000 for a course of treatment.

There was PUBLIC OUTRAGE in November when The Guardian and other media organisations revealed that dozens of big companies were benefiting from sweetheart tax deals agreed with the Luxembourg authorities over many years, in some cases cutting their tax liabilities to less than 2%.

Rethinking prosperity Civil society urges politicians to transform the finance system Manifesto backed by 11 organisations calls on next government to create a finance system that serves society, the environment and the wider economy Public outrage at the failures of the finance system was at its height during the Occupy protests.

"Hidden and unregistered trials are compromising patient care, and, rightly, causing PUBLIC OUTRAGE .

Amid PUBLIC OUTRAGE and ridicule over drastically raising the price of a life-saving drug earlier this year, Shkreli, who is now facing criminal securities fraud charges and questions about his finances , tweeted that he gave "millions" to charity.

Aún en el año 2015, prosiguieron los recortes y la coocurrencia de public y cut volvió a adquirir relevancia por primera vez desde 2012 (*logDice* 7,881). En todos los casos se observan referencias a “public spending cuts” en tiempos presentes y con un sentimiento negativo. Aún en el año 2015, los recortes en el gasto público siguen a la orden del día. Los ejemplos evidencian la rabiosa actualidad de los eventos y es fácil relacionar el uso de esta colocación con las políticas de la UE sobre Grecia (quinto ejemplo) o la austeridad autoimpuesta por los gobiernos (primer ejemplo). Los efectos sobre la población y la indignación popular causada por estos se evidencian en algunos de ellos, como es el caso del cuarto ejemplo.

" A more radical approach to debt: do nothing It also warned Osborne against heaping most of the pain from PUBLIC spending CUTS on the poorest, saying that cuts should be fair, adding: "Although the composition of measures is yet to be defined, it is important that they mitigate distributional effects.

" He added that PUBLIC spending CUTS and rising interest rates are likely to bite hard outside of London.

This, plus another apparently doctrinal bias in favour of PUBLIC spending CUTS rather than tax increases, explains why we are now witnessing all too many examples in the public sector of false economies: cuts that hurt now, but can be guaranteed to lead to the need for emergency expenditure in future, expenditure which may well exceed the savings made in the short term.

And placing some four-fifths of the burden of the austerity programme on PUBLIC spending CUTS was very political indeed.

"I hope it will make a difference," she said as thousands of people chanting "no ifs, no buts, no PUBLIC sector CUTS ," streamed past.

After a weekend of intense talks with EU officials in Brussels, the final bid to keep Greece afloat will boil down to a tug-of-war deal over the country's staggering debt load, which Athens has insisted must be reduced as the price for further CUTS in PUBLIC spending.

Para concluir, el uso de la colocación de “public” y “debt” (*logDice* 7,782) vuelve a ser frecuente en este año, por primera vez desde el año 2012. Si bien se encuentran usos en tiempos presentes y pasados, en todos los casos su utilización es negativa (usos de “drag”, “debt... increasing”, “turbulence”, etc. en su vecindario). En todos los casos de mencionan los altos niveles de deuda pública durante dicho intervalo. Algunos de los ejemplos son los siguientes:

Third, while ECB policies keep borrowing costs lower, private and PUBLIC DEBT in the periphery countries, as a share of GDP, is high and still rising, because the denominator of the debt ratio – nominal GDP – is barely increasing.

"In economies with high PUBLIC DEBT , the pace of fiscal consolidation needs to strike an appropriate balance between debt reduction and imposing a drag on economic activity.

Further fiscal adjustment is needed to further reduce vulnerabilities from high PUBLIC DEBT, particularly given the increased risk of financial market turbulence.

PUBLIC sector net DEBT was increasing significantly during a period when it was being greatly reduced in most other industrialised countries.

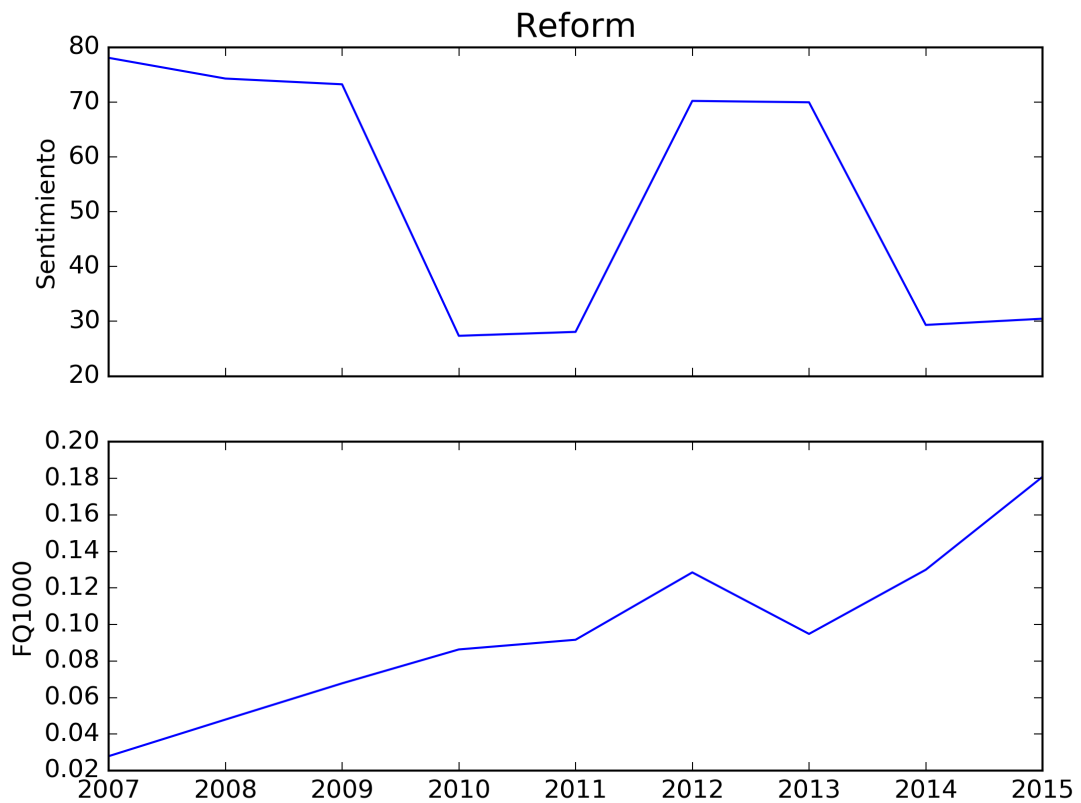
The most serious mistake was not to get across with sufficient emphasis that by far the biggest contribution to a rise in PUBLIC sector DEBT was caused by the banking crisis.

Third, not only is fixed income more illiquid, but now most of these instruments – which have grown enormously in number, owing to the mushrooming issuance of private and PUBLIC DEBTS before and after the financial crisis – are held in open-ended funds that allow investors to exit overnight.

## 7.8 Término 8: *Reform*

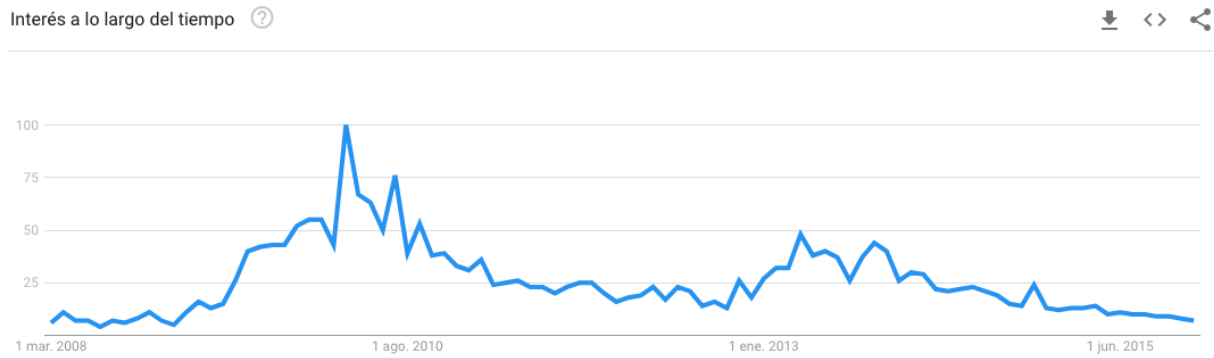
El término “reform” hace referencia principalmente a las políticas dirigidas a lograr mejoras en la eficiencia económica según el pensamiento económico del gobierno de turno. Este término no es exclusivo del dominio financiero y puede encontrarse también en otros dominios adyacentes como el legislativo o el político. Las reformas relativas al dominio económico pueden producirse mediante la modificación de las políticas de un sector o de toda la economía, p. ej., la política fiscal y la política de competencia. Generalmente, cuando se habla de “reform” suele hacerse desde dos perspectivas: (a) la desregulación y eliminación de la presencia del gobierno en los sectores, en lugar de nuevas o mayores regulaciones; o (b) la creación de políticas en las que la regulación estatal maximice el rendimiento económico.

El siguiente paso consiste en observar los factores que condicionaron el movimiento de la orientación semántica de los valores de sentimiento del contexto del término “reform”. A partir de la gráfica de medias de puntuación de sentimiento (Figura 24) y los datos estadísticos de los textos (Tabla 47), se pueden observar cuatro segmentos claramente diferenciados en la evolución del sentimiento del contexto del término: (1) una primera etapa en el trienio 2007-2009 que corresponde a los primeros años de la crisis con puntuaciones de sentimiento altas con valores medios de sentimiento se mantienen por encima de la franja de los 70 puntos (78,10, 74,31, 73,26, respectivamente); (2) un descenso repentino al umbral de sentimiento negativo durante el bienio 2010-2011 (27,36 en 2010 y 28,08 en 2011), (3) un nuevo ascenso a la franja positiva en el bienio 2012-2013 (puntuaciones 70,23 y 69,98, respectivamente ) y, finalmente, (4) un segundo descenso al umbral de negatividad en el bienio 2014-2015 (puntuaciones de 29,34 y 30,48).



**Figura 24** Gráfico longitudinal de las medias anuales de las puntuaciones de sentimiento para “reform” y la frecuencia relativa del término por cada 1000 tokens

En cuanto a las frecuencias relativas, en este caso no ocurre una correlación aparente entre frecuencia y sentimiento. En este caso se produce una tendencia progresiva desde una frecuencia muy baja de partida (FQ1000 0,03 en 2007) a una frecuencia relativa final de 0,18 en 2015. Es importante reseñar que la frecuencia relativa es mucho menor que en otros casos de estudio. Se concluye aquí que el término “reform” no suele aparecer en el corpus si no está de actualidad informativa por el transcurrir de los eventos políticos y económicos. Con el fin de realizar una comparación, una observación de la Figura 25 muestra también escasos valores, especialmente en las fechas correspondientes al primer intervalo. Esta gráfica corresponde a la frecuencia de búsqueda de noticias del término “reform” en *Google Trends* a nivel mundial para el intervalo 2008-2015.



**Figura 25** Número de búsqueda de noticias del término “reform” en *Google* durante el intervalo 2008-2015 (*Google Trends*)

Año	N	FQ1000	Sentimiento	pos_items	neg_items	neu_items	none_items
2007	48	0,03	78,10	0,88	0,73	12,67	11,50
2008	89	0,05	74,31	1,21	1,06	12,21	12,28
2009	135	0,07	73,26	1,23	1,13	13,76	12,37
2010	194	0,09	27,36	1,31	1,49	14,53	13,34
2011	248	0,09	28,08	1,38	1,60	14,55	12,98
2012	320	0,13	70,23	1,41	1,41	14,73	12,70
2013	341	0,09	69,98	1,36	1,34	13,48	11,73
2014	510	0,13	29,34	1,40	1,56	14,80	12,16
2015	785	0,18	30,48	1,54	1,55	15,64	13,05

**Tabla 47** Resumen estadístico del análisis de las oraciones que incluyen “reform”

### Primer intervalo de estudio

La primera fase de análisis corresponde a la época inmediatamente previa a la explosión de la crisis (año 2007) y a los dos años inmediatamente posteriores: 2008 y 2009. El análisis de los datos realizado con *Lingmotif* aquí ofrece una media de sentimiento positiva. Como se puede apreciar, existen escasas coocurrencias significativas durante el primer trienio y, en relación con su sentimiento, estas son neutras y relativas al ámbito de la economía o de la política. Teniendo esto en cuenta, junto con los bajos valores de frecuencia relativa se observa que durante este tiempo “reform” no fue un término evento de especial relevancia en el discurso.

El análisis realizado mediante *The Sketch Engine* ofreció una lista de colocaciones que permiten observar las relaciones con distintas palabras a lo largo de este intervalo, así como las distintas categorías gramaticales que coocurren con el término objetivo, es el caso de las siguientes:

- Sustantivos genéricos: *care* (2), *debate*, *government*, *health* (2), *immigration*, *plan* (2), *system*.
- Sustantivos específicos: *banking*, *bill*, *card*, *company*, *compensation*, *market*, *mortgage*, *policy*.

- Adjetivos genéricos: *comprehensive, last, public, radical*.
- Adjetivos específicos: *economic (3), financial (2), regulatory, structural*.
- Verbos genéricos: *back, be, call, have (2), make, push (2)*.

2007			2008			2009		
Palabra	Cooc	logDice	Palabra	Cooc	logDice	Palabra	Cooc	logDice
care-n	6	8,797	economic-j	9	8,269	health-n	22	9,926
health-n	7	8,622	policy-n	3	7,430	care-n	14	9,861
push-v	3	8,546	plan-n	4	6,921	regulatory-j	6	9,445
economic-j	5	8,415	call-v	3	6,376	structural-j	3	9,254
should-x	3	6,907	financial-j	3	5,630	immigration-n	3	9,218
government-n	3	6,500	last-j	4	5,315	radical-j	3	9,158
be-v	17	3,511	market-n	4	5,205	comprehensive-j	3	9,158
have-v	6	3,340	make-v	3	4,718	debate-n	3	8,666
			have-v	13	4,315	compensation-n	4	8,569
			company-n	3	3,956	system-n	9	8,386
						back-v	3	8,281
						bill-n	4	8,045
						push-v	3	7,790
						card-n	3	7,660
						banking-n	4	7,660
						financial-j	11	7,492
						mortgage-n	3	7,297
						plan-n	5	7,188
						economic-j	4	6,822
						public-j	3	6,785

**Tabla 48** Principales colocaciones de “reform” durante el primer intervalo

Dado el escaso número de colocaciones encontrado en el intervalo de estudio, es pertinente aquí hacer una revisión del sentimiento de una muestra de oraciones que incluyen el término “reform” en el corpus durante este intervalo. Tras observar las muestras, se destaca que una mayoría de las oraciones tiene una valencia positiva y coocurre con colocaciones asimismo positivas como p. ej., “legitimizing”, “approve”, “refinancing” o “long-awaited”, entre otras. Durante la crisis, la prensa y las instituciones demandaban políticas para paliar los efectos de la crisis. Los ejemplos son los siguientes:

The REFORMS will end up legitimizing the system, not weakening it.

Meanwhile, Beijing approved a health REFORM plan worth 850 billion yuan, or \$124 billion, last week that sets out to provide free basic health care to the country's 1.3 billion inhabitants by 2011.

"We absolutely believe this is one of the best ideas for Medicare REFORM that's around," said Constance Row, executive director of the American Academy of Home Care Physicians.

The FSA is also to preparing to consult on further raft of reforms intended to give customers total confidence that their savings are secure.





"We have an economy that has atrophied because of the lack of REFORM ," said Gerardo Esquivel, an economist at the Colegio de México.

The topics included regulatory REFORM , mortgage refinancing, the "stress tests" that federal regulators are now applying to large banks, and the proposed program to buy their troubled assets.

This mandate will strengthen the hand of the central government to unlock long-awaited REFORMS .

Then in April, with the economy in free fall, Paulson announced a plan for complete REFORM of the nation's financial regulators.

Insurers, whose stocks have been depressed by the difficult economy and visions of health reform , have price-to-earnings ratios in the single digits, he said, but the pharmacy benefit managers tend to command ratios in the teens.

Se procede, entonces, a estudiar las colocaciones más relevantes de este periodo y se observa que en estos años el término “care” (*logDice* 8,797 en 2007 y 9,861 en 2009) va seguido por la colocación “health” (*logDice* 8,622 en 2007 y 9,926 en 2009) en cuanto a significancia. Como es obvio, las colocaciones hacen referencias a las reformas en “health care” (“sanidad”, “sistema de salud”). Esta colocación no responde al evento histórico de la crisis global en sí (aunque en muchos ejemplos se observan relaciones), sino que hace referencia a las reformas en sanidad propuestas por el gabinete de gobierno de Obama en los EE. UU., el cual fue el germen del llamado Obamacare y que aún hoy en día sigue siendo un tema de debate polémico en ese país. Se evidencia una correspondencia entre las puntuaciones positivas medias generales en este intervalo y la valencia de las oraciones que incluyen esta colocación. Como se puede observar en los ejemplos, en una mayoría de los casos la prosodia de semántica de la colocación incluye términos positivos como por ejemplo “influence”, “strategic”, “growth” o “exciting”.

We're running this campaign like this was a presidential campaign, and our candidate is HEALTH CARE REFORM ," said Mr. Rivera, whose union is hugely resented by Republicans for doing so much to elect Mr. Obama.

Despite the program's limited reach the HR Policy group hopes it can help "shape and influence" the national debate on HEALTH CARE REFORM , said Ronald A.

In the next meeting, or maybe even in the same meeting, you might have to move to a very high, strategic question – for example, the president wants to REFORM the HEALTH CARE system and the economy is struggling.

The stimulus package provides a short-term buffer for the economy, and other policy measures such as HEALTH CARE and land REFORMS will be a long-term growth driver.

This is obviously an exciting time for HEALTH CARE REFORM .

With lawmakers still fighting over how to finance HEALTH CARE REFORM , Mr. Nadler said he hoped his bill might find an audience.

"You can find out a lot more about buying a washing machine than buying health care," said Senator Ron Wyden, an Oregon Democrat who is pushing HEALTH CARE REFORM.

Otra colocación que comienza a ser frecuente en el año 2009 es el adjetivo “radical” (*logDice* 9,158), el cual por sí no está marcado como positivo ni negativo. Una observación de la prosodia semántica en los ejemplos del corpus evidencia a través de sus colocaciones que se hace referencia a un sentimiento positivo. Durante la crisis, los voceros consideraban necesario materializar una serie de reformas fiscales, de un calado político u otro, que permitieran de nuevo el flujo de la economía. Algunas de sus colocaciones durante este periodo, observables aquí pueden ser “chance” o “out of poverty”.

If we accept these simple ideas, and take the conversation out of the gilded ghetto, we might have a chance of RADICAL REFORM .

Because we believe that work is the best way out of poverty we are continuing with RADICAL REFORM to the welfare system.

Despite warnings from industry Jeremiahs and calls for change from Washington, Wall Street is betting the ratings services will avoid the RADICAL REFORMS their most ardent critics are urging.

Para concluir este segmento de estudio, al contrario que en los ejemplos anteriores, el uso del término “mortgage” en el contexto de “reform” suele usarse con un sentimiento negativo. Como se pudo ver en el contexto relacionado con la crisis de la vivienda (“housing”) para este intervalo, las referencias al término estaban muy marcadas por la actualidad de la crisis tras la explosión de la burbuja inmobiliaria y en este contexto ocurre de manera similar. Se observan aquí tres ejemplos, de los cuales los dos últimos tienen una orientación semántica negativa:

This will save the Mitsubishi deal and buy time for successfully implementing the recapitalization and MORTGAGE REFORM programs.

However, any optimism about the speed of any housing market recovery is expected to be dampened by today's report from the Financial Services Authority, which will outline plans for REFORM of the MORTGAGE lending industry.

FSA report to outline REFORM plans for MORTGAGE lenders Reseach by property website Rightmove revealed that asking prices for homes in the capital had jumped 6.5% compared with last year.

### **Segundo intervalo de estudio**

La segunda de las cuatro fases de este análisis corresponde al bienio 2010-2011. Durante este periodo, las medias de puntuación de sentimiento fueron negativas con una tendencia moderadamente ascendente en su frecuencia relativa. A partir de esto último, creció el número de coocurrencias significativas y, además de los términos propios del dominio aparecieron unidades con sentimiento negativo (i.e. “draconian”) que ofrecen indicios de la relación del uso del término con un periodo de resultados económicos muy desfavorables. Otra muestra de esto es el ascenso de “radical” a la cabeza de la lista en el

año 2010 (*logDice* 9,841), por encima de otras colocaciones adjetivales aparentemente neutras como “structural”.

El análisis realizado mediante *The Sketch Engine* devolvió una serie de colocaciones para cada año que permiten evidenciar las relaciones léxicas del término de estudio con distintas unidades. Además, se pueden observar las distintas categorías gramaticales que coocurren con el término objetivo:

- Sustantivos genéricos: *agenda, bill, franchise, health, hearing, implementation, labour (2), legislation, measure (2), oversight, pace, pledge, rule, taxis*.
- Sustantivos específicos: *package, pension, tax, welfare*.
- Adjetivos genéricos: *deep, draconian, fundamental, radical (2), urgent*.
- Adjetivos específicos: *economic, financial, fiscal (2), regulatory (2), structural (2)*.
- Verbos genéricos: *adopt, implement, introduce, outline, propose, sign*.

2010			2011		
<i>Palabra</i>	<i>Cooc</i>	<i>logDice</i>	<i>Palabra</i>	<i>Cooc</i>	<i>logDice</i>
structural-j	10	10,421	radical-j	7	9,385
radical-j	7	9,841	structural-j	6	9,196
hearing-n	6	8,944	implement-v	6	8,905
health-n	10	8,858	labour-n	6	8,830
<b>draconian-j</b>	3	8,784	implementation-n	4	8,766
regulatory-j	5	8,710	urgent-j	4	8,752
labour-n	4	8,669	measure-n	10	8,707
bill-n	8	8,664	regulatory-j	6	8,699
oversight-n	3	8,654	pledge-n	4	8,557
welfare-n	3	8,590	propose-v	6	8,502
outline-v	3	8,529	agenda-n	4	8,449
<b>fundamental-j</b>	3	8,524	pace-n	4	8,359
rule-n	8	8,472	package-n	5	8,122
fiscal-j	4	8,338	deep-j	4	8,120
franchise-n	3	8,304	tax-n	13	7,947
adopt-v	3	8,281	pension-n	5	7,925
economic-j	10	8,110	fiscal-j	4	7,865
legislation-n	3	8,109	introduce-v	5	7,780
measure-n	5	8,090	taxis-n	4	7,773
sign-v	4	8,042	financial-j	15	7,691

**Tabla 49** Principales colocaciones de “reform” durante el segundo intervalo de estudio

Dado que existe una mayoría de elementos léxicos neutros en la clasificación, siendo el patrón predominante en el año 2011, también es considerado pertinente realizar una observación exploratoria de una muestra aleatoria del sentimiento de las oraciones en

este segmento del corpus. Una mayoría de las oraciones contienen una orientación semántica negativa y entre sus colocaciones se encuentran unidades léxicas negativas como “limit”, “exacerbate”, “aftermath” o “crisis”, entre otras. Más abajo se pueden revisar los ejemplos:

Banks are to meet campaign groups in Zurich next month to discuss possible REFORM that would limit the funding of energy projects that exacerbate climate change and make funding decisions more transparent.

Further REFORM of the principles would be a matter for individual bank groups which could destroy the existing consensus, the source added.

The landmark paper will increase pressure on US politicians to regulate the mortgage industry, which Washington insiders say has so far been immune from meaningful financial REFORM in the aftermath of the bank crisis.

In Washington, the House of Representatives oversight and government REFORM committee has scheduled a hearing next week entitled Toyota Gas Pedals: Is the Public at Risk? amid growing concern.

Athens has pledged to roll back its budget deficit from 12.7% to 2.8% by 2012 through REFORMS including spending cuts and increasing tax revenue.

This week the EU is expected to pile on the pressure further by setting out further budget cuts and financial REFORMS that it will ask Athens to meet by May this year.

Como se vio en el intervalo anterior, el uso de “radical” en el contexto de “reform” es considerado generalmente deseable en el discurso de la prensa durante la crisis. Durante este intervalo, sin embargo, se habla en ella en términos de necesidad y campaña por una reforma radical del sistema bancario a través de la acción política (“need”, “crusaders”, “push ahead”, “calls”...). Como es bien sabido, durante esta época, más decididamente en los Estados Unidos, se intentaron depurar responsabilidades y evitar que situaciones como las evidenciadas a partir de 2008 volviesen a ocurrir. Otro factor es la reforma del sistema de pensiones en el Reino Unido. Sin embargo, la prosodia semántica aquí refleja un sentimiento negativo dada la situación política y económica. Como se puede observar a continuación en los ejemplos del corpus, aquí se observan unidades léxicas con sentimiento negativo como “nervous”, “insolvency”, “collapse” o “risky”:

Both King and Turner have become unlikely crusaders for RADICAL REFORM, as nervous politicians adopt a more cautious approach.

The National Association of Pension Funds, which represents schemes worth more than £800bn, argued today that RADICAL REFORM of pension accounting rules was needed to prevent companies from closing schemes, often to avoid being pushed into insolvency.

Instead, it has afforded a golden opportunity for the 57-year-old to push ahead with a RADICAL programme of REFORM that will, he hopes, change the way Greece works and the way its people think.

The protests, which despite a lower than usual turnout attributed by unionists to the summer heat, came as parliament prepared to debate legislation outlining RADICAL REFORMS to the pension system.

While the reforms were regarded as hitting Barclays and bailed-out Royal Bank of Scotland hardest, bankers conceded that the recommendations by Sir John Vickers and his Independent Commission on Banking (ICB) could have been worse given some of the calls for RADICAL REFORM made in last year's election campaign.

Those hoping for RADICAL and immediate REFORM, almost four years after the run on Northern Rock and nearly three years to the day that Lehman Brothers collapsed, were disappointed after Vickers conceded that rapid implementation of his reforms was not practical.

The accounting regulation system needs RADICAL REFORM so that banks are not encouraged to invest in risky assets to make themselves seem more profitable than they really are.

En la misma línea de necesidad de reformas que en este intervalo expresa “radical” está el adjetivo “urgent” (*logDice* 8,752 en 2011), el cual no tiene un sentimiento inherente. Una observación exploratoria de los ejemplos encontrados en el corpus permite evidenciar colocaciones con sentimiento negativo en su vecindario, como es el caso de “worrying”, “toxic” o “risk”.

"The British Retail Consortium agreed that the situation was worrying and is calling for a "retail-friendly" budget next month, placing the URGENT REFORM of the business rate system at the top of its list of priorities.

Amid concerns about the health of Europe's banking system, Oakeshott told a debate on the Vickers reforms in the Lords that "this reminds us how much toxic banking risk remains in the system, and how URGENT radical REFORM is".

But doing so requires more than just better national policymaking in Europe and America; it is also time for URGENT and deep REFORM of the multilateral system and its main institutions.

Por último, “draconian” (*logDice* 8,784 en 2010) emerge como adjetivo calificativo de “reform” en este intervalo. Si bien está dentro del ámbito especializado, esta unidad dista de ser terminología aséptica e incluso su definición por el diccionario *Collins* Cobuild (2019) hace uso de léxico negativo “Draconian laws or measures are extremely harsh and severe.” En este caso, todos los ejemplos hacían referencia a las medidas impuestas desde la Unión Europea como condición del rescate de Grecia y todos tienen valencias negativas. Los ejemplos pueden consultarse abajo:

Greeks were told today that they faced seeing their country go bankrupt unless they accepted more austerity measures, as the prime minister ignored public outrage and pushed ahead with a RADICAL new package of budgetary REFORMS .

Public sector workers to pay a heavy price for Greek bailout Prime minister, George Papandreou insisted that his government has no choice but to enforce the DRACONIAN economic REFORMS Greece , eurozone's weakest link has exposed the single currency bloc's vulnerabilities.

But under the deal the Greek government must also freeze wages, slash pensions and do away with a plethora of bonuses and allowances – DRACONIAN REFORMS not seen since the second world war.

### Tercer intervalo de estudio

La tercera etapa de este análisis corresponde al bienio 2012-2013. Durante este periodo, las medias de puntuación de sentimiento fueron positivas y supusieron una bajada temporal de 3 puntos en su frecuencia relativa. No obstante, durante este intervalo sí se puede hablar de que “reform” se convirtió en una palabra evento. Esto se justifica en que el número de coocurrencias significativas fuera de su dominio específico aumentó y es especialmente llamativo el hecho de que solo “banking” sea el único sustantivo perteneciente al dominio de la economía y las finanzas.

La tendencia en este intervalo temporal coincide con las puntuaciones medias positivas de otros términos estudiados como “credit”, “housing” o “investor”. No obstante, estas puntuaciones no parecen reflejar con una época de gran turbulencia política provocadas, de hecho, por las reformas planteadas por distintos actores políticos. De hecho, estas reformas se llevaron a cabo confrontando a grandes sectores de la población en una época plagada de huelgas (p. ej., Grecia o España) y protestas (p. ej., las manifestaciones del 15-O de 2011 en las principales capitales de Europa). Por ello, es interesante observar las colocaciones y su uso en el corpus y determinar si esto queda reflejado en el corpus.

En este intervalo sí se observan varias unidades léxicas que tienen un sentimiento positivo (p. ej., “necessary”, “ambitious”) y estas superan numéricamente a las unidades que portan una valencia negativa (p. ej., “austerity”). Sin embargo, la observación cualitativa a nivel detalle de los ejemplos permite observar que, aunque muchas de estas colocaciones son positivas, estas se usan en contextos negativos y reflejan justamente pinceladas de esta situación histórica.

Más abajo, se puede observar la lista de colocaciones que coocurren con “reform” clasificada tanto por categoría gramatical como por año (Tabla 50):

- Sustantivos genéricos: *agenda, austerity, bill, blueprint, campaigner, care, commission, education, health, immigration, labour (2), law, legislation, pension, planning, programme, proposal, root, tax.*
- Sustantivos específicos: *banking (2).*
- Adjetivos genéricos: *ambitious, bold, economic (2), fundamental, necessary, radical.*
- Adjetivos específicos: *fiscal, regulatory, structural (2).*
- Verbos genéricos: *delay, enact, implement (2), introduce, pass, push.*

2012			2013		
Palabra	Cooc	logDice	Palabra	Cooc	logDice
structural-j	20	10,618	structural-j	21	10,562
labour-n	17	9,965	implement-v	14	9,554
implement-v	13	9,728	banking-n	24	9,480
banking-n	23	9,426	bill-n	15	9,003
<b>fundamental-j</b>	7	8,987	radical-j	5	8,567
fiscal-j	9	8,848	tax-n	24	8,489
health-n	13	8,793	agenda-n	6	8,447
legislation-n	6	8,708	campaigner-n	5	8,409
care-n	8	8,693	immigration-n	4	8,268
proposal-n	7	8,456	<b>delay-v</b>	4	7,944
<b>enact-v</b>	4	8,453	law-n	10	7,868
<b>austerity-n</b>	6	8,387	<b>bold-j</b>	3	7,867
root-n	4	8,317	programme-n	7	7,860
introduce-v	7	8,295	push-v	6	7,780
regulatory-j	5	8,205	labour-n	4	7,747
education-n	5	8,119	<b>necessary-j</b>	4	7,726
blueprint-n	3	8,101	pension-n	5	7,701
commission-n	5	8,100	<b>ambitious-j</b>	3	7,681
economic-j	14	8,063	economic-j	13	7,625
pass-v	5	8,025	planning-n	3	7,535

**Tabla 50** Principales colocaciones de “reform” durante el tercer intervalo de estudio

Para comenzar, se ofrece aquí también una visión general del uso de “reform” en el corpus para este intervalo. En él se encuentran una mayoría de oraciones relacionadas con las reformas exigidas por la UE a los países miembros (p. ej., la reforma laboral española o reforma bancaria en el Reino Unido). A nivel global, aparecen oraciones con sentimiento positivo (con palabras en su vecindario como “positive” y “wellbeing”), junto con otras de carácter neutro y un solo ejemplo de oración con sentimiento negativo, situada en el último lugar de este bloque.

The so-called troika of the I.M.F., the European Commission and the European Central Bank issued positive statements but urged Mr. Rajoy's government to maintain momentum on REFORM .

Key political and economic figures, as well as water specialists, now recognise that a water REFORM plan is important to both a country's economic growth and its social and environmental wellbeing.

In 2009 and 2010, the RBS subsidiaries also lobbied on a range of banking REFORMS including the Dodds-Frank Wall Street Act, the largest of the US banking REFORM acts.

Unlike US lobbying regulations , David Cameron's REFORMS will not list in-house lobbyists, nor will they show amounts spent or subjects involved.

In the 12th Five-Year Plan, presented last year, the government committed itself to "deepening reform of China's cultural industries.

In November, Newman used a speech to say Australia's wage rates were high by international standards and workplace REFORM was needed.



"The I.M.F. has learned a lot how to design programs and structural REFORM measures and how to embed them in the local political system," Mr. Berglof said.

RBS has spent most of its efforts this year lobbying on attempts to REFORM "swipe" card fees.

Substantive and well supported REFORM programs have been developed and are underway in each location.

Other REFORMS being considered by ministers include ending the "cosy cartel" that allows top executives to set each other's pay, by bringing outsiders on to remuneration committees, and introducing rules to force executives to take more of their pay in shares that cannot be cashed in for at least five years.

In a report in September – its fourth on the issue – the select committee found that the pubcos had "wasted a final opportunity " to REFORM their practices, offered to them by the committee in a previous report in 2010.

El verbo “enact” (“promulgar”) es una colocación propia de “reform” (*logDice* 8,453), pues hace referencia al proceso presidencial o parlamentario que las habilita legalmente para ser aplicadas. La aprobación de una reforma es considerada, generalmente, un hecho deseable en contraposición con su obstrucción o bloqueo. En estos casos se habla de dos reformas principales, la primera la del sistema de sanidad estadounidense y la segunda relativa a las exigencias de la UE al estado helénico. Como se puede observar en los ejemplos, existen coocurrencias con valencia positiva (“dares”, “growth”) y, además, se evidencia que su negación (“not enacting”) coocurre con unidades léxicas negativas como “risk”. A continuación, pueden revisarse los ejemplos:

Final legislation is due to allow REFORMS to be ENACTED .

It has managed to transform a programme of endless austerity into a programme that dares [to ENACT ] REFORMS and will lead to growth," added a tired looking but visibly relieved Samaras.

Unless the pace or nature of medical innovation changes, this trend is likely to continue, regardless of structural REFORMS we ENACT for Medicare .

These steps clarified the state's financial position, but the legislature has yet to ENACT significant pension REFORM.

“As this represents a new run mechanism that does not exist under the status quo, the fees-and-gates alternative may actually increase run risk relative to not ENACTING further REFORM ,” Mr. Rosengren wrote.

El uso del adjetivo “necessary” (*logDice* 7,726) hace referencia a la idoneidad de promover las reformas. En este caso, las distintas cabeceras aquí hablan de unas reformas necesarias que no terminan de llevarse acabo (“being abandoned”, “delay in implementing” y “fail to implement”), que junto a una prosodia semántica con sentimiento negativo (“backlash”, “weakened, pockets of poverty”) promueven un sentimiento opuesto al reflejado por la unidad léxica. Más abajo pueden encontrarse los ejemplos revisados:

As our Paris correspondent Kim Willsher writes: Moody's said it was not confident Hollande's government could – or would – introduce the NECESSARY structural REFORMS and spending cuts to improved its rating



in the medium term and expressed concern over France's exposure to risks from other ailing eurozone countries.

Jin Liqun, chair of the supervisory board of the \$480bn (£300bn) China Investment Corporation (CIC), warned that undue harshness risked a backlash which could end with NECESSARY economic REFORMS being abandoned.

Unlike China, which has shown clear signs of stabilisation since mid-2012, there is no clear evidence of recovery in India yet, as delay in implementing NECESSARY REFORMS, among other factors, has weakened the economy's competitiveness.

Countries that fail to implement the NECESSARY REFORMS would become permanent pockets of poverty and dependency, much like Italy's Mezzogiorno region today.

Two NECESSARY conditions for these structural REFORMS are to improve our competitiveness and to build a new, investment friendly, business environment.

La tercera unidad de estudio seleccionada es “radical”, que en este intervalo se utiliza como una unidad léxica con sentimiento negativo, al contrario que en los intervalos anteriores. Esto se justifica en la aplicación de medidas de reforma exigidas por la UE y el BCE a sus países miembros, los cuales fueron consideradas impopulares por muchos sectores políticos y, como se revisó en el Capítulo 2, generaron una ola de protestas por amplios sectores de la población. Estos imponían medidas de austeridad, privatizaciones y recortes. Como se puede observar, en este caso las “radical reforms” se relacionan con unidades negativas como “austerity”, “refused” o “dark pool”.

In the letter, the officials policing Greece's compliance with the austerity package imposed in return for the bailout insist on RADICAL labour market REFORMS, from minimum wages to overtime limits to flexible working hours, that are likely to worsen the standoff between the government and organised labour in Greece.

Ed Balls, the shadow chancellor, said: "It confirms George Osborne is continuing to duck the RADICAL banking REFORM we need and which the cross-party commission has demanded.

Among the countries driving the initiative have been Germany and the UK, with the British chancellor, George Osborne, and his finance minister counterpart, Wolfgang Schäuble, issuing a joint statement on the fringes of a G20 meeting in Berlin in November calling for RADICAL REFORM.

The German argument is that it needs RADICAL REFORM in order to survive.

When governments refused to back RADICAL REFORM, the JP Morgans and Citibanks took their trading to these so-called dark pools where no regulator could see what was going on.

La última colocación estudiada en este intervalo es “austerity”, la cual hace referencia a las políticas de recorte impuestas desde la *troika* del FMI, BCE y la Comisión Europea a los distintos estados miembros de la Unión Europea. Todos los ejemplos hacen referencia a estos eventos históricos, en especial las condiciones exigidas del segundo rescate de Grecia que imponían recortes (“cuts”) en los presupuestos, paquetes de privatizaciones de empresas públicas e infraestructura y control directo sobre sus finanzas. La austeridad fue una palabra evento por sí y siempre relacionada con un

sentimiento negativo en las enunciaciones. Tal es así, que todas las oraciones tienen un sentimiento negativo y existen múltiples unidades léxicas relacionadas con esa polaridad. Entre estas últimas se encuentran adjetivos como “unpopular” o “recessionary” y sintagmas nominales como “spending cuts”.

In the event that the new general elections scheduled for 17 June fail to produce a government with a mandate to continue with the EU-IMF [International Monetary Fund] programme of fiscal AUSTERITY and structural REFORM, an exit of Greece from EMU would be probable," Fitch said.

Postponing the exit after the June election, with a new government committed to a variant of the same failed policies (recessionary AUSTERITY and structural REFORMS ), will not restore growth and competitiveness.

The IMF has U-turned in recent months and urged government to allow their AUSTERITY REFORMS to be planned over a longer period to lessen the impact on growth.

And Venizelos also defended the €13.5bn package of tax rises, spending cuts , labour REFORMS and other AUSTERITY - saying Athens has 'no choice' but to comply.

"A very grey and very dark period for Greece closed definitively," said the prime minister, whose fragile coalition had put its future on the line by endorsing a fresh round of unpopular AUSTERITY MEASURES and structural reforms in return for further assistance from the EU and IMF.

#### Cuarto intervalo de estudio

Durante la última etapa de este análisis (2014-2015), el contexto de “reform” se caracteriza por tener más coocurrencias con unidades del dominio específico que en el intervalo anterior y por un descenso repentino de las puntuaciones medias de sentimiento al umbral de la negatividad. Muchas de las colocaciones hacen referencia también aquí a las condiciones del tercer rescate griego y a la situación previa al *Quantitative Easing* promulgado por el BCE. En el año 2014 predominan las colocaciones con sentimiento neutro, con ausencia de colocaciones claramente negativas y tan solo aparece una sola colocación considerada positiva dentro del dominio específico (“consolidation”). Mientras tanto, en el año 2015 aparecen las colocaciones “cut” y “austerity” con un sentimiento claramente negativo.

Por otro lado, la puntuación de colocación de palabras neutras como p. ej., “structural”o “implement” permanecen generalmente estables durante este periodo. En cuanto a las colocaciones, se encuentran las siguientes, junto al número de subcorpus anuales en los que coocurre con “reform”:

- Sustantivos genéricos: *agenda, austerity, consolidation, governance, government, labour (2), list, pension (2), programme (2), set.*
- Sustantivos específicos: *banking (2), creditor, cut, tax (2), VAT.*
- Adjetivos genéricos: *fundamental, radical (2), regulatory (2), (2), sweeping.*

- Adjetivos específicos: *economic* (2), *fiscal* (2), *structural*.
- Verbos genéricos: *demand*, *implement* (2), *need*, *propose* (2), *push*, *submit*.

<i>Palabra</i>	<i>2014</i>		<i>Palabra</i>	<i>2015</i>	
	<i>Cooc</i>	<i>logDice</i>		<i>Cooc</i>	<i>logDice</i>
structural-j	56	11,323	structural-j	78	11,356
implement-v	15	9,308	implement-v	41	10,246
governance-n	12	9,026	pension-n	33	9,502
radical-j	10	8,988	economic-j	48	8,859
fiscal-j	10	8,618	fiscal-j	17	8,841
programme-n	14	8,597	list-n	18	8,763
fundamental-j	8	8,588	labour-n	17	8,754
regulatory-j	10	8,573	propose-v	18	8,748
tax-n	27	8,539	demand-v	15	8,712
<b>consolidation-n</b>	7	8,538	banking-n	23	8,671
propose-v	9	8,346	programme-n	19	8,611
sweeping-j	5	8,139	tax-n	42	8,507
set-n	7	8,129	<b>austerity-n</b>	13	8,493
agenda-n	6	8,067	creditor-n	15	8,450
government-n	32	8,009	VAT-n	10	8,440
banking-n	10	8,004	submit-v	9	8,263
labour-n	7	7,944	<b>cut-n</b>	18	8,232
push-v	9	7,916	regulatory-j	12	8,226
pension-n	7	7,853	need-v	37	8,185
economic-j	17	7,788	radical-j	8	8,123

**Tabla 51** Principales colocaciones de “reform” durante el cuarto intervalo

Para comenzar, se ofrece aquí también una visión general del uso de “reform” en el corpus para este intervalo. Una revisión general de las oraciones extraídas aleatoriamente del subcorpus permite observar que la mayoría de las oraciones tienen un sentimiento neutro o negativo. Algunas de las unidades léxicas negativas que se observan a simple vista en las oraciones son “snag”, “spending cuts”, “hard-hitting” o, directamente, “negative”.

Saker Nusseibeh, chief executive officer at Hermes Asset Management, who sits on a CFA advisory council, is also calling for REFORM, saying the body should be under a lot of external pressure to rethink what they teach their guys.

That way, it could also make companies responsible for labour abuses in their supply chains in the same way they have through similar REFORMS to corporate bribery.

The IMF has released a hard-hitting report on western governments' banking REFORMS.

In areas such as the implementation of resolution frameworks or structural REFORMS, countries have adopted policies without much co-ordination.

In Europe, politicians are aware of the need for major structural REFORMS, including those required to solidify regional integration; but companies lack adequate clarity about the components of such reforms.

Payment of the rescue loans depend on the country meeting criteria in spending cuts, tax increases and REFORMS, and Greece's progress in meeting the targets is reviewed regularly by the debt inspectors, collectively known as the 'troika'.

Talks had snagged on several issues, including public sector firings and market REFORMS.

For companies that are able to take advantage of the US lobbying system, campaign finance REFORM is bound to be a net negative.

El término “consolidation” (*logDice* 8,538) hace referencia a “fiscal consolidation”, el cual se refiere a las políticas emprendidas por los gobiernos para reducir sus déficits y la acumulación de reservas de deuda. Aunque en la lista que recopila las colocaciones del año 2014 este término está resaltado en verde como positivo, se puede observar que las oraciones que hacen referencia al mismo suelen tener un sentimiento negativo. El uso de “fiscal consolidation and structural reform” es usado casi metonímicamente para referenciar al plan de *Expansión Cualitativa* promulgado por la presidencia de Draghi. Su uso aquí se centra en las desventajas de este plan: principalmente el riesgo de deflación. En el vecindario de estos términos se encuentran adjetivos como “painful”, “deflationary” o “spectre”. A continuación, se encuentran las oraciones que incluyen la colocación:

If QE and looser short-term fiscal policies boost demand, growth, and employment, governments may be more likely to implement politically painful structural REFORMS and long-term fiscal CONSOLIDATION .

According to this moral-hazard argument, the authorities will push ahead with painful labor- and product-market REFORMS and pursue the fiscal CONSOLIDATION needed to gain the confidence of investors only if they remain under the gun.

And structural REFORM and fiscal CONSOLIDATION are both deflationary in the short run.

Indeed, Europe does stand to reap significant benefits from fiscal consolidation and structural REFORM – but in the long run, not in the short run, when the problem is deflation.

Ireland's request should include a commitment to undertake "decisive actions in the areas of fiscal CONSOLIDATION, structural REFORMS and financial sector restructuring" in agreement with international partners, the letter added.

It follows that the best thing the ECB could do to encourage fiscal consolidation and structural REFORM is to banish the spectre of deflation.

Those who object that such audacity would discourage fiscal CONSOLIDATION and structural REFORM should consider what deflation would do.

El uso de “sweeping” es muy similar al uso de “radical” en este contexto y hace referencia a medidas drásticas necesarias en la política económica. Sin embargo, una observación de las oraciones en las que esta colocación aparece en el corpus permite observar que una mayoría contiene, en global, un sentimiento negativo. Se observan

unidades como “controversial” o “spending cuts”. Más abajo se pueden observar las oraciones extraídas del corpus que incluyen esta colocación:

"It will be very hard for him to please both sides of this equation," said Mr. Brzeski, alluding to the restive Greek electorate – who voted in large numbers for Mr. Tsipras to ditch austerity – and to the country's creditors, who are demanding that Greece enact SWEEPING REFORMS before being given more bailout money.

Greece's creditors want the Tsipras government to make deeper spending cuts, as well as faster and more SWEEPING REFORMS to the Greek pension system, according to a leaked version of the creditors' counter-proposals that is very similar to an earlier version rejected by the Greeks on Wednesday.

Such SWEEPING energy REFORMS are not likely to be enacted by Congress soon.

However, the search for a new boss is on hold until group members have voted on controversial proposals for a SWEEPING REFORM of governance throughout the organisation.

Renzi's promise to institute SWEEPING REFORMS to business and labour markets appeared to be more than hot air following the appointment of Pier Carlo Padoa-Schioppa as finance minister.

These steps have yet to have an effect on the ground, and economists and executives are looking for more SWEEPING REFORMS to reinvigorate capital investment, increase business confidence and encourage industry.

El sustantivo “austerity” vuelve a aparecer en el listado de colocaciones del año 2015 con un coeficiente muy similar al de 2012 (*logDice* 8,493). En los ejemplos se puede observar también que existen referencias claras a la situación política y económica en la Zona Euro y que el sentimiento en el contexto de la colocación es unánimemente negativo. En cuanto a la prosodia semántica de los ejemplos extraídos para su análisis cualitativo, el vecindario de la colocación incluye unidades léxicas como “weakening”, “backlash”, “unpopular” o “fatigue”.

Some more hawkish ECB officials worry that QE will lead to moral hazard by weakening governments' commitment to AUSTERITY and structural REFORMS.

Indeed, the social and political backlash against AUSTERITY and REFORM is stronger when there is no income or job growth.

Slow job creation and income growth may continue to fuel the populist backlash against AUSTERITY and REFORM.

And AUSTERITY and REFORM fatigue in the eurozone periphery has been matched by bailout fatigue in the core, boosting support for a range of anti-EU parties in Germany, the Netherlands and Finland.

The move came as the Greek government reacted angrily to what was seen as an ultimatum from its creditors – including the IMF – that demanded further AUSTERITY and unpopular REFORMS to VAT, pensions and wage bargaining as the price for €7.2bn in fresh financial help.

Finalmente, el término “cut” (*logDice* 8,232) aparece en 2015 y refleja la frecuencia con la que las reformas exigidas por la *troika financiera* a los países periféricos de Europa (principalmente Grecia) implicaban recortes en los salarios, gasto público o las empresas públicas. Aunque es muy llamativo que este término no apareciese entre las

colocaciones anteriormente, el término refleja fielmente el espíritu detrás del cambio del sentimiento tras el contexto de “reform”: una oleada de cambios en la actualidad relacionados con la crisis que modifican la prosodia semántica de un término en concreto. Alrededor de “cuts” orbitan términos con un sentimiento negativo y a través del examen de una muestra de oraciones extraída del corpus se observan términos como “hard-hitting”, “bailout”, “painful” o “humiliation”.

Facing a rebellion over the terms of a fresh €86bn (£61bn) rescue package, Tsipras has considered sidelining the increasingly belligerent leftwing faction in Syriza in favour of a broader coalition to push through spending CUTS and painful REFORMS .

The troika of lenders insisted on a broad set of REFORMS and public spending CUTS as the price for the loans under a programme that ran out in February and was extended until the end of June.

The intractable problem is that many governments do not trust the Greek government to implement a €12bn (£8.6bn) programme of spending CUTS and REFORMS that will be delivered as part of a bailout.

Tsipras must now take the measures , which include VAT REFORM, spending CUTS, a pensions overhaul and €50bn in privatisation, to a hostile Greek parliament . "This is indeed the politics of humiliation," said Varoufakis.

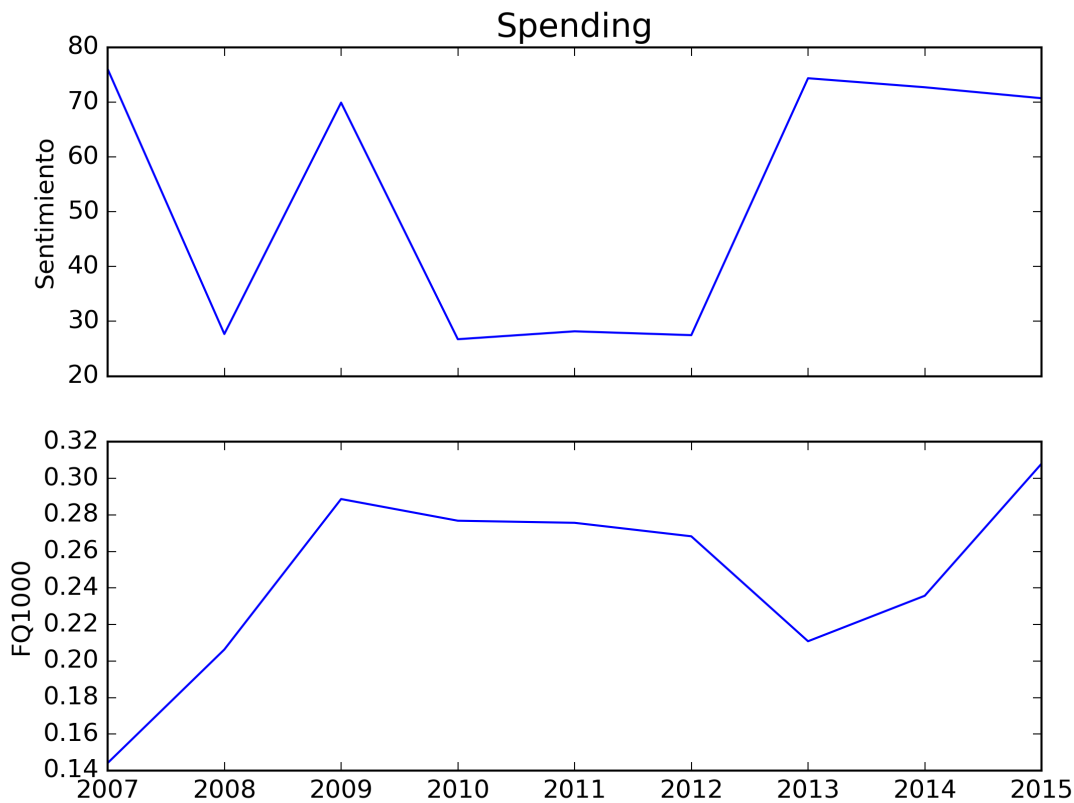
Greek ministers are spending this weekend, almost five grinding years since Athens was first bailed out, wrangling over the details of the spending CUTS and economic REFORMS they have drawn up to appease their creditors.

The ECB has lobbied for hard-hitting structural REFORMS that include wage CUTS to make local economies more competitive.

## 7.9 Término 9: *Spending*

El noveno término analizado en este trabajo es el sustantivo “spending” (“gasto”). El gasto es un concepto clave para diagnosticar el estado de la economía de la población que hace referencia a la suma de toda actividad económica que implique la selección y uso de bienes y servicios. En contraste, “investment” (“inversión”) se refiere al gasto para la adquisición de ingresos futuros.

El interés del estudio de este término recae, principalmente, en el desequilibrio que se provocó entre consumo e ingresos durante esta época. El estallido de la crisis el consumo se congeló a nivel mundial, provocando un efecto en cadena que afectó inmediatamente a la ciudadanía. Por otro lado, la disminución del gasto público (i.e., el realizado por los gobiernos), especialmente en Europa, debido a la crisis supuso la paralización de los logros en los objetivos de desarrollo humano, que dependían de los aumentos de gasto logrados en la década de 1990 y en la primera parte de la década de los 2000 y provocaron un déficit de financiación en salud, educación, seguridad e infraestructura (Brumby y Verhoeven, 2010).



**Figura 26** Gráfico longitudinal de las medias anuales de las puntuaciones de sentimiento para “spending” y la frecuencia relativa del término por cada 1000 tokens.

Se procede entonces a observar el movimiento de la orientación semántica del contexto del término “spending”. A través de los datos aportados por la observación de los valores aportados por la serie temporal de medias de puntuación de sentimiento (Figura 26) y el análisis estadístico de los textos (Tabla 52) se pueden observar varias etapas en las tendencias de la evolución del sentimiento del contexto del término en cuestión:

- (1) una primera etapa anterior a la crisis relativa al año 2007 con una media de sentimiento positivo (76,07).
- (2) una segunda etapa entre 2008 y 2012 con forma de uve doble (27,69, 69,92, 26,74, 28,18 y 27,48), un patrón similar al que sigue la evolución del sentimiento del contexto de “markets”. Con la incertidumbre causada por la crisis la puntuación media de las oraciones que incluyen la palabra “spending” cayó 48,38 puntos hasta la zona negativa. Se produce una subida al umbral de positividad puntual y repentina en el periodo entre la crisis del crédito y la crisis de la deuda en el año 2009, para volver a descender y permanecer durante tres años en el umbral



negativo entre 2010 y 2012. Este descenso entre 2007 y 2008 se justifica por el número de elementos negativos, el cual subió de manera muy significativa y pasó de 2,25 en 2007 a 3,54 en 2008.

- (3) Entre los años 2013 y 2015, los valores medios de sentimiento en las oraciones que incluyen “spending” ascienden súbitamente al umbral de la positividad (74,38, 72,72 y 70,73). La media de unidades positivas aumenta progresivamente, pasando de 3,28 en 2012 a 3,77 en 2013, 3,92 en 2014 y 4,23 en 2015.

La media de las frecuencias relativas oscila en un rango entre 0,14 en 2007 y 0,31 en 2015. Al contrario que en el uso de otros términos, aquí un aumento de la frecuencia relativa no supone generalmente un descenso brusco de las puntuaciones medias de sentimiento. El uso de “spending” se duplicó entre 2007 y 2009 (FQ1000 0,14 y 0,29, respectivamente) y a partir de entonces ha mantenido unos valores moderadamente estables. En el intervalo 2008-2013 la frecuencia relativa mantiene una tendencia ligeramente descendiente (0,28, 0,28, 0,27 y 0,21, respectivamente). Finalmente, se produce un nuevo pico al final del estudio, 2015 (FQ1000 0,31).

Año	N	FQ1000	Sentimiento	pos_items	neg_items	neu_items	none_items
2007	248	0,14	76,07	2,75	2,25	2,00	1,01
2008	383	0,21	27,69	2,75	3,54	2,06	1,11
2009	575	0,29	69,92	3,18	3,69	2,32	1,27
2010	622	0,28	26,74	3,15	4,71	2,71	1,24
2011	746	0,28	28,18	3,33	4,52	2,67	1,29
2012	668	0,27	27,48	3,28	4,57	2,55	1,28
2013	758	0,21	74,38	3,77	3,15	2,51	1,36
2014	925	0,24	72,72	3,92	3,46	2,71	1,44
2015	1337	0,31	70,73	4,23	4,26	2,78	1,60

**Tabla 52** Resumen estadístico del análisis de las oraciones que incluyen “spending”

### Primer intervalo de estudio

La primera fase de análisis corresponde al año previo a la explosión de la crisis del crédito. En este año, se obtienen una media de sentimiento positiva para las oraciones que incluyen el término “spending”. Entre las unidades léxicas con sentimiento dentro del dominio específico, puede encontrarse la colocación “cuts” y marcadores de la direccionalidad en el sentimiento (“reduce” y “rise”).

El análisis realizado mediante *The Sketch Engine* ofreció una lista de colocaciones que permiten observar las relaciones con distintas palabras a lo largo de este intervalo, así como las distintas categorías gramaticales que coocurren con el término objetivo. En este caso son las siguientes:



- Sustantivos genéricos: *ad, asignment, care, education, household, health, habit, tourist*.
- Sustantivos específicos: *consumer, consumption, dollar, money, saving*.
- Adjetivos genéricos: *political, online*.
- Adjetivos específicos: *public*.
- Verbos genéricos: *estimate, rise*.

<i>Palabra</i>	2007	
	<i>Cooc</i>	<i>logDice</i>
education-n	6	9,030
consumer-n	20	8,951
household-n	5	8,858
<b>cut-v</b>	9	8,823
public-j	12	8,811
health-n	10	8,709
reduce-v	8	8,700
estimate-v	7	8,696
care-n	8	8,671
online-j	8	8,433
assignment-n	3	8,395
habit-n	3	8,375
saving-n	4	8,272
tourist-n	3	8,249
ad-n	9	8,179
money-n	13	8,163
political-j	4	8,138
consumption-n	3	8,097
dollar-n	6	8,000
rise-v	7	7,995

**Tabla 53** Principales colocaciones de “spending” durante el primer intervalo de estudio

Dado que no hay una suerte mayoritaria de colocaciones con sentimiento positivo o negativo en este intervalo, es pertinente aquí observar una muestra de oraciones que incluyan el término “spending” en el subcorpus. El uso del término en los ejemplos responde a su utilización normal dentro de reportes de noticias económicas, sin una relación clara con eventos relevantes en la actualidad económica del momento. Se puede observar que la mayoría de las oraciones tienen un sentimiento positivo o neutro y no se observan patrones de colocaciones relacionadas con unidades léxicas relacionadas con una orientación semántica determinada. A continuación, se pueden revisar algunas de las oraciones aleatorias revisadas.

Ad SPENDING last year totaled \$2.8 million, according to TNS Media Intelligence, reflecting the absence of TV commercials.

They found that in more diverse cities and counties, the share of local government SPENDING on public goods – in this case, roads, sewage treatment, trash clearance and education – was generally lower than it was in more homogeneous localities.

It has increased SPENDING on kindergarten to \$3.3 billion this year, from \$2.75 billion, partly using money transferred from its \$350 billion State Pension Fund, once known as the Petroleum Fund.

Public SPENDING in law and order might not go down.

These were the consistency of both sales and earnings growth, and SPENDING in three categories: research and development, advertising and capital expenditures.

And the quickest, surest way to get more consumer SPENDING is to put more cash into the hands of people who live hand-to-mouth.

Over the past two years, the government has more than doubled total health care SPENDING to some \$7 billion, but that still works out to only about 3.4 percent of all government SPENDING, and the World Health Organization recommends at least 5 percent.

El uso de “cut” como verbo (“recortar”) no tenía en 2007 las mismas connotaciones que tendría tras el inicio de la crisis cuando su uso se orientó principalmente a los recortes en servicios públicos. Si bien existen algunos ejemplos en esa dirección (p. ej., el primero del bloque), la mayoría de los ejemplos habla de eliminar gastos superfluos (“extravagant spending”) o de reducir la producción de un producto (“cut back spending on its cars”). A continuación, se ofrecen seis ejemplos encontrados en el corpus para esta colocación:

However, he says Nicolas Sarkozy, or any other politician wanting to CUT back public SPENDING, will have a fight on their hands.

In effect, the report argues for SPENDING the money to CUT emissions because future generations have as much claim on resources as current generations.

But with demanding corporate parents, their Wall Street partners and restive shareholders all looking to CUT down on extravagant SPENDING, studio executives are finding that it is hardly business as usual.

"It would be a tough decision to accept a job that required me to go farther," she said, adding that she and her husband were CUTTING back on discretionary SPENDING until she is employed again.

In 2004, Ford sold a record 939,000 full-size pickups, and nearly two-thirds of its overall sales in the United States were trucks, vans and S.U.V.'s. At the same time it invested in the growing truck market, Ford CUT back SPENDING on its cars.

But there are signs that many Americans are feeling the pinch, and business economists are worried that rising gasoline prices will CUT consumer SPENDING this holiday season.

El resto de los ejemplos destacados aquí para este intervalo están relacionados con unidades de direccionalidad de los marcadores económicos (“reduce” y “rise”). El aumento o el descenso del consumo o gasto será positivo o negativo dependiendo de la entidad a la que haga referencia. Por ejemplo, un aumento moderado del gasto en sanidad o educación será considerado positivo o viceversa. Para ilustrar esto, en el caso de “reduce spending”, mientras que “reduce health care spending” en el primer ejemplo el

sentimiento es considerado positivo dadas las unidades que aparecen su vecindario, en el cuarto ejemplo no lo es. En términos generales, se puede observar que la gran mayoría de ellos son usados en un contexto positivo (“improve”, “reduge budget deficit” o “achieve profitability”).

For Wal-Mart, the payoff could be significant: if it succeeds, the initiative could improve employee morale, and therefore productivity; REDUCE health care SPENDING on a work force with higher rates of heart disease and diabetes than the general public; and improve Wal-Mart's reputation with the image-conscious consumers it is courting with costlier merchandise.

Ethnic diversity doesn't inevitably REDUCE SPENDING on public goods.

The OECD, a thinktank for 30 of the world's richest countries, called on the government to use SPENDING restraint to REDUCE the budget deficit.

But the U.A.W. broke from that tradition in 2005, when it agreed to help General Motors and the Ford Motor Company REDUCE health care SPENDING by shifting some costs onto workers, but refused to grant Chrysler the same deal.

In a statement Monday, Peter Meldrum, the chief executive of Myriad, said the discontinuation of Flurizan would REDUCE SPENDING and allow the company to achieve profitability in the fiscal year that ends June 30, 2009.

SPENDING on technology was REDUCED.

Igualmente, y para concluir este corto intervalo, el uso de “rise” in “spending” dependerá aquí de la entidad con la que se relacione. Igualmente, un alza en el gasto en educación o en el consumo de los turistas será positivo. De los cuatro ejemplos observados, tres de ellos son positivos y en su prosodia semántica aparecen colocaciones con una orientación semántica positiva (“best showing” o “fastest-growing”). A continuación, las oraciones revisadas:

For December, consumer SPENDING ROSE 0.7 percent, the best showing in five months, while incomes rose by 0.5 percent, both figures matching Wall Street expectations.

Education SPENDING in England to RISE from £60bn this year to £64bn, £67bn, £70bn and £74bn in successive years: education to be a right for every young person to the age of 18

SPENDING by Chinese tourists overseas ROSE 11.8 percent to \$24.3 billion in 2006, the fastest-growing of the top 10 nations apart from South Korea, according to the World Tourism Organization.

Like MasterCard, ComScore, a research firm, found that online SPENDING ROSE steadily to \$26.3 billion.

Japan's inflation rose at the fastest pace in more than nine years in November while industrial production and household SPENDING declined, signaling that RISING oil costs might derail the economy's longest postwar expansion.

## Segundo intervalo de estudio

La segunda etapa de estudio corresponde los cinco años que siguieron al estallido de la crisis (2008-2012). A partir de la contracción del sistema crediticio y sus consecuencias, las puntuaciones medias de las oraciones que incluyen a “spending” caen bajo umbral de negatividad, con la salvedad del año 2008. De hecho, a partir del año 2010 la colocación más significativa es “cuts” que reemplaza a su colocación neutro habitual el resto de los años: “consumer”. Se puede apreciar referencias a los recortes del gasto y el consumo relativos a la crisis (“cut”, “slowdown”, “cutback”, “squeeze” o “downturn”), escasas unidades léxicas con sentimiento positivo, destacando “stimulus” en el año 2008 y marcadores descendentes de la direccionalidad en el sentimiento (“reduce” y “rise”).

El análisis realizado mediante *The Sketch Engine* ofreció una lista de colocaciones que permiten observar las relaciones con distintas palabras a lo largo de este intervalo, así como las distintas categorías gramaticales que coocurren con el término objetivo. En este caso son las siguientes:

- Sustantivos genéricos: *ad* (3), *advertising*, *bill*, *construction*, *consumer* (5), *defence*, *government* (4), *habit*, *health*, *household* (4), *increase* (4), *infrastructure* (3), *pattern*, *power* (2), *review*, *rise*, *shopper* (3), *taxis* (4).
- Sustantivos específicos: *capital* (3), *cut* (5), *cutback*, *downturn*, *money*, *slowdown*, *squeeze* (2), *stimulus* (2), *tax* (4).
- Adjetivos genéricos: *comprehensive*, *less*, *military* (3), *much*, *overall*, *wasteful*, *weak*.
- Adjetivos específicos: *discretionary* (4), *public* (5).
- Verbos genéricos: *increase* (5), *reduce* (2), *rise*, *warn*.
- Verbos específicos: *boost*, *cut* (5), *rein*, *slash* (2), *slow*, *track*.

La primera colocación que se analizará este intervalo es la unidad “cuts”. Mientras los ejemplos del año anterior hacían uso de otras construcciones (p. ej., “cuts in spending” o “cuts in health spending”), a partir del año 2008 (*logDice* 8,562) se produce una fijación entre los dos términos y no solo aumenta el uso, sino que se asienta la colocación “spending cuts” la cual hace referencia clara a los recortes causados por la crisis económica (*logDice* 11,257 en 2012). Se observa, entonces, una muestra aleatoria de la colocación durante los años de este intervalo. En primer intervalo de estudio existía mayor diversidad en los agentes que realizan los recortes (empresas, estados, instituciones) pero, sin embargo, en este intervalo todos los agentes hacen referencias a gobiernos de estados

soberanos de manera directa o indirecta (“Washington”, “The Germans”), lo que expresa claramente la relación con un evento importante en marcha. Además, en todos los casos el sentimiento de las oraciones es negativo y muchas unidades léxicas del vecindario de la colocación son también claramente negativas (“procastinates”, “make it more difficult”, “jobs are likely to go”).

And it will stay that way all the while turmoil exists in euroland and Washington procrastinates over planned tax rises and SPENDING CUTS due in January that could plunge the US into recession.

However, there are fears in the industry that the government's SPENDING CUTS could make it more difficult for companies to come up with the £10bn needed to develop the new plants and pipelines.

The Germans refuse to budge, saying they want all eurozone countries to contribute to the insurance fund and SPENDING CUTS to deal with the situation.

Impending military SPENDING CUTS were the talk of the show . Italy's decision to reduce its order of Eurofighter jets by 25 planes to save almost \$2.6bn was viewed as a sign of things to come.

Unemployment is still rising across the economy, and thousands of public sector jobs are likely to go as part of the £6bn SPENDING CUTS pledged by the Conservatives.

Entre otros términos negativos que aparecieron en el ranking de colocaciones del año 2008 está “slowdown” (“ralentización”), el cual tiene un sentimiento negativo en el dominio de la economía. El uso del término hace referencia a la ralentización del consumo y las consecuencias nefastas que esto tenía para las empresas privadas. Tras una revisión cualitativa, la mayoría de los ejemplos que incluyen esta colocación tienen sentimiento negativo y a su alrededor aparecen unidades léxicas como “ominous”, “victim” o “tightness”.

SPENDING SLOWDOWN drags on air travel Jet aeroplane taking off at night.

"We're aware that people are talking about the consumer SPENDING SLOWDOWN but we're just not seeing it," said ABF finance director John Bason.

The chain appears to be the first major American retailer to make job cuts because of a SLOWDOWN in consumer SPENDING , and the move may be ominous sign of what is to come across the industry.

Aronson, the former chief executive of Saks Fifth Avenue and a managing director at Kurt Salmon Associates, a retail consulting firm, said the credit tightness and consumer SPENDING SLOWDOWN have only wiped out the "bottom tier" companies in retailing.

NexCen appears to be the latest victim of a tight credit market and a SLOWDOWN in consumer SPENDING , which have helped push nearly a dozen big chains – like Linens 'n Things and The Sharper Image – into bankruptcy over the last year.

2008			2009			2010		
<i>Palabra</i>	<i>Cooc</i>	<i>logDice</i>	<i>Palabra</i>	<i>Cooc</i>	<i>logDice</i>	<i>Palabra</i>	<i>Cooc</i>	<i>logDice</i>
consumer-n	82	10,852	consumer-n	104	10,889	<b>cut-n</b>	94	11,347
<b>rein-v</b>	10	9,631	<b>cut-n</b>	24	9,584	consumer-n	78	10,286
<b>slowdown-n</b>	10	9,223	<b>cut-v</b>	33	9,580	public-j	57	10,037
<b>cut-v</b>	21	9,223	household-n	17	9,438	review-n	30	9,959
increase-v	18	9,052	public-j	28	9,402	<b>cut-v</b>	41	9,925
<b>cutback-n</b>	6	8,840	<b>stimulus-n</b>	15	9,223	military-j	15	9,331
habit-n	6	8,787	infrastructure-n	12	9,133	government-n	55	9,041
much-j	13	8,780	government-n	45	8,909	ad-n	22	9,036
household-n	7	8,771	increase-v	18	8,773	capital-n	19	8,992
back-a	17	8,726	rise-v	18	8,719	increase-v	22	8,809
<b>slash-v</b>	6	8,684	military-j	9	8,670	tax-n	24	8,769
discretionary-j	5	8,642	reduce-v	15	8,660	comprehensive-j	9	8,741
<b>cut-n</b>	9	8,562	increase-n	12	8,597	<b>stimulus-n</b>	10	8,720
public-j	11	8,560	<b>wasteful-j</b>	7	8,540	increase-n	15	8,656
infrastructure-n	5	8,470	overall-j	9	8,535	taxis-n	11	8,579
ad-n	11	8,405	taxis-n	9	8,504	discretionary-j	7	8,408
<b>downturn-n</b>	7	8,339	less-j	13	8,441	track-v	8	8,394
capital-n	9	8,211	tax-n	18	8,423	<b>slow-v</b>	8	8,368
shopper-n	5	8,147	bill-n	10	8,413	<b>weak-j</b>	9	8,365
construction-n	5	8,132	health-n	12	8,348	<b>warn-v</b>	12	8,293
2011			2012					
<i>Palabra</i>	<i>Cooc</i>	<i>logDice</i>	<i>Palabra</i>	<i>Cooc</i>	<i>logDice</i>			
<b>cut-n</b>	94	11,205	<b>cut-n</b>	94	11,257			
consumer-n	136	10,693	consumer-n	79	10,199			
<b>cut-v</b>	38	9,552	increase-n	32	9,602			
tax-n	46	9,391	<b>cut-v</b>	31	9,436			
household-n	23	9,370	government-n	69	9,242			
public-j	35	9,177	public-j	32	9,178			
ad-n	24	9,109	tax-n	39	9,109			
government-n	60	8,994	household-n	16	9,008			
<b>squeeze-n</b>	13	8,957	power-n	20	8,799			
discretionary-j	12	8,919	increase-v	25	8,765			
taxis-n	15	8,780	infrastructure-n	12	8,675			
increase-n	22	8,752	taxis-n	13	8,619			
increase-v	26	8,697	<b>boost-v</b>	11	8,569			
shopper-n	14	8,639	military-j	9	8,553			
rise-n	16	8,543	advertising-n	12	8,392			
money-n	26	8,480	<b>slash-v</b>	8	8,350			
capital-n	15	8,400	defence-n	8	8,322			
<b>power-n</b>	18	8,363	discretionary-j	7	8,309			
reduce-v	18	8,308	shopper-n	10	8,307			
pattern-n	8	8,189	<b>squeeze-n</b>	7	8,252			

**Tabla 54** Principales colocaciones de “spending” durante el segundo intervalo de estudio

La media de puntuaciones positivas para el año 2009 puede justificarse a partir del análisis de “stimulus”. Al igual que pasó con las puntuaciones de “markets” durante el año 2009 se tomaron algunas medidas económicas de estímulo fiscal, especialmente en los Estados Unidos, que supusieron un breve alivio para unas economías famélicas tras la contracción del crédito el año anterior. Estas fueron el *Economic Stimulus Act* de 2008 promulgada por el gabinete Bush y la *Recovery Act* de 2009 promulgada por el gabinete Obama. El objetivo de estas leyes era la salvación de los empleos existentes y la creación de nuevos empleos. A pesar del entusiasmo con el que la prensa las recibió, las medidas tomadas tuvieron un éxito muy moderado pretendían implementar programas de ayuda temporal para los más afectados por la crisis y la inversión en infraestructura, educación, salud y energías renovables.

El término “stimulus” tiene un sentimiento inherentemente positivo en el dominio de la economía. Además, su prosodia semántica también es positiva y entre las colocaciones que se observan en los ejemplos se encuentran elementos como “enthusiasm”, “fresh” o “success”. A continuación, se pueden observar algunos ejemplos:

"The new enthusiasm for fiscal STIMULUS , and particularly government SPENDING, represents a huge evolution in mainstream thinking," said Janet Yellen, president of the Federal Reserve Bank of San Francisco.

" The STIMULUS SPENDING bill signed last month includes \$4.5 billion for job training.

The \$787 billion STIMULUS SPENDING bill signed by President Obama last month is expected to generate fresh demand for goods and services.

The funding is needed to help finance the hundreds of billions of dollars that the government is SPENDING on bank bailouts and STIMULUS .

Analysts said this is due to the success of official STIMULUS packages, including INCREASED spending on infrastructure projects and cash handouts to encourage consumers to keep spending.

The new STIMULUS plan involves emergency SPENDING of ¥1.9 trillion to protect jobs and retrain unemployed workers, according to a draft released by the Liberal Democratic Party before the formal announcement set for Friday.

El verbo “slash” tiene un significado similar a “cut” pero su uso tiene una connotación de mayor vehemencia. Su uso fue especialmente significativo en los años 2008 y 2012 (*logDice* 8,684 y 8,350) en los que el gasto público se redujo considerablemente en la Unión Europea. Se encuentran referencias a los dolorosos recortes de gasto público que se produjeron en los distintos países afectados por la crisis. Todos los ejemplos observados en los que aparece “slash” son negativos y entre sus colocaciones se encuentran elementos negativos como “forced”, “deteriorated”, “crisis”



o “painfully”. A continuación, se pueden observar algunos de los ejemplos extraídos del corpus:

The housing collapse, coupled with the global credit crisis, forced the Irish government to SLASH SPENDING to keep its deficit in check and pushed the benchmark stock index to fall more than any other in western Europe this year.

A report by experts at Citigroup this week suggested that if conditions deteriorated further, Ford and GM could be forced into "drastic SPENDING cuts" including SLASHING jobs and making new pay deals with unions – or even debt-for-equity exchanges, in which the banks would take ownership of part of the two companies.

Europe's 30-month effort to save the euro by SLASHING SPENDING and debt levels risks turning into a crisis of political legitimacy after EU leaders' strategies collided spectacularly with the wishes of voters in Greece and France.

The catch is that this is a "medium term" objective, and progress has been painfully slow even with the need in most countries to SLASH government SPENDING .

It has been forced to raise taxes and SLASH SPENDING for years – and that won't stop until at least 2015.

El último término de este segundo intervalo de análisis es “squeeze” (“presión”) que tuvo un auge en su uso en el intervalo 2011-2012 (*logDice* 8,957 y 8,252). El uso del término hace referencia, principalmente, al descenso en el consumo de los hogares tras la caída del poder adquisitivo que provocó serios problemas económicos a las empresas privadas (p. ej., “Flybe” o “BSkysB”). Según la observación de los ejemplos en el corpus, la colocación de dichos términos tiene lugar en oraciones con sentimiento negativo. Más abajo, se puede observar un bloque con algunos ejemplos:

Mervyn King , governor of the Bank of England, recently warned that families face the worst SQUEEZE on their SPENDING power since the 1920s, with inflation-adjusted wages down over the past six years.

That will fall to £92bn in the next financial year, the OBR believes, and £21bn, or 1.1% of GDP by 2016-17 as the SPENDING SQUEEZE takes hold in earnest.

UK quoted companies have just limped through their toughest three months since the financial crisis, as the SQUEEZE on household SPENDING , soaring commodity prices and rising inflation have hammered consumer and business spending.

But Keith Bowman, analyst at Hargreaves Lansdown, said that the continuing SQUEEZE on consumer SPENDING was bad news for leisure companies, which were also coping with higher costs as coffee and food prices continue to rise.

Flybe profit warning sends share price crashing down Flybe shares drop 25% as airline admits cash-strapped consumers are cutting back on air travel Flybe shares slumped after the Exeter-based carrier issued a profit warning on 5 May 2011, blaming the SQUEEZE on UK consumer SPENDING.

Signs of BSKyB slowdown bring more bad news for investors BSKyB investors braced for disappointing growth forecasts with full-year results BSKyB's full-year results are expected to show signs of a slowdown in growth as the consumer SPENDING SQUEEZE hits Sky subscribers.



### Tercer intervalo de estudio

La tercera etapa corresponde los tres últimos años de este estudio (2013-2015). En este periodo, Europa seguía sumida en una crisis no solo política, sino también institucional. Sin embargo, las puntuaciones medias de las oraciones que incluyen a “spending” vuelven al umbral de positividad. Para ilustrar esto, a partir del año 2014 la colocación más significativa del ranking pasa de “cuts” a “consumer”, la colocación neutra que encabezaba durante el primer intervalo de estudio. Se puede apreciar que, si bien persisten unidades léxicas relativas a la crisis, estas son más escasas (esencialmente “cut” y “slash”), y aparecen referencias al crecimiento (“growth”, “spree” o “boost”), destacadamente positivas durante una etapa de crisis.

El análisis del corpus realizado con la herramienta *The Sketch Engine* devolvió una serie de colocaciones (Tabla 55) que permiten observar las relaciones con distintas palabras a lo largo de este intervalo, así como las distintas categorías gramaticales que coocurren con el término objetivo. En este caso son las siguientes:

- Sustantivos genéricos: *ad* (2), *bill*, *care*, *consumer* (3), *government* (3), *growth* (2), *health* (2), *household* (3), *increase* (2), *infrastructure* (2), *plan*, *power*, *programme* (2), *review*, *spree*, *tax* (3), *taxis*, *welfare*.
- Sustantivos específicos: *banking*, *capital*, *cut* (3), *reform*.
- Adjetivos genéricos: *discretionary*, *high*, *military* (2), *overall*, *total* (2).
- Adjetivos específicos: *public* (3), *structural* (3).
- Verbos genéricos: *increase* (2), *rise*.
- Verbos específicos: *boost*, *cut* (3), *slash*.

El término “cut” sigue utilizándose en este intervalo tanto con tiempos en presente como con tiempos en pasado, con referencias históricas a los eventos de años anteriores. En todos los casos examinados, su uso tiene un sentimiento general negativo. Más adelante, se ofrecen algunos ejemplos:

A fiscal consolidation begun in 2010 is continuing, in which tax increases and SPENDING CUTS are being applied to bring the overall budget deficit down to 3 percent of gross domestic product in 2013, from an estimated 4.5 percent in 2012.

The troika of lenders insisted on a broad set of reforms and public SPENDING CUTS as the price for the loans under a programme that ran out in February and was extended until the end of June.

The price rise of 34% since mid-January has largely been fueled by CUTS to capital SPENDING and falling US rig counts, which have yet to result in a fall in near-term production.

The intractable problem is that many governments do not trust the Greek government to implement a €12bn (£8.6bn) programme of SPENDING CUTS and reforms that will be delivered as part of a bailout.

But Carney conceded that the government's SPENDING CUTS had been "material" and economists warned that there will be further pressures on the UK economy when the next round of austerity is revealed by chancellor George Osborne at his spending review later this month.

Greek ministers are spending this weekend, almost five grinding years since Athens was first bailed out, wrangling over the details of the SPENDING CUTS and economic reforms they have drawn up to appease their creditors.

2013			2014			2015		
Palabra	Cooc	logDice	Palabra	Cooc	logDice	Palabra	Cooc	logDice
<b>cut-n</b>	52	10,348	consumer-n	108	9,907	consumer-n	218	10,724
consumer-n	73	9,611	<b>cut-n</b>	47	9,809	<b>cut-n</b>	126	10,580
government-n	71	9,070	public-j	45	9,195	public-j	69	9,407
discretionary-j	13	8,969	infrastructure-n	25	9,122	capital-n	56	9,338
tax-n	38	8,870	household-n	27	9,082	<b>cut-v</b>	67	9,291
military-j	14	8,850	structural-j	18	8,938	household-n	41	9,221
banking-n	22	8,839	tax-n	40	8,838	structural-j	29	9,117
increase-v	34	8,758	government-n	63	8,824	government-n	108	9,054
<b>cut-v</b>	20	8,554	ad-n	19	8,540	increase-v	62	8,951
health-n	23	8,515	<b>care-n</b>	17	8,530	<b>boost-v</b>	32	8,902
household-n	14	8,513	health-n	23	8,432	review-n	33	8,871
ad-n	19	8,479	military-j	12	8,400	tax-n	60	8,765
increase-n	18	8,419	increase-n	21	8,378	plan-n	51	8,741
public-j	24	8,344	<b>cut-v</b>	23	8,372	programme-n	25	8,458
bill-n	14	8,324	programme-n	16	8,294	reform-n	22	8,332
total-j	12	8,167	taxis-n	12	8,213	infrastructure-n	19	8,268
structural-j	8	8,142	high-j	36	8,132	rise-v	40	8,258
<b>growth-n</b>	26	8,105	<b>welfare-n</b>	10	8,122	<b>slash-v</b>	16	8,238
<b>spree-n</b>	7	8,084	total-j	13	8,052	power-n	27	8,222
overall-j	10	8,083	increase-v	25	8,031	<b>growth-n</b>	59	8,197

**Tabla 55** Principales colocaciones de “spending” durante el tercer intervalo de estudio

En los rankings de los años 2013 y 2015 aparece el término “growth” que, como se ha visto a lo largo de este trabajo, indica un factor deseable dentro del dominio económico. Dado que en ocasiones se habla de p. ej., “slowdown in the growth of spending” y el contexto de esta colocación no es siempre positivo, en una mayoría de los casos se habla de previsiones de crecimiento gracias al consumo. Algunos de los ejemplos son los siguientes:

However, the forecast said inflation in major building materials would affect the comparison between 2012 and 2013; it said SPENDING GROWTH in 2012, on a constant dollar basis, was 2.9 percent; and it predicted that SPENDING GROWTH would jump 3.4 percent this year.

Despite a recent slowdown in the GROWTH of SPENDING on overall health care, Republicans and Democrats alike have warned that rising health costs will eventually overwhelm the federal budget and make basic health care unaffordable for many Americans.

This has, in turn, affected consumer SPENDING and business GROWTH, resulting in a substantial decrease in the UK's GDP.

More money circulating in the economy should lead to higher prices, and help generate a positive cycle of more investment, profits, wages, SPENDING and GROWTH .

Summers of Harvard, President Obama's former chief economic adviser; and Brad DeLong of the University of California, Berkeley – proposed in a recent paper that at the rock-bottom interest rates that prevail today, government SPENDING to encourage GROWTH would in fact pay for itself.

One consequence could be higher short- and longer-term interest rates as markets reassess the UK's willingness to accumulate sovereign debt, and the Bank of England responds to higher near-term GROWTH thanks to stronger government SPENDING.

El conjunto “spending spree” (*logDice* 8,084) se puede traducir al castellano de varias maneras y hace referencia a “oleadas de gasto” u “oleadas de consumo” con una connotación positiva. Si bien no se encuentran patrones que lo correlacionen con ningún evento en concreto, se pueden observar referencias a la mejora de los resultados del deprimido mercado inmobiliario debido a la inversión asiática (como en el caso de “investor” en 7.6). Algunos de los ejemplos en los que aparece en el corpus son los siguientes:

A foreign SPENDING SPREE , led by Russians, Chinese, Americans, Emiratis and Indians, resulted in the number of overseas buyers overtaking the number of Britons buying London properties worth more than £2m in the 12 months to April 2013, according to new research by estate agent Knight Frank.

And despite the public's belief that Washington is engaged in a SPENDING SPREE , the deficit in the United States narrowed to 7 percent of G.D.P. in 2012 from 10.1 percent in 2009.

But David Kern, chief economist at the British Chambers of Commerce, said a SPENDING SPREE was a good start to a more sustained upturn.

It could be that the business sector, already using outdated software and ageing equipment, is about to embark on an investment SPENDING SPREE .

But analysts believe the muted start to the traditional festive SPENDING SPREE also reflects the squeeze on real incomes caused by prices rising more rapidly than earnings.

Las políticas de *Quantitative Easing* impulsadas por el BCE, junto a las bajas cifras de inflación en el año 2015 fueron uno de los factores por los que “spending” obtiene una media de puntuaciones de sentimiento positivas en este último intervalo. La colocación “boost” (“impulsar” o “disparar”), con un coeficiente *logDice* de 8,902, es usada en contextos positivos en la mayoría de los dominios y es recurrente en los usos de la prensa económica. Aunque tras la observación de algunos ejemplos aleatorios extraídos del corpus se evidencian opiniones que mezclan factores positivos y negativos (p. ej., “can boost confidence and unlock spending – but that tends to wane after a while”), en la mayoría de los casos el sentimiento es positivo. Algunos de los ejemplos son los siguientes:

He said the steep fall in UK inflation – expected to turn negative in the spring – would BOOST consumer SPENDING, and in turn economic growth.

The governor said lower energy and food costs would underpin the fastest growth in real (inflation-adjusted) take-home pay for a decade in 2015, BOOSTING consumers' SPENDING power.

National governments could embark on a SPENDING spree to BOOST activity, but most would breach EU borrowing limits.

There were tears for Boohoo investors as it warned profits would be more than 25% below City expectations after heavy discounting meant extra SPENDING on marketing failed to BOOST sales.

However, lower oil prices – which have more than halved since last summer – are expected to significantly BOOST consumer SPENDING.

Economists say, in contrast to the struggling manufacturing and construction sectors of the economy, consumer SPENDING is being BOOSTED by a gradual pick-up in wages, continued record low interest rates and low inflation.

The powerful message QE sends, that the central bank is determined to get to grips with an economic crisis, can BOOST confidence and unlock SPENDING – but that tends to wane after a while.

Finalmente, en el año 2015 aparecen 27 coocurrencias de “spending” y “power” (*logDice* 8,222), que en casi la totalidad de los casos hace referencia al término “spending power” (“poder adquisitivo”), que hace referencia a la posibilidad de gasto de una entidad (p. ej., individuo, familia, etc.). Tras más de un lustro con caídas en el consumo, el poder adquisitivo aumentó ese año debido a varios factores (p. ej., una tasa de baja inflación) y se vio claramente reflejado en las menciones a “spending” en la prensa. De los ejemplos extraídos para realizar un análisis cualitativo, todos tienen un sentimiento general positivo y son muchas las coocurrencias que se observan con un sentimiento positivo (p. ej., “boost”, “recovering”, “strong ... confidence”). Una muestra de los ejemplos observados se expone a continuación:

The worry is therefore that weak pay growth means the economy is reliant on ultra-low inflation to boost consumer SPENDING POWER.

British consumers are at their most upbeat for nine years on the back of improving job prospects and low inflation, according to the latest report to suggest SPENDING POWER is recovering.

Low inflation, which we expect to remain below 1% for the rest of the year, has given household incomes a much-needed boost and greater SPENDING POWER .

Ian Stewart, chief economist at consultants Deloitte commented: Falling prices raise consumer SPENDING POWER and help keep interest rates low.

A stellar increase in activity in the business and professional service sector and retail sales bounding ahead are clear indications of strong business and consumer confidence and increased SPENDING POWER .

Falling oil prices have been treated as a boon in the UK, at least for now, because they have increased the SPENDING POWER of consumers.

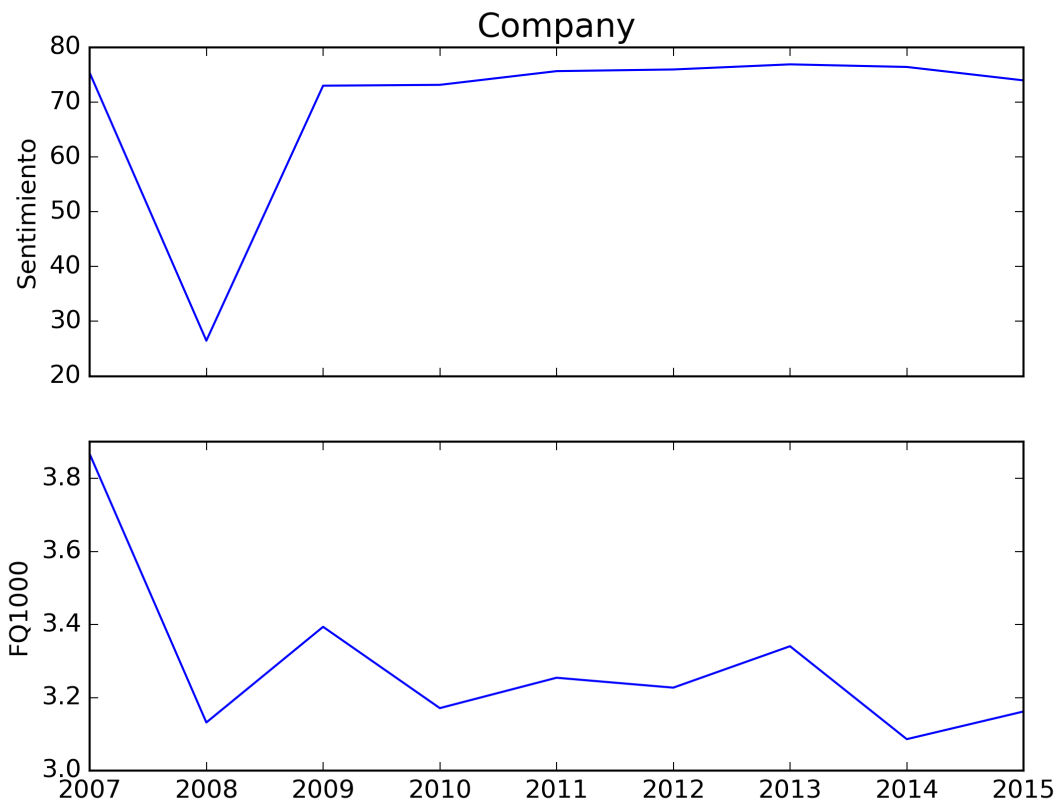
## 7.10 Término 10: *Company*

El décimo término analizado es el sustantivo “company” (“empresa” o “compañía”) y hace referencia a las entidades legales formadas por una asociación de personas, físicas o jurídicas, para llevar a cabo un emprendimiento comercial con vistas a obtener beneficios.

En este caso, no se puede considerar que el sustantivo “company” fuera una *palabra evento*. Este término tiene una frecuencia de aparición significativamente más alta que el resto de las unidades analizadas. Las cifras tan altas y estables de frecuencia se justifican en que las empresas son uno de los objetos de estudio centrales de la prensa económica y su existencia o existencia no estuvo puesta en cuestión, ni fue foco de debate durante los años de la crisis.

El interés del estudio de la evolución del sentimiento de este término recae en la naturaleza de no ser una *palabra evento*. A través de los datos aportados por la observación de los valores aportados por la serie temporal de medias de puntuación de sentimiento (Figura 27) y el análisis estadístico automatizado de los textos (Tabla 56) se pueden observar que, sin contar la puntuación del año 2008 (26,47), las puntuaciones permanecen estables en el umbral de la positividad en un rango de 76,91 en 2013 y 73,01 en 2009. La media de puntuación del año 2008 se puede justificar con el varapalo económico que supuso el mayor colapso bancario desde el *Crack del 27* y el shock que eso supuso, aunque es pertinente realizar una observación cualitativa de una pequeña muestra de las oraciones para profundizar en sus detalles. De hecho, la cifra de ítems negativos pasó de 0,80 en 2007 para descender a 0,93 en 2008. Si bien en 2009 y 2010 estas cifras también fueron muy altas (0,94 y 0,99), los ítems positivos se aumentaron rápidamente pasando del mínimo histórico 0,89 en 2008 a 0,97 en 2009.

La media de las frecuencias relativas oscila en un rango entre un mínimo de 3,09 en 2014 y un máximo 3,87 en 2007. La prensa económica pasó de hablar del devenir cotidiano de las empresas con sus alzas y caídas a narrar el colapso del sistema crediticio y sus severas consecuencias durante el resto de la serie histórica. Por ello, al contrario que en la mayoría de los casos estudiados anteriormente, el dato de mayor frecuencia relativa tiene lugar al inicio del estudio, la cual es muy superior al resto de las frecuencias tras el inicio de la crisis.



**Figura 27** Gráfico longitudinal de las medias anuales de las puntuaciones de sentimiento para “company” y la frecuencia relativa del término por cada 1000 tokens

Año	N	FQ1000	Sentimiento	pos_items	neg_items	neu_items	none_items
2007	6654	3,87	75,38	0,91	0,80	0,85	10,36
2008	5814	3,13	26,47	0,89	0,93	0,87	10,23
2009	6759	3,39	73,01	0,97	0,94	0,87	10,56
2010	7125	3,17	73,17	1,06	0,99	0,97	10,94
2011	8806	3,25	75,67	1,13	0,88	0,99	10,77
2012	8035	3,23	75,97	1,14	0,88	0,99	10,87
2013	12010	3,34	76,91	1,21	0,85	0,98	10,79
2014	12109	3,09	76,43	1,25	0,90	1,04	11,22
2015	13724	3,16	73,98	1,12	0,97	1,00	10,99

**Tabla 56** Resumen estadístico del análisis de las oraciones que incluyen “company”

También existen diferencias con el resto de las unidades de estudio cuando se observan los rankings de colocaciones para este término. En cuanto a las colocaciones, no se observan unidades léxicas que no sean neutras. Además, hay que destacar que gracias al ordenamiento por coeficiente *logDice* (que permite destacar colocaciones exclusivas, pero no raras) se puede observar que, prácticamente, no existen colocaciones destacables. De hecho, no se ha dado ningún caso en el que destaquen verbos auxiliares



o adjetivos modificadores (p. ej., “other”, “many”) o con menores implicaciones en su significado léxico. Por ello, se harán una revisión general de ejemplos para las tres etapas: (a) el año 2007 previo al estallido de la crisis, (b) una puntual bajada al umbral de la negatividad en el año 2008 con la explosión de la crisis, (c) un nuevo ascenso a las medias de sentimiento positivo durante el resto del estudio 2009-2015.

### Primer intervalo de estudio

En este primer intervalo de estudio correspondiente al año 2007 destacan unidades léxicas no específicas, principalmente verbos, como “have”, “say” o “be” y algunas unidades pertenecientes al dominio con sentimiento neutro: “executive”, “business” o “share”. El listado de colocaciones extraídas con la suite de análisis de corpus es el siguiente:

- Sustantivos específicos: *business, executive, share, year*.
- Verbos específicos: *buy, run, sell*.
- Adjetivos genéricos: *large, last, like, many, more, new, other*.
- Adjetivos específicos: *private*.
- Verbos genéricos: *be, have, include, make, say*.

Una revisión de una muestra aleatoria de oraciones procedentes del subcorpus 2007 devuelve una mayoría de oraciones con un sentimiento general positivo o neutro, lo que corresponde con las puntuaciones obtenidas con el análisis con *Lingmotif*. Las oraciones tienen relación con el cotidiano en las noticias de empresas y no se puede observar ningún patrón narrativo relacionado con la crisis. Algunas de las oraciones observadas son las siguientes:

In any transaction involving Iraqi oil, Ms. Woertz wrote that the company should consider the "identity, experience and reputation of the selling COMPANY," as well as "any deviation of the proposed pricing basis or margin for the transaction from historical practice.

One of the key growth targets for the new COMPANY will be in the booming internet end of the travel market.

The target of the lawsuit was a COMPANY controlled or owned by Mr. Zong and his wife and daughter, who are listed as the company's legal representatives and who live in California.

Even so, workers at a COMPANY like Google are in a highly competitive environment.

He is survived by two sons, Christian, who is now the COMPANY president, and Olivier, its managing director, both of Paris, and five granddaughters.

The combined COMPANY would have estimated proven reserves of about 2 trillion cubic feet of natural gas equivalents, with about 70 percent of that natural gas, and daily production of 520 million cubic feet of equivalent.



Mel Karmazin, the chief executive of Sirius, who has argued that a merger will bolster shareholder value, would become chief executive of the combined COMPANY, while Gary Parsons, currently chairman of XM, would become its chairman.

"In cases such as these, we must look at the actions of a COMPANY'S insiders to help us gain information that is not forthcoming in more tangible disclosures," Mr. Gabele said.

<i>Palabra</i>	<i>2007</i>	
	<i>Cooc</i>	<i>logDice</i>
have-v	1008	10,265
say-v	840	10,256
be-v	1778	10,015
other-j	197	9,331
make-v	204	9,224
sell-v	167	9,191
like-i	176	9,172
large-j	144	9,078
executive-n	148	8,982
year-n	199	8,940
new-j	134	8,783
buy-v	120	8,770
last-j	131	8,718
more-j	130	8,653
share-n	108	8,636
many-j	112	8,632
business-n	118	8,567
include-v	106	8,538
private-j	91	8,483
run-v	89	8,441

**Tabla 57** Principales colocaciones de “company” durante el primer intervalo de estudio

### Segundo intervalo de estudio

Este segundo tramo de análisis corresponde al año 2008, en el que se produjo el estallido de la crisis y las puntuaciones descendieron repentinamente durante este año. Como en el intervalo anterior, también destacan unidades léxicas no específicas. Sin embargo, aquí llama la atención la presencia del sustantivo “insurance” (“seguro”), en referencia a las compañías aseguradoras. Este término se repite también en 2009 y puede suponer un indicio para observar cualitativamente los factores de la caída de las puntuaciones de sentimiento. El listado de colocaciones extraídas con la herramienta de análisis de corpus es el siguiente:

- Sustantivos genéricos: *oil*, *year*.
- Sustantivos específicos: *business*, *executive*, *insurance*, *share*.
- Adjetivos genéricos: *big*, *large*, *other*.

- Adjetivos específicos: *financial*.
- Verbos genéricos: *be, become, have, make, own, say, sell, use*.
- Verbos específicos: *buy, pay*.

<i>Palabra</i>	<i>2008</i>	
	<i>Cooc</i>	<i>logDice</i>
have-v	919	10,078
say-v	707	10,002
be-v	1592	9,785
other-j	170	9,234
large-j	138	9,171
share-n	137	9,059
make-v	164	9,042
executive-n	138	9,040
oil-n	114	8,933
big-j	120	8,906
year-n	188	8,885
buy-v	110	8,779
sell-v	100	8,602
insurance-n	82	8,591
financial-j	99	8,586
pay-v	89	8,486
use-v	87	8,404
become-v	81	8,381
own-v	73	8,381
business-n	90	8,308

**Tabla 58** Principales colocaciones de “company” durante el segundo intervalo de estudio

La observación de una muestra aleatoria de oraciones obtenida del subcorpus 2008 devuelve una mayoría de oraciones con un sentimiento general negativo o neutro, lo que corresponde con las puntuaciones obtenidas con el análisis con *Lingmotif*. Las oraciones aquí tienen relación con el cotidiano en las noticias de empresas durante la crisis ya que se observan aquí muchas referencias a la caída de las acciones de empresas privadas. Entre las palabras del vecindario de esta colocación se encuentran unidades léxicas negativas, entre otras: “debt rose”, “failed”, “unnerved”, “dangerous” o “lost”. Algunas de las oraciones observadas son las siguientes:

The cost to insure the COMPANY'S debt rose to a record high Wednesday, another sign that investors are increasingly nervous about the company's ability to pay back its loans.

In 2001, the COMPANY got a bit of a publicity lift when Mr. Armstrong went to the White House and presented President Bush with a Trek.

In quarterly results last week, the COMPANY said it had delivered only 84 aircraft over the three months to September, falling 35 planes short of its 119 scheduled deliveries.

The creeping devaluation, far from reassuring Russians, has unnerved COMPANIES and individuals alike.

The COMPANY also said it would write off \$3 billion from its exposure to derivatives linked to the U.S. mortgage market, including \$1.6 billion related to risks in residential housing and \$800 million related to U.S. bond insurance companies.

"It is clear to me that financial services and insurance has gotten away from its core competence and that is dangerous," said Dinallo, who compared such COMPANIES' strategies to driving a car while putting on make-up, looking at a BlackBerry and having a conversation.

An equities trader at Renaissance Capital, a Russian investment bank, lost the COMPANY about \$10 million by placing unauthorized bets on the stock market just as it collapsed, the bank said Monday.

The COMPANIES' share prices have plunged about 90 percent this year, wiping out about \$70 billion of shareholder value.

La única colocación destacable en este intervalo es “insurance”. Esta colocación ocurre como referencia a las pérdidas que tuvieron que asumir las aseguradoras con el colapso de las empresas financieras y, destacadamente, a la quiebra de AIG, la principal aseguradora del mundo, como se relata en el Capítulo 2. La mayoría de las colocaciones observadas a partir de una muestra aleatoria obtenida del corpus tiene un sentimiento negativo.

En resumen, a fines de 2008, el gobierno de los Estados Unidos asumió el control de AIG tras rescatarla por 180 mil millones de dólares como medida de evitar el colapso de los “big five”. Esto tiene relación con lo revisado en el análisis de “credit” durante este mismo año, AIG vendió masivamente los productos denominados “credit default swaps” sin depositar la garantía inicial. En cuanto a la prosodia semántica, la colocación tiene lugar cerca de elementos negativos como “hammered”, “downgraded” o “disastrous”. Algunos ejemplos encontrados en el corpus son los siguientes:

Some of its INSURANCE COMPANIES ended up with mortgage-backed securities on their books, but the real trouble involved the insurance that its financial products unit offered investors for complex debt securities.

Greenberg , has won a victory in his fight against the INSURANCE COMPANY that he ran for nearly four decades.

With most of Woolworths' businesses in administration, and the remaining operations expected to be sold off, trustees must work out the cost of buying out liabilities through a deal with an INSURANCE COMPANY .

With defaults rising, COMPANIES that sold default INSURANCE are getting hammered.

Lee's family controlled the group by holding a big stake in its amusement-park operator, Everland, which holds a stake in Samsung's life INSURANCE COMPANY , which owns shares of Samsung Electronics, which in turn has a stake in the credit card company.

The urgency of the talks grew by late Monday as A.M. Best Company, a credit rating organization specialized in INSURANCE and health care COMPANIES, downgraded the credit of A.I.G. and several of its major subsidiaries.

The INSURANCE COMPANY Aegon said Tuesday that it had taken an investment of 3 billion euros or \$3.7 billion from the Dutch government, acting to shore up its cash position in the face of the continuing financial crisis.

The US government was on the brink of a deal to take control of the world's largest INSURANCE COMPANY, AIG, in a last-ditch intervention to avert a potentially disastrous corporate collapse.

Société Générale also said Thursday that it would take a write-down of 2.05 billion euros in the fourth quarter, including 1.1 billion euros related to the United States housing market and 550 million euros related to American bond INSURANCE COMPANIES .

The governor of AIG's home state of New York, David Paterson, has granted the COMPANY an exemption from INSURANCE regulations to allow it to raid its subsidiaries' reserves for \$20bn.

### Tercer intervalo de estudio

El último intervalo es el más largo de los observados en este trabajo y abarca siete años: 2009-2015. Durante este intervalo, tanto las medias de puntuaciones de sentimiento como las cifras de frecuencia relativa como las colocaciones observadas (Tabla 59) son muy similares. Entre las colocaciones extraídas con *The Sketch Engine* para cada año destacan unidades léxicas no específicas, principalmente verbos, como “have”, “say” o “be” y algunas unidades pertenecientes al dominio con sentimiento neutro. Además del sustantivo “insurance” en el año 2009 (ya analizado), la única unidad que pueda contener sentimiento en este dominio es el verbo “help”, el cual aparece solo en el año 2011.

A continuación, se puede observar una lista de colocaciones clasificada por categoría gramatical, por frecuencia y por año:

- Sustantivos genéricos: *energy, oil* (6), *world, year* (5).
- Sustantivos específicos: ‘business’ (3), *executive* (7), *insurance, market, profit, share* (6),
- Adjetivos genéricos: *big* (7), *large* (7), *last, many* (6), *more* (6), *new* (7), *other, small* (2).
- Verbos genéricos: *be* (7), *do* (6), *have* (7), *help, include* (4), *make* (7), *say* (6), *take* (5), *use* (5), *work* (3).
- Verbos específicos: *buy* (2), *own* (3), *sell* (5).

Ya que tras la observación las colocaciones de este intervalo no se detectan tendencias positivas o negativas en las valencias, es pertinente aquí analizar cualitativamente una muestra aleatoria de oraciones del corpus. Las oraciones extraídas no reflejan ninguna relación con algún evento relacionado con la actualidad de la crisis, sino que son testigo de la rutina periodística empresarial. A simple vista, el sentimiento de las oraciones tomadas como muestra es, en su gran mayoría, positivo o neutro por lo que coincide con los datos obtenidos a través del procesamiento del lenguaje con *Lingmotif*. Algunas de las unidades léxicas positivas encontradas en el vecindario de las

palabras son las siguientes: “helping”, “good” o “benefited”. Una muestra de las oraciones observadas cualitativamente son las siguientes:

One example of cloud computing services mentioned in the court briefs are the music-storage lockers offered by COMPANIES like Google and Apple.

The COMPANY said Thursday that it had "not been advised of the intention of any federal agency to suspend or debar the COMPANY in connection with the plea agreement.

To a certain extent, Gap's other brands are helping the COMPANY make up for its flagship label's flailings.

Here are some recent "country of origin" entries from my credit card statements: First Great Western: Luxembourg EasyJet: Ireland The Trainline: Luxembourg These are, of course, good old British COMPANIES .

Food waste has become a hot-button issue, and restaurants, grocery chains, food processing COMPANIES , municipalities and the federal government are adopting strategies for reducing it.

Today, a half-dozen low-cost carriers control close to 40 percent of the intra-European air travel market, while in purely domestic markets like France, the British COMPANY easyJet is now a close second to Air France at many of the country's busiest airports, including Paris Orly.

Shares in the UK's largest hotel and restaurant group has dropped 1.5%, or 25p, to £16.95 despite the COMPANY's assertion that it had "delivered good total sales growth" of 11.0% for the 50 weeks to 16 February.

The COMPANY said it would inform shareholders on plans to use \$3bn of excess capital, previously earmarked for the RSA deal, when it publishes annual results.

How was the COMPANY doing?

Lakshmi Mittal, the company's chief executive, said in a conference call with reporters on Friday that the COMPANY had benefited from higher demand in Europe and North America, where it ships two-thirds of its steel.

Alibaba chose to debut on the NYSE after it could not get the Hong Kong exchange to agree a rule change that would allow the COMPANY to nominate the majority of its own board even after the share sale.

UK COMPANIES also pay out dividends to investors, something which pension funds and the like rely on

2009			2010			2011		
Palabra	Cooc	logDice	Palabra	Cooc	logDice	Palabra	Cooc	logDice
have-v	1065	10,146	say-v	843	10,069	have-v	849	9,562
say-v	799	10,058	have-v	1111	10,068	say-v	571	9,507
be-v	1853	9,880	be-v	1958	9,788	be-v	1477	9,146
other-j	204	9,330	oil-n	165	9,180	make-v	186	9,041
big-j	148	9,023	other-j	177	9,049	other-j	150	8,993
make-v	174	8,963	large-j	149	9,020	oil-n	116	8,988
executive-n	146	8,934	make-v	194	9,004	big-j	128	8,975
large-j	132	8,916	big-j	151	8,969	large-j	120	8,949
year-n	217	8,881	year-n	228	8,883	many-j	114	8,754
sell-v	134	8,856	executive-n	139	8,801	executive-n	108	8,677
own-v	110	8,762	sell-v	121	8,667	do-v	171	8,662
many-j	124	8,727	new-j	144	8,664	sell-v	97	8,629
oil-n	104	8,693	do-v	171	8,644	share-n	93	8,552
new-j	134	8,666	share-n	116	8,619	new-j	116	8,471
take-v	121	8,572	many-j	106	8,445	use-v	95	8,447
do-v	152	8,569	own-v	89	8,408	<b>help-v</b>	83	8,426
insurance-n	92	8,559	use-v	105	8,404	share-n	139	8,593
include-v	109	8,527	buy-v	96	8,373	use-v	148	8,588
more-j	126	8,521	market-n	112	8,373	business-n	166	8,485
share-n	101	8,511	more-j	116	8,350	world-n	127	8,472
2012			2013			2014		
Palabra	Cooc	logDice	Palabra	Cooc	logDice	Palabra	Cooc	logDice
have-v	850	9,660	have-v	1375	9,816	have-v	1484	9,831
say-v	570	9,625	say-v	880	9,761	say-v	963	9,803
be-v	1559	9,313	be-v	2506	9,445	be-v	2625	9,414
large-j	136	9,144	large-j	230	9,245	large-j	226	9,156
other-j	148	8,987	other-j	248	9,096	other-j	248	9,044
big-j	112	8,801	make-v	274	8,977	big-j	221	9,004
executive-n	111	8,754	oil-n	161	8,869	make-v	280	8,904
do-v	174	8,712	do-v	295	8,826	do-v	291	8,794
year-n	169	8,600	executive-n	170	8,747	more-j	214	8,632
new-j	118	8,566	year-n	271	8,694	year-n	285	8,608
make-v	129	8,556	big-j	161	8,665	share-n	160	8,606
small-j	83	8,519	sell-v	156	8,651	new-j	208	8,596
oil-n	78	8,501	more-j	193	8,569	take-v	190	8,595
many-j	93	8,483	help-v	152	8,566	many-j	198	8,595
take-v	103	8,449	work-v	159	8,555	include-v	206	8,564
work-v	124	8,495	take-v	169	8,518	executive-n	188	8,549
more-j	147	8,487	use-v	200	8,536	many-j	156	8,532
last-j	129	8,427	include-v	190	8,491	sell-v	179	8,513
use-v	119	8,408	new-j	215	8,489	energy-n	174	8,497
own-v	99	8,403	business-n	244	8,456	work-v	174	8,231

<i>2015</i>		
<i>Palabra</i>	<i>Cooc</i>	<i>logDice</i>
say-v	1263	9,918
have-v	1,74	9,846
be-v	2952	9,459
oil-n	285	9,251
big-j	301	9,226
large-j	262	9,163
make-v	338	9,043
other-j	281	9,037
share-n	247	8,930
executive-n	203	8,707
more-j	252	8,687
new-j	238	8,658
year-n	360	8,609
business-n	292	8,533
take-v	237	8,513
sell-v	202	8,509
include-v	215	8,481
small-j	177	8,363
buy-v	175	8,351
profit-n	177	8,333

**Tabla 59** Principales colocaciones de “company” durante el tercer intervalo de estudio





## 8. Resultados y discusión

« "When I use a word", Humpty Dumpty said, in rather a scornful tone, "it means just what I choose it to mean—neither more nor less".  
"The question is", said Alice, "whether you can make words mean so many different things".  
"The question is", said Humpty Dumpty, "which is to be master—that's all". »  
Lewis Carroll (1832-1898), *Through the Looking-Glass* (1872)

El presente capítulo pretende resumir los principales hallazgos observados en el Capítulo 7 y, además, reflexionará sobre las implicaciones teóricas de los mismos. Por último, plantea una pequeña discusión sobre las implicaciones deontológicas en la aplicación de algoritmos en los estudios humanísticos.

Gracias al desarrollo de *LexiEcon* y el posterior procesado de los textos con *Lingmotif* se ha podido realizar un seguimiento de la evolución de la orientación semántica de distintos términos clave en intervalos temporales cortos y observar la influencia de la crisis en las connotaciones del lenguaje especializado utilizado por la prensa en lengua inglesa. Tras el procesado y posterior análisis se observan principalmente dos hechos interesantes. El primero se refiere al enorme aumento de la frecuencia de uso de los términos en relación con los eventos clave de la crisis. Los términos se convierten en *palabras evento*. A partir de estos eventos, se producen cambios significativos en el uso de estos y el tipo de colocaciones varía, entrando en juego nuevas unidades léxicas con una orientación semántica diferente y que, en general, suelen ser menos especializadas que las fijaciones tradicionales relacionadas con términos. El segundo es que las medias anuales de la orientación semántica de un término contextualizado permiten observar con claridad fluctuaciones importantes en el sentimiento embebido en el discurso. Una triangulación de los datos cuantitativos con sus colocaciones más significativas y los eventos, ya históricos, relacionados con la *Gran Recesión* permite concluir que la orientación semántica de los términos del dominio económico-financiero es muy susceptible de variar a medida que se desarrollaron los eventos de la crisis financiera.

En el discurso, un evento (p. ej., “the credit crunch”) se define por una ruptura en el orden ordinario de las cosas que determina un antes y un después (Quenet, 2010).

Escapando de controversias sociológicas y lingüísticas sobre la definición del evento, los datos cuantitativos y cualitativos estudiados a lo largo del Capítulo 7 ponen sobre la mesa la magnitud y relevancia la de la crisis y sus efectos sobre el discurso. De acuerdo con López Muñoz y Verine (2014), lo que define un evento es su tratamiento como tal por el discurso de la prensa a la que la opinión pública confía unos criterios mínimos de objetividad, autenticación y de los hechos relatados. Es decir, se espera que la prensa cumpla con la responsabilidad de detectar e informar sobre los eventos (por ejemplo, informar de manera fidedigna sobre los malos resultados de un banco).

En el caso del periodismo económico, este es un discurso mediado, en el que los periodistas interactúan con diversas categorías de expertos (políticos, economistas, usuarios, etc.). La cadena de transmisión experto-periodista-público pone sobre la mesa la relevancia y el uso de las unidades léxicas especializadas, convirtiendo a los términos en eventos. Desde una perspectiva semántica, es esta circulación la que hace que un hecho mundano se transmute en un evento, asignándole un significado social memorable en torno a unos términos, con unas connotaciones determinadas a lo largo de un intervalo de tiempo.

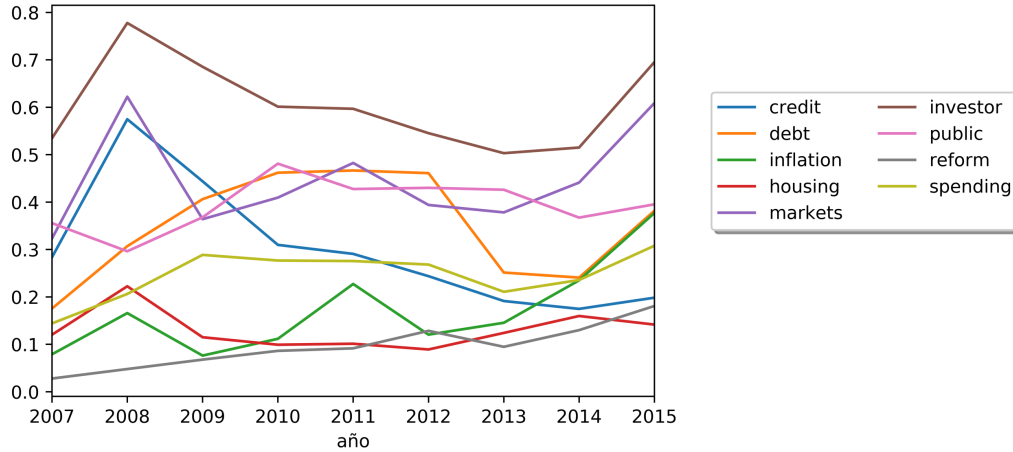
La tarea de análisis realizada trata de identificar brechas en los usos lingüísticos que constituyen los eventos y, como se ha visto con profundidad en el Capítulo 7, establecer las causas. La observación cuantitativa de la frecuencia relativa del uso de los distintos términos (Figura 28) permite observar que en la mayoría de los casos el estallido de la crisis en el año 2008 fue muy relevante, generándose un pico en el uso de los términos en cuestión. En muchos casos, además, esto ha ido acompañado de un cambio radical en sus connotaciones de orientación semántica. En efecto, como se puede observar en la Figura 29 en la mayoría de los casos se pasa de una media positiva de sentimiento previa a la crisis (año 2007) a una caída de las medias al umbral de la negatividad. La separación de la crisis del crédito con la crisis de la deuda es también evidente tanto en las frecuencias como en el sentimiento y se observan patrones similares en la mayoría de los casos. De hecho, en el año 2009 se produce un aumento puntual en la media de los datos de sentimiento para, posteriormente, volver a caer en el año 2010 al aumentar la deuda y, consecuentemente, el deterioro en las condiciones de vida de la ciudadanía. Posiblemente a partir de una mejora leve en las condiciones macroeconómicas de los países del norte de la Zona Euro y en gran medida con el espíritu y giro discursivo que supuso el lanzamiento de políticas que supusieron un alivio económico relativo, en especial la *expansión cuantitativa* por el Banco Central Europeo iconizada en el famoso

discurso “Whatever it takes” de Mario Draghi, se observó un aumento significativo en las puntuaciones de sentimiento, que en la mayoría de los casos se prolongó desde 2011 o 2012 hasta el final del estudio.

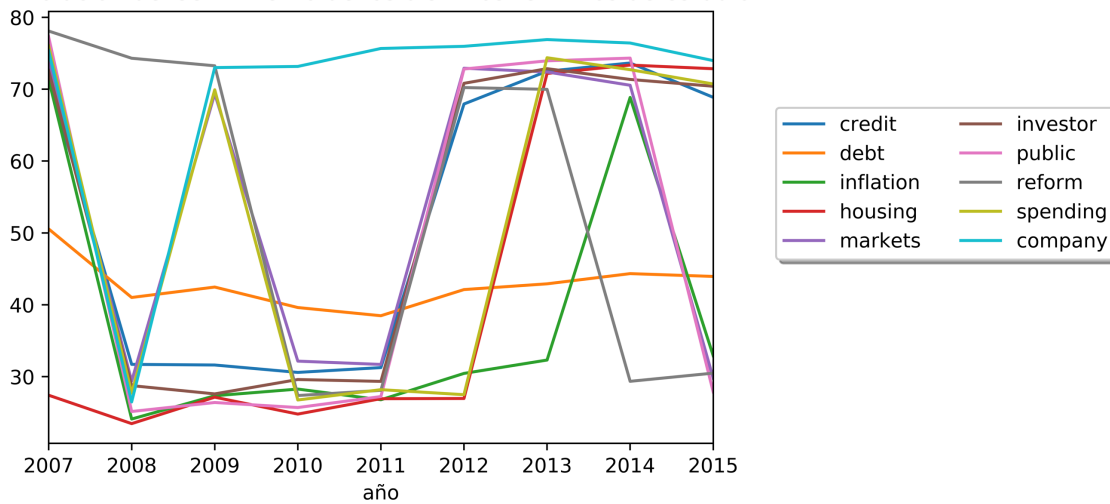
A partir del estudio basado en corpus se constata que la prosodia semántica es una propiedad del léxico que no es estática, sino variable en el tiempo. De acuerdo con Bublitz (1996) o Morley et al. (2009), se observa que un conjunto de colocaciones evoluciona para adquirir un matiz de orientación semántica positiva o negativa lo largo de un intervalo temporal. Como se ha visto a lo largo del análisis llevado a cabo en el Capítulo 7, los cambios cuantitativos en el sentimiento surgidos a partir de la crisis han ido de la mano de la aparición de colocaciones con sentimiento positivo o negativo según la fluctuación de las medias de sentimiento de los contextos de los términos y coincidiendo con los eventos económicos. De esta forma, a partir de 2008, la significancia (en cuanto al coeficiente *logDice*) de colocaciones neutras y asentadas, como la de “credit” con “card” fue superada por colocaciones relacionadas con la crisis (p. ej., “crunch” o “default”). Esto evidencia el uso de los términos en contextos más negativos. Se produce a todos los efectos, utilizando las palabras de Steward (2010), un “contagio” inmediato del significado pragmático o actitudinal del término a causa de un evento repentino y violento como es la crisis, que permaneció constante mientras duraban sus efectos. Por ello, la mera mención de “credit” durante este intervalo implicaba en la mayoría de los casos connotaciones negativas, de forma que su uso mantenía la armonía semántica de las enunciaciones.

Como contraste y evidencia que confirma esta hipótesis está el uso de “company”, un término que no formó parte de esta transformación. En efecto, el término fue considerado *palabra evento* dada la escasa variación de su frecuencia relativa. Con la excepción del colapso de la media de sentimiento en 2008 (hecho que no hace sino constatar la magnitud discursiva del evento), tanto la frecuencia relativa como la media de sentimiento de las oraciones que incluían el término permanecieron estables. De hecho, entre sus colocaciones no destacan fijaciones de ningún tipo y, al contrario que en el resto de los términos de estudio, prevalecen los verbos auxiliares y sustantivos que no indican ninguna eventualidad. En efecto, estas colocaciones permanecen prácticamente inalteradas durante todo el intervalo de estudio.

Evolución de la frecuencia relativa de los distintos términos de estudio

**Figura 28** Evolución de la frecuencia relativa de los distintos términos de estudio<sup>59</sup>

Evolución del sentimiento de los distintos términos de estudio

**Figura 29** Evolución del sentimiento de los distintos términos de estudio

El análisis detallado de grandes conjuntos de datos de manera automatizada, de forma cuantitativa y cualitativa, permite obtener una mayor perspectiva del uso del lenguaje durante un evento de calado mundial como fue la crisis global. A nivel teórico, se pueden extraer tres conclusiones principales:

1. Los resultados del análisis de rendimiento de *LexiEcon* y el posterior análisis de los textos evidencian las enormes posibilidades que ofrece el estudio de los lenguajes especializados desde una perspectiva escasamente explorada como es la expresión del sentimiento.

<sup>59</sup> Se omite la frecuencia relativa de *company* al ser esta mucho mayor y estable.

2. Existen términos del dominio económico-financiero que adquieren gran relevancia durante eventos de gran calado social, como es el caso de la Gran Recesión.
3. En relación con estos términos, se pudo observar que este sentimiento inherente no es en ninguno de los casos un valor estático, sino que en mayor o menor medida este varía con las condiciones socioeconómicas.

La observación de las puntuaciones de sentimiento, los rankings de colocaciones y su evolución a lo largo del intervalo de estudio no es sino una puesta en escena de los planteamientos de la lingüística sistémica donde la lengua es un sistema de elecciones por el cual el hablante elige de una forma explícita qué expresión (palabras, términos, etc.) utilizará en detrimento de la totalidad de sus alternativas. En sintonía con Halliday et al. (2014), el buen rendimiento de *LexiEcon* en el análisis de los textos muestra que el lenguaje especializado no es una excepción y está organizado de una forma no arbitraria y enteramente funcional, en consonancia con su uso. Desde este punto de vista, el lenguaje es un recurso semiótico social que permite a los hablantes opciones para construir significados, destacando aquí la expresión de la evaluación que no escapa de la gramática.

A través de la observación del uso de una serie de unidades terminológicas en un contexto divulgativo del lenguaje económico-financiero, se ha podido evidenciar que el uso evaluativo de los términos está presente en todos los niveles lingüísticos. De hecho, el uso evaluativo no es ninguna excepción para el lenguaje especializado y su puesta en escena cumple en su totalidad con la definición de Alba-Juez y Thompson (2014) vista en el Capítulo 3, ya que desde este dominio es posible la expresión estructurada de los sentimientos sobre las entidades de las que habla, con vistas a influir a una audiencia de forma que se refleja el conjunto de valores de la comunidad de hablantes.

En cuanto a la teoría de los lenguajes de especialidad y la terminología, en consonancia con lo propuesto por Weissenhofer (1995) o Ten Hacken (2015), el uso cotidiano de la terminología no es del todo *unívoco*. Como se ha podido observar, la puesta en uso de los términos corresponde a dos factores principales en el ámbito del dominio. Por un lado, la estricta y necesaria expresión científico-técnica en el dominio económico. Por otro lado, especialmente aquí, los usos prototípicos que emergen del discurso cotidiano del periodista especializado como mediador entre el experto y el público. Por poner un ejemplo relacionado con el uso, este sería el caso de la necesidad

de especificar con un adjetivo en “bad debt” de forma que quedase clara la naturaleza extraordinaria y “tóxica” de la deuda. Es necesario resaltar que uso divulgativo del lenguaje por parte de la prensa exige niveles menores de especialización que la precisión requerida por los textos dirigidos a expertos (p. ej., en divulgaciones corporativas). En los textos analizados, el aspecto evaluativo juega un papel fundamental. A través de la opinión y la discusión, se pueden expresar percepciones subjetivas sobre el estado de la actualidad económica y aún así cumplir con las condiciones suficientes y necesarias para describir la realidad de la crisis sin perder rigor, y teniendo en cuenta la dimensión dialógica que señale las fricciones entre distintos marcos teóricos e ideológicos.

La observación de las colocaciones, cotexto y contexto de los términos evidencia fluctuaciones en el sentimiento de manera casi inmediata como reacción a los eventos económicos. A partir de una perspectiva computacional con base lingüística, la observación del sentimiento de las unidades léxicas especializadas permite también avanzar desde una terminología conceptual eminentemente estructural a la consideración de estas como unidades también comunicativas y emocionales. De acuerdo con las nociones contemporáneas, se propone realizar algunas aportaciones a los planteamientos clásicos de la terminología, de forma que se supere la visión aséptica y estática del término para desplazarse a un conjunto teórico que refleje una naturaleza dinámica y poliédrica de la terminología que incluya significados que vayan más allá de la linealidad concepto/definición. Los resultados coinciden con las aportaciones críticas con la consideración de los términos como compartimentos estancos de carácter sincrónico (p. ej., Cabré, 2000 o Temmerman, 2000) que no tienen en cuenta ningún tipo de evolución del término en el transcurso del tiempo.

Asimismo, como se ha podido observar, se plantea un uso de los términos con un pie dentro y otro fuera del registro formal profesional en el que quepan más significados, eminentemente comunicativos, que la mera definición del concepto como la orientación semántica. De esta manera, el uso *vivo* de los términos evidencia también elementos semánticos cognitivos y pragmáticos. Esto es, tras el procesado y posterior análisis cualitativo se puede afirmar que, además de transmitir conocimientos, los usos cotidianos de los términos en el discurso especializado permiten un uso evaluativo del lenguaje y que, al menos en el dominio económico-financiero, este uso evaluativo del lenguaje especializado evoluciona con los cambios políticos y sociales.

Por otro lado, es necesario discutir sobre la cuestión deontológica. El problema de la relación entre la clasificación de la información y la ideología no es novedoso (p. ej.,



Quine, 1951). Sin embargo, los desarrollos actuales permiten realizar nuevas discusiones sobre el papel de estos en el mundo. Al pasar del paradigma preciso, *aséptico* y libre de variación a una perspectiva que acepta el carácter funcional (especialmente la subjetividad) y las variaciones en el uso de la terminología, es necesario tener en cuenta el papel que juega el paradigma ideológico del terminógrafo en la selección de los términos que contienen sentimiento y la necesidad de establecer criterios para asignar valores de sentimiento a las unidades terminográficas. De hecho, en gran parte de la literatura los productos terminográficos (p. ej., ontologías o tesauros) se presentan como entes estructurados, estandarizados y lejanos al discurso. La informatización de los textos ha impactado de pleno en el trabajo de indexación y a finales de la década de 2010 está jugando un papel polémico y fundamental en el desarrollo de la vida cotidiana (p. ej., los *rankings* de búsqueda en *Google*). Por ello, se da por un lado la necesidad de estandarizar, para promover intercambios dentro de una lengua y entre lenguas; por otro lado, sin embargo, normalizar significa imponer una visión del mundo que responde necesariamente a una construcción ideológica del imaginario colectivo.

Sin que sea un objetivo de trabajo de esta tesis, es necesario evidenciar que la discusión ideológica a la hora de clasificar y etiquetar la información ha sido fundamental a la hora de desarrollar *LexiEcon* para *Lingmotif*. En muchos casos la labor de investigación de ejemplos en corpus textuales y la observación de resultados ha sido fundamental a la hora de poder determinar el sentimiento expresado en los textos. En la *guerra de la información*, gran parte de las bases técnicas fundamentales toman como punto de partida el estudio de los niveles de clasificación revisados en el Capítulo 4.<sup>60</sup> De acuerdo con Aussenac-Giles y Condamines (2005), previo a este desarrollo teórico, la disciplina de trabajo terminográfico se ha considerado tradicionalmente como un apoyo al acceso a los datos ya existentes. Se puede concluir que las distintas perspectivas o intereses suponen cambios significativos en la manera de tratar y clasificar el conocimiento, al interponer factores antes secundarios (por desgracia, principalmente el lucro) en un área previamente considerada como un servicio y/o una labor eminentemente académica.

---

<sup>60</sup> En el momento de escribir este texto se produce el escándalo por la extracción y clasificación de datos personales de los usuarios de redes sociales por *Cambridge Analytica* utilizados para fines políticos. En paralelo, para afrontar este desafío, se implanta el *Reglamento General de Protección de Datos* de la UE.

Todos los algoritmos esconden una opinión y, camuflado entre las “frías matemáticas”, se encuentra un trasfondo ideológico. De una manera casi apocalíptica, O’Neil (2017) pone sobre la mesa la cuestión humana y epistemológica en la era de los algoritmos: “Los privilegiados son analizados por personas; las masas, por máquinas”. Sobre esto la autora comenta que el papel de estos algoritmos, muchas veces de manera bienintencionada, en la crisis (financiera, de vivienda, de empleo, etc.) dejó claro que las fórmulas matemáticas no solo están involucradas en gran parte de los problemas del mundo, sino que en muchas ocasiones los agravan. Si bien la velocidad de procesamiento de los datos da luz y evidencias basadas en datos sobre muchos de los problemas humanos otrora inaccesibles, el análisis está incompleto sin un análisis cualitativo, casi artesanal, que triangule y ponga en cuestión el calado de dichos problemas y sus implicaciones éticas o sociales. El mundo que nos rodea es ya dependiente de potentes algoritmos como el propuesto por *Lingmotif*. Gracias al desarrollo de herramientas de análisis de sentimiento, la tarea de determinar la frecuencia y la polaridad agregada de los términos ha sido sorprendentemente rápida y esclarecedora. La resolución de este problema hace tres décadas hubiese planteado un reto metodológico y técnico inalcanzable. Sin embargo, en 2018, en pocos minutos se procesan decenas de miles de oraciones y se obtienen gráficas de la evolución del sentimiento del contexto de los términos de estudio, junto con una decena de datos estadísticos adicionales.

Este trabajo quedaría del todo incompleto sin un análisis pormenorizado del funcionamiento y de las implicaciones lingüísticas, económicas y sociales que hay de fondo y que se han discutido en estos capítulos. El estudio de los ejemplos en el corpus ha sido esclarecedor y fundamental para obtener conclusiones y comprender el trasfondo de los cambios en la evolución semántica. Por eso, la labor de las humanidades digitales debe en todo caso enfocarse a observar los problemas y poner rostro humano a la gran cantidad de nuevas preguntas que nos facilita el mundo nuevo que se abre con el procesamiento automatizado de la información.

En resumen, el modelo llevado a cabo a lo largo del Capítulo 7 parece adecuado para cumplir con el objetivo general de esta tesis: la detección tendencias y fluctuaciones en el sentimiento de los términos, para así establecer vínculos con factores externos que permitan tanto detectar cambios en los patrones lingüísticos y sociológicos. Los datos cuantitativos y cualitativos analizan una gran cantidad de información textual y aportan solidez a la argumentación de que los eventos de gran calado social, como es el caso de la Crisis Global fueron los desencadenantes de los cambios en la evolución de la

---

orientación semántica de los términos. Las claras variaciones de su uso y la aparición de colocaciones nuevas a lo largo de un intervalo de tiempo marcado por las adversidades económicas muestran con claridad la presencia de una dimensión evaluativa dentro del uso comunicativo de los términos económicos y financieros.



## 9. Conclusiones finales

« A story has no beginning or end: arbitrarily one chooses that moment of experience from which to look back or from which to look ahead. »

Graham Greene (1904-1991), *The End of the Affair* (1951)

### 9.1 Consecución de objetivos

Esta sección trata sobre la consecución de los propuestos y dar respuesta de manera concisa a las dos preguntas formales realizadas en la introducción de la tesis. Esta labor se realiza como recopilación de las discusiones, reflexiones y hallazgos adelantadas a lo largo de las conclusiones de cada uno de los capítulos de esta tesis. De esta forma, se pretende dar cierre al trabajo realizado y abrir la puerta a las nuevas preguntas que emergen a partir de aquí.

El presente trabajo es de carácter altamente multidisciplinar, pues combina una discusión el rol eminentemente lingüístico del estudio de las lenguas de especialidad y el cambio en la orientación semántica con los factores socioeconómicos de la mayor crisis económica de las últimas ocho décadas, el dinamismo de la actualidad informativa y el rigor de los últimos desarrollos técnicos en el análisis de sentimiento, una de las disciplinas más punteras que abarca el procesamiento del lenguaje natural. A su vez, el desarrollo del análisis de sentimiento como disciplina armoniza técnicas terminográficas fundamentadas en la teoría en conjunción con los últimos desarrollos en algoritmos matemáticos y ciencia de datos.

Gracias a esta perspectiva, observar los cambios en el lenguaje se convierte en una tarea fascinante, ya sea a partir de enfoques estructurales más clásicos como las variaciones en la morfología, la sintaxis o más vinculadas al estudio del significado, principalmente a través de la ventana que abre el estudio de la prosodia semántica. Todo cambio en los distintos aspectos de la naturaleza humana responde a una motivación histórica construida de manera colectiva que pone de manifiesto la relación hiperdinámica latente entre el uso comunicativo del lenguaje y los dispositivos culturales que nos rodean. El acto discursivo revela la naturaleza cognitiva y la posición del hablante con el mundo,

mostrando especialmente la percepción subjetiva del estado de las cosas. En pocas palabras, la observación de las relaciones contextuales del lenguaje ordinario a través de una serie temporal permite conocer en profundidad las relaciones estructurales entre el discurso y el espíritu de los tiempos y observar con sencillez la fluidez del sentimiento que emana de ellas. El desarrollo metodológico longitudinal cumple con su doble propósito de describir patrones de cambio y explicarlos a través de su relación con los eventos sociopolíticos relacionados con la crisis.

Los objetivos marcados para la presente tesis se cumplen ampliamente. Con respecto al objetivo general, “Analizar las fluctuaciones en la orientación semántica de determinados términos económicos dentro del período 2007-2015 a través de la caracterización del impacto de los eventos de mayor orden en las variaciones semánticas de las unidades léxicas”, este se pudo llevar a cabo gracias a la velocidad y eficacia que actualmente ofrece *Lingmotif*. A lo largo del Capítulo 7 fue posible analizar los datos cuantitativos relacionados con la evolución de la orientación semántica y la frecuencia relativa de diez términos financieros dentro del período 2007-2015. Gracias a este estudio cuantitativo se observa que el sentimiento de los términos no es estático a lo largo de intervalos temporales, sino que se producen variaciones (*peyoraciones* y *mejoras*) rápidas y pronunciadas de la orientación semántica que emergen en el uso de la comunidad de hablantes como reacción inmediata ante cambios drásticos en la situación socioeconómica.

El análisis de sentimiento o análisis de opinión depende profundamente del contexto y, además, del dominio específico en cuestión. Al abrir la puerta comunicativa del uso del lenguaje especializado se puede observar el entramado prosódico inscrito en el uso y este se convierte en un instrumento excelente para observar la proyección del sentimiento a lo largo de intervalos cortos, de forma que se puede obtener una *foto fija* del uso contemporáneo del lenguaje. Este análisis refleja que el significado connotativo de las palabras está vivo e íntimamente relacionado con el estado mental y la opinión de la comunidad de hablantes. Dada la necesidad divulgativa de la población, el periodista económico debe realizar una labor mediadora entre el experto y el público. Los resultados hacen explícitos que, además de la transmisión técnica de la información, la opinión va en todo caso inscrita en la *traducción* de las abstractas cifras macroeconómicas a enunciaciones lingüísticas.

Además de abrir nuevos horizontes de discusión en el dominio lingüístico, este ofrece aplicaciones prácticas en economía. En resumidas cuentas, el estudio de la

evolución del sentimiento, además, sirve como termómetro de la opinión de los expertos en torno a los indicadores económicos. Como se plantea en la introducción de este trabajo, el análisis de sentimiento ofrece otra perspectiva de observación de la política económica que, si bien aún no se alcanza a dominar, se comienza a vislumbrar su potencial y abre la puerta a nuevos retos en la investigación multidisciplinar.

El primer objetivo específico es “Determinar un lexicón de sentimiento de dominio específico en lengua inglesa para el dominio de las noticias económicas de alta precisión y exhaustividad a partir de un corpus de texto especializado”. Este resultado puede ser considerado como una aportación a la comunidad científica interesada tanto en el análisis del discurso económico como en la metodología de análisis de discurso especializado. De hecho, el desarrollo de *LexiEcon* aporta uno de los lexicones de dominio económico-financiero en lengua inglesa más completos de los disponibles en la actualidad. Como resultado se presenta *LexiEcon*, un lexicón *plug-in* de dominio específico para la lengua inglesa que en la actualidad incluye 6.843 entradas (palabras y expresiones multipalabra). Como se revisó con profundidad en el Capítulo 6, los resultados de su evaluación a través de un *gold standard* devolvieron unas medias excelentes de precisión y exhaustividad. Además, dadas las especificaciones técnicas y la filosofía *open data*, *LexiEcon* es un producto que ofrece gran sencillez de ampliación y adaptación según las necesidades específicas de otros sistemas y modelos de investigación.

Para el proceso de compilación del *plug-in*, se construyó un corpus de textos de prensa económica-financiera en inglés estructurado mensualmente para el período 2007-2015, a disposición bajo demanda, que contribuye a la valoración académica del estudio realizado. El apartado 6.3 trata sobre el diseño y recopilación del *Esmeraldas Corpus of Business English* (ECBE), la fuente de datos principal de este trabajo que incluye más de 24 millones de *tokens*. El corpus está recopilado según los principios de equilibrio y representatividad tradicionales de la disciplina de lingüística de corpus: electrónico, con textos auténticos, equilibrado y representativo de un idioma o variedad del lenguaje. La colección clasifica de manera cronológica todos los textos en línea de la sección *Business* de las cabeceras *The Guardian* y *New York Times* a lo largo de los años en los que tuvo lugar la crisis global hasta el año de inicio de este estudio (2007 y 2015). Para la extracción de los textos se utilizaron *arañas web* y los textos se clasificaron posteriormente de manera semiautomática. El producto final fue *tokenizado*, *lematizado* y empaquetado en *The Sketch Engine* y está disponible bajo demanda. Tras su utilización



se reafirma el planteamiento de que la prensa generalista online es, sin duda, uno de los mejores recursos posibles. Esta refleja fielmente el discurso de una época y se convierte en un instrumento fácil de utilizar para obtener información veraz de la lengua en uso. A diferencia de los informes de empresas o la complejidad del estudio del sentimiento en la lengua oral o en redes sociales (cargada de p.ej., ironías y usos coloquiales), el estudio de la prensa aporta una interpretación de la situación político-económica con un grado de imparcialidad, pero aún así cargada de opinión, que sirve de puente entre la opinión del *parqué* financiero y de la opinión en los debates a pie de calle.

Asimismo, el investigador en búsqueda de metodología de análisis de discurso especializado podría encontrar interesante la adaptación a su dominio de especialidad del conjunto de metodologías y procedimientos que se usaron para extraer la información lingüística y terminológica con el afán de analizar el sentimiento y su representación en el léxico de dominio específico a partir de un corpus especializado. Como se detalla en el apartado 6.5, se combinaron técnicas cuantitativas y cualitativas que permitiesen la extracción de terminología. Una amplia lista de 30.000 candidatos a término (15.000 palabras y expresiones multipalabra) se extrajo del ECBE gracias al extractor terminológico *KeyWords*. Con la ayuda de un experto en el dominio, cada uno de los candidatos fue comparado con los valores de las entradas del dominio general recogidas por *Lingmotif-Lex*, seleccionándose solo aquellas entradas cuyo sentimiento en el dominio específico difirieran y, posteriormente, a cada una de ellas se le asignó la valencia correspondiente al dominio específico. A partir del corpus, se recopiló una serie elementos direccionales (p. ej., “rise”, “fall”, “decline”, etc.) en conjunción con una lista de indicadores económicos potencialmente positivos o negativos (p. ej., “debt”, “spending”) a los que se asignó el sentimiento correspondiente según el formato establecido por *Lingmotif*. De esta manera la entrada <rise>\_3\_debt permite cubrir casos que utilizan el mismo patrón y comparten sentimiento como es el caso de “rise in bad debts”, “rise in the national debt” o “rising level of debt”.

El segundo objetivo específico es “Definir un subconjunto de datos longitudinal en forma de oraciones que contienen los términos de estudio y que serán el “input” del análisis de sentimiento”. Este objetivo se cumplió mediante la utilidad de extracción de oraciones ofrecido por el buscador de concordancias de *The Sketch Engine*, que para cada uno de los términos clave recopiló la totalidad de las oraciones etiquetadas por fecha en un fichero de texto plano. El conjunto de datos cumple con los requisitos técnicos de *Lingmotif* y está recopilado bajo estándares abiertos. Al igual que el corpus, estos

conjuntos son fácilmente modificables y reutilizables, facilitan el estudio automatizado del dominio económico-financiero y están disponibles bajo demanda.

El tercer objetivo específico es “Para los términos económicos-financieros objetos de estudio, identificar los factores (sociales, políticos, económicos, etc.) que han acompañado a cambios en el uso, desde la perspectiva de sentimiento, durante el período”. Además de las particularidades de cada término de estudio recogidas extensamente en el Capítulo 7, en el Capítulo 8 se discutió que la frecuencia relativa del uso de los distintos términos permite concluir que en la mayoría de los casos el estallido de la crisis en el año 2008 hizo que estos términos ganasen en popularidad al relacionarse con los eventos (p. ej., se destaca el colapso de Lehman Brothers). Este fenómeno, además, ha ido acompañado de un descenso del umbral de negatividad de las medias de sentimiento para las oraciones que incluyen dichos términos. De hecho, en casi todos los casos se parte de una media positiva de sentimiento en 2007 que es seguida en el año 2008 de una caída brusca de las medias al umbral de la negatividad.

Otro factor general es la separación de la época influida por el “credit crunch” (2007-2008), de carácter más global, con la crisis de la deuda en la Zona Euro (a partir de 2010) que, si bien influyó en el exterior, fue un fenómeno eminentemente europeo. Para todos los términos, las cifras de frecuencia y las medias de sentimiento ofrecen patrones similares. De hecho, en el año 2009 se produce un aumento puntual en la media de los datos de sentimiento para posteriormente volver a caer en el año 2010 al aumentar la deuda y, como consecuencia de las políticas de austeridad que pretendían paliar sus efectos, el deterioro en las condiciones de vida de la ciudadanía. En el bienio 2010-2011 un aumento de la media de puntuaciones de sentimiento hacia el umbral de la positividad en la mayoría de los casos, producida por una mejoría relativa en las condiciones macroeconómicas principalmente en los países del norte de la Zona Euro. Es necesario destacar el espíritu de alivio infligido por las políticas de expansión cuantitativa llevadas a cabo por el Banco Central Europeo y enfocadas eminentemente a los capitales. Se recoge en la discusión que la potencia perlocutiva del espíritu del discurso “Whatever it takes” de Mario Draghi en la rueda de prensa del Banco Central Europeo, que según el consenso de los expertos “salvó el euro”, quedó reflejada también en la orientación semántica de los términos estudiados. El espíritu de las políticas de alivio supuso en muchos casos un aumento significativo en las puntuaciones de sentimiento, de forma que se prolongó hasta el final del estudio.

El cuarto y último objetivo específico es “Analizar cualitativamente las posibles variaciones en la prosodia semántica de los términos de estudio a través de los años”. Acompañando al estudio cuantitativo del sentimiento, el diseño de la investigación propone un análisis cualitativo del uso del lenguaje de forma que sirva como triangulación de los resultados y ofrecer mayor robustez a las conclusiones del estudio. Las lenguas de especialidad tienen una terminología, expresiones y construcciones propias que distan de ser inalterables, y que además reflejan el espíritu de los tiempos. En la gran mayoría de los casos los términos han pasado a tener colocaciones anuales de carácter más negativo o positivo. Todos estos factores salen a la luz en la manera en la que se expresan las evaluaciones y opiniones en los textos. Por ello, se está de acuerdo con las enunciaciones revisadas a lo largo del trabajo en la línea de distintos autores, p. ej., Moesner (2003) o Morley et al. (2009), por las que el contenido semántico de las palabras se modifica a medida que se asocian con conceptos más o menos favorables de lo que solían hacerlo. Como se observa a través de los *rankings* anuales de colocaciones y la posterior observación de los ejemplos, la orientación semántica del contexto de los términos evoluciona al intensificar su positividad o negatividad, y se da el caso invertir su polaridad por completo a lo largo de un periodo determinado. Además, se da el caso de que mientras estas unidades se han considerado *palabra evento*, suelen relacionarse con unidades léxicas de carácter no especializado. Por poner un ejemplo, en 2007, durante el estallido de la burbuja inmobiliaria y previo al desplome de las hipotecas *subprime*, el término “housing” se relacionaba con unidades de carácter negativo, entre otras, “downturn”, “slump”, “weaken”, “bubble”, “slowdown”, “slow” o “trouble”. Con una recuperación del mercado inmobiliario en el mundo anglosajón en el año 2013, la orientación semántica de las colocaciones es totalmente opuesta: “affordable”, “boom”, “recovery”, “allowance”.

En otros casos, el uso de las colocaciones no fue tan evidente. Por ello, se realizó una revisión completa a través de los ejemplos obtenidos del corpus, lo que en los casos menos evidentes fue esclarecedor, ya que prevalecían colocaciones de carácter neutro. Al observar de manera cualitativa el sentimiento de los textos a nivel oración, y posteriormente en conjunto, se evidenció la prevalencia de una orientación semántica determinada, coincidente con los valores de sentimiento medios.

Además del planteamiento de los objetivos, se realizaron dos preguntas que constituyeron el punto de inicio e hilo conductor de la tesis. De manera sintética, se dará una respuesta a cada una de ellas de acuerdo con lo expuesto a lo largo de los ocho

capítulos anteriores. Para la facilitar el trabajo de revisión, a continuación, las preguntas se reproducirán de manera textual.

Ambas preguntas pueden responderse en conjunto: “¿Varían las unidades léxicas a medida que se asocian con conceptos más o menos favorables que antes?”, “¿Impulsan los eventos históricos de mayor orden, como es el caso de las crisis cambios rápidos en la orientación semántica de los términos?” Como se ha podido observar a lo largo del Capítulo 7, existe una relación entre el aumento de la frecuencia relativa y un cambio en la media de la orientación semántica. Como afirma Bybee (2015), una vez que una palabra entra de lleno en el habla diaria de sus hablantes, con mayor o menor magnitud esta sufrirá cambios de naturaleza semántica. Existe una tensión entre la necesidad de estabilidad en el significado de forma que la comunidad lingüística use con normalidad estos términos y la necesidad de adaptar los términos a nuevos usos. Aunque la correlación no implica causa, se ha observado aquí que los cambios en el sentimiento de las palabras suelen ir acompañados de un aumento súbito en su frecuencia de uso.

Existen *palabras evento*, de acuerdo con Moirand (2007), que parten de un carácter inicialmente vinculado a las lenguas de especialidad, es decir, tienen inicialmente un punto de vista superficialmente aséptico y técnico. A partir de un hecho destacable, p. ej., un giro político, el uso de estos términos comienza a evolucionar al convertirse en medios de transmisión de los hechos comunicativos especializados para el gran público. Finalmente terminan adquiriendo características léxicas que son permanentes en mayor o menor grado a partir de los eventos ya que sus repercusiones permanecen en la memoria colectiva.

Casi en la totalidad de los casos, una media negativa de puntuaciones de sentimiento ha acompañado a situaciones económicas desfavorables relacionadas con los factores que implican el uso del término y, a su vez, la aparición de colocaciones más negativas, y a la inversa. Las connotaciones de determinadas unidades léxicas evolucionan a partir de contextos con valores de sentimiento distintos que son observables a través de sus distintas relaciones léxicas: reformulaciones discursivas, colocaciones, coligaciones, patrones, etc. Como consecuencia, las unidades léxicas tienden a adquirir connotaciones más positivas o negativas debido a que su prosodia semántica varía a lo largo de un intervalo temporal determinado. Se puede concluir que a través del “aura” del término, es decir la prosodia semántica, en términos de Louw (1993), el significado connotativo del término fluctúa en relación con el transcurso de los eventos y sus factores socioeconómicos. Para ilustrar esto, en el caso de “public”, en plena crisis del crédito se

observa que la mayoría de las oraciones engloban un sentimiento negativo. Además, las colocaciones que se encuentran en el vecindario de los distintos conjuntos también suelen serlo. Por ejemplo, el término coocurre puntualmente con el sustantivo “cuts”, y a su vez aparecen frecuentemente colocaciones que mantienen la armonía del sentimiento, por ejemplo: “rig”, “conspiring” o “miserable”.

Como se ha visto a lo largo del Capítulo 3, el lenguaje evaluativo permea a través del texto (palabra, sintagma, oración, texto y contexto), es acumulativo y se manifiesta de formas muy diversas. El estudio de patrones gramaticales mediante corpus (p. ej., Hunston et al., 2000) facilita la detección una serie de estructuras relacionales (p. ej., colocación, coligación o prosodia semántica) susceptibles de incluir segmentos evaluativos. Sin embargo, es necesario poner sobre la mesa la dificultad que en ocasiones supone la diferenciación entre enunciaciones evaluativas y factuales. En el caso del dominio económico, especialmente, existen enunciaciones que son aparentemente factuales y que sin embargo contienen sentimiento. Por ejemplo, este es el caso de “Analysts predict wages will keep rising” que, sin embargo, contienen un sentimiento positivo desde un punto de vista social.

A partir de los resultados obtenidos en la evaluación de *LexiEcon* se puede observar que el sentimiento es altamente dependiente del dominio. Los sistemas de análisis de sentimiento suelen obtener una precisión mucho menor debido a que los patrones evaluativos tienen particularidades propias de cada dominio. Tras observar los resultados del análisis de las oraciones de dominio económico-financiero del *gold standard*, se evidencian estas profundas diferencias pues utilizándose en conjunto con el lexicón de *Lingmotif* obtiene un *Valor-F* medio de 0,735, en torno a un 20% que los resultados que ofrece *Lingmotif* sin *plug-in* cuando clasifica los textos del dominio económico-financiero.

Dependiendo del dominio, las unidades léxicas que a través de la polisemia están presentes en distintos dominios pueden tener un sentimiento distinto dependiendo del contexto. Este es el motivo principal por el que es necesario adaptar los sistemas de análisis de sentimiento según el dominio. Por ejemplo, mientras que “launder” tendría un sentimiento neutro en el dominio general, la observación de su uso en el dominio económico haría referencia al blanqueo de capitales, el cual es una actividad económica delictiva. Este es uno de los motivos, por lo que un sistema de análisis de sentimiento basado en léxico necesita un amplio trabajo lexicográfico relativo al dominio objetivo. En el caso de este dominio, es esencial extraer una amplia lista de terminología, modismos

y jerga que incluya los tres elementos principales: indicadores de rendimiento, direccionalidad de los indicadores de rendimiento y otros términos propios del dominio.

Por otro lado, la naturaleza especializada de los textos facilita el análisis de sentimiento basado en léxico y ofrece mejores resultados que en el dominio general. A pesar de que el uso de los términos no es íntegramente unívoco, los buenos resultados se justifican por la naturaleza del lenguaje especializado ya que tiende a un bajo nivel de ambigüedad. Si bien en el discurso de la prensa semiespecializada se da una falta de precisión estricta, estos niveles de precisión en el uso del lenguaje son mucho mayores en comparación con el uso del lenguaje en el dominio general (p. ej., el uso de ironías o anáforas). Además, la redacción de los artículos económicos favorece la claridad de las enunciaciones con el fin de allanar al público la comprensión de una información con un alto nivel de abstracción.

Como cierre, este trabajo ha descrito de manera exitosa el desarrollo de una metodología para la observación de la evolución del sentimiento del lenguaje económico-financiero durante un evento de máximo calado: la mayor crisis económica y social de las últimas ocho décadas. Tras una revisión completa del estado del arte en los campos del lenguaje evaluativo, los lenguajes de especialidad y el análisis de sentimiento, este desafío se ha llevado a cabo a través de un diseño metodológico basado en la evidencia, que combina el análisis de sentimiento con técnicas cualitativas. Además de las aportaciones científicas aquí expuestas, el trabajo aporta un valor añadido para la comunidad mediante el desarrollo técnico y divulgación de un lexicón de sentimiento en lengua inglesa para el dominio económico con altos niveles de precisión y exhaustividad.

## 9.2 Limitaciones y perspectivas de futuro

Los límites y nuevos desafíos que ofrece el fin de esta investigación se pueden dividir en tres campos: ampliación de fuentes de datos, ampliación del objeto de estudio y ampliación de la cuestión técnica.

En cuanto a las fuentes de texto elegidas, el diseño de este trabajo ha contemplado el corpus creado como un bloque único, aún proviniendo de dos fuentes distintas en cuanto la variedad del inglés. Por un lado, sería interesante el estudio comparativo del uso del lenguaje evaluativo entre los dos periódicos y descubrir los factores lingüísticos, periodísticos, sociales e ideológicos que influyen en las posibles diferencias. Dado que

ambas fuentes comparten en gran medida el mismo espectro ideológico, también se podrían contrastar estos resultados con otras publicaciones de distinta ideología.

Por otra parte, está la cuestión del estudio de la expresión del lenguaje económico desde distintos niveles de especialización. Cabría además estudiar el calado del uso del lenguaje económico y observar sus cambios en el discurso de la población a partir de fuentes menos formales que las cabeceras periodísticas como es el contenido generado por los usuarios: blogs, comentarios en periódicos o redes sociales. De hecho, en relación con la oleada de movimientos sociales antiausteridad de esta década, se podría estudiar si se ha producido una *rebelión* en el léxico del hablante medio durante la crisis a la hora de adquirir el léxico especializado y ser la figura bajo la cual recaen los efectos de la crisis. Sería también pertinente observar las posibles variaciones en el uso del lenguaje altamente especializado a partir de documentos emitidos y destinados a especialistas como es el caso de las divulgaciones corporativas internas. De esta forma, se podrían observar las divergencias léxicas, cognitivas y afectivas entre distintos niveles de especialización: discurso especialista, lenguaje periodístico / divulgativo y público consumidor de los términos.

Por último, se plantean varios desafíos en torno al desarrollo técnico de soluciones para el análisis de sentimiento y las aplicaciones prácticas de *LexiEcon*. Tras la ponderación de las ventajas y desventajas de las técnicas de extracción terminológica descritas aquí, se podrían aplicar en un corpus en lengua castellana de forma que esta solución sea multilingüe. Se podría, además, diseñar una aplicación en traductología para tomar en cuenta la conservación o la variación de la prosodia semántica de términos específicos considerados como traducciones sobre un período dado. Por otro lado, ampliar los ensayos realizados (Fernández-Cruz y Ortega Palomo, 2019 en prensa) en el estudio de la relación entre el discurso y los rendimientos económicos, de forma que se puedan establecer una metodología de trabajo que alcance a determinar el impacto del discurso en la economía para así realizar análisis y predicciones que beneficien al bienestar comunitario.



## Referencias

- Adelstein, A. (1998). Representación de unidades terminológicas: el modelo de término wüsteriano. *Organon*, 12(26).
- Ahmad, K., Cheng, D., & Almas, Y. (2006). Multi-lingual sentiment analysis of financial news streams. *PoS*, 001.
- Aijmer, K., & Altenberg, B. (2014). *English Corpus Linguistics*. Nueva York: Routledge.
- Alba-Juez, L., & Thompson, G. (2014). The many faces and phases of evaluation. *Evaluation in Context*, 242, 1.
- Alcaraz-Mármol, G., & Soto Almela, J. (2016). The semantic prosody of the words inmigración and inmigrante in the Spanish written media: A corpus-based study of two national newspapers. *Revista Signos*, 49(91), 145-167. <https://doi.org/10.4067/S0718-09342016000200001>
- Alim, H. S. (2011, diciembre 21). What if We Occupied Language? The New York Times. Recuperado de <http://opinionator.blogs.nytimes.com/2011/12/21/what-if-we-occupied-language/>
- Alonso Campos, A., & Renau Araque, I. (2013). Corpus Pattern Analysis in determining specialised uses of verbal lexical units. *Terminàlia*, 26–33.
- Alonso, L. E., & Fernández Rodríguez, C. J. (2013). Debemos aplacar a los mercados: el espacio del sacrificio en la crisis financiera actual. *Vínculos de Historia*, (2).
- Altamirano, C. (Ed.). (2002). *Términos críticos de sociología de la cultura* (1ª. edn.). Buenos Aires: Paidós.
- Altman, R. C. (2009, febrero 3). The Great Crash, 2008. *Foreign Affairs*. Recuperado de <https://www.foreignaffairs.com/articles/united-states/2009-01-01/great-crash-2008>
- Amiri, H., & Chua, T.-S. (2012). Mining sentiment terminology through time. *Proceedings of the 21st ACM International Conference on Information and Knowledge Management - CIKM '12*, 2060. <https://doi.org/10.1145/2396761.2398572>
- Andor, J. (2004). The master and his performance: An interview with Noam Chomsky. *Intercultural pragmatics*, 1(1), 93–111.
- Antweiler, W., & Frank, M. Z. (2004). Is all that talk just noise? The information content of internet stock message boards. *The Journal of finance*, 59(3), 1259–1294.
- Arntz, R., & Picht, H. (1995). *Introducción a la terminología*. Madrid: Fundación Germán Sánchez Ruipérez.
- Atkinson, A. C., Koopman, S. J., & Shephard, N. (1997). Detecting shocks: Outliers and breaks in time series. *Journal of Econometrics*, 80(2), 387-422. [https://doi.org/10.1016/S0304-4076\(97\)00050-X](https://doi.org/10.1016/S0304-4076(97)00050-X)
- Aue, A., & Gamon, M. (2005). Customizing sentiment classifiers to new domains: A case study. *Proceedings of recent advances in natural language processing (RANLP)*, 1, 2–1. Citeseer.
- Aussenac-Gilles, N., & Condamines, A. (2007). *Corpus et terminologie*. Cepadues.
- Baccianella, S., Esuli, A., & Sebastiani, F. (2010). Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. *Lrec*, 10, 2200–2204.
- Bajtin, M. (1934). Discourse in the Novel. *Literary theory: An anthology*, 2, 674–685.
- Bakers, L., & Sassard, S. (2012). *Insight: How the Greek Debt Puzzle was Solved*. Reuters.



- Balasubramaniam, N. (2009). User-Generated Content. *Business Aspects of the Internet of Things, Seminar of Advanced Topics*, 28-34. Zurich: Swiss Federal Institute of Technology.
- Baroni, M., Kilgarriff, A., Pomikálek, J., Rychlý, P., & others. (2006). WebBootCaT: instant domain-specific corpora to support human translators. *Proceedings of EAMT*, 247–252.
- Barr, C. (2009, mayo 27). The \$4 trillion housing headache. Recuperado 19 de abril de 2019, de CNN Money website: <https://money.cnn.com/2009/05/27/news/mortgage.overhang.fortune/index.htm>
- Barro, R. (1997). *Macroeconomics*. Cambridge, EE.UU.: MIT Press
- Batra, S., & Rao, D. (2010). Entity based sentiment analysis on twitter. *Science*, 9(4), 1–12.
- BBC Mundo. (2018, marzo 10). ¿Por qué el hombre «más odiado de Estados Unidos», Martin Shkreli, irá a la cárcel? Recuperado 20 de mayo de 2019, de BBC News Mundo website: <https://www.bbc.com/mundo/noticias-internacional-43357779>
- BBC News. (2014, junio 5). ECB imposes negative interest rate. *BBC Mundo*. Recuperado de <https://www.bbc.com/news/business-27717594>
- BBCMundo. (2009, septiembre 14). Después de la caída del gigante - BBC News Mundo. *BBC Mundo*. Recuperado de [https://www.bbc.com/mundo/economia/2009/09/090914\\_lehman\\_caída\\_rojas\\_ip](https://www.bbc.com/mundo/economia/2009/09/090914_lehman_caída_rojas_ip)
- Bednarek, M. (2006). *Evaluation in media discourse: Analysis of a newspaper corpus*. Londres: Continuum.
- Bednarek, M. (2008). *Emotion Talk Across Corpora*. Basingstoke/Nueva York: Palgrave Macmillan.
- Bednarek, M., & Caple, H. (2012). *News discourse* (Vol. 46). Londres: A&C Black.
- Bell, A. (1991). *The language of news media*. Oxford: Blackwell.
- Benamara, F., Taboada, M., & Mathieu, Y. (2017). Evaluative language beyond bags of words: Linguistic insights and computational applications. *Computational Linguistics*, 43(1), 201–264.
- Benes, E. (1969). Zur Typologie der Stilgattungen der wissenschaftlichen Prosa. *Deutsch als Fremdsprache*, 6, 225-233.
- Bergenholtz, H., Nielsen, S., & Tarp, S. (Eds.). (2009). *Lexicography at a crossroads: Dictionaries and encyclopedias today, lexicographical tools today*. Berna / Nueva York: P. Lang.
- Bernard, J. R. L. B. (1986). *The Macquarie Thesaurus*. Sydney: Macquarie.
- Biber, D. (1993). Representativeness in Corpus Design. *Literary and Linguistic Computing*, 8(4), 243-257. <https://doi.org/10.1093/lhc/8.4.243>
- Biber, D., Douglas, B., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.
- Biber, D., & Finegan, E. (1988). Adverbial stance types in English. *Discourse processes*, 11(1), 1–34.
- Biber, D., & Finegan, E. (1989). Styles of stance in English: Lexical and grammatical marking of evidentiality and affect. *Text-interdisciplinary journal for the study of discourse*, 9(1), 93–124.
- Bifet, A., Holmes, G., Pfahringer, B., & Gavalda, R. (2011). Detecting Sentiment Change in Twitter Streaming Data. *JMLR: Workshop and Conference Proceedings*, 17, 5-11.
- Blood, D. J., & Phillips, P. C. (1995). Recession headline news, consumer sentiment, the state of the economy and presidential popularity: A time series analysis 1989–1993. *International Journal of Public Opinion Research*, 7(1), 2–22.

- Bloom, K., & Argamon, S. (2010). Automated learning of appraisal extraction patterns. En J. Gries, S.Th., Wulff, S. & Davies, M. (Eds.), *Corpus-linguistic applications* (pp. 249-260). Amsterdam / Nueva York: Rodopi.
- Boiy, E., & Moens, M.-F. (2009). A machine learning approach to sentiment analysis in multilingual Web texts. *Information retrieval*, 12(5), 526–558.
- Bollen, J., Mao, H., & Pepe, A. (2011). Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena. *Fifth International AAI Conference on Weblogs and Social Media*.
- Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of computational science*, 2(1), 1–8.
- Bourigault, D. (1992). Surface Grammatical Analysis for the Extraction of Terminological Noun Phrases. *Proceedings of the 14th Conference on Computational Linguistics - Volume 3*, 977–981. <https://doi.org/10.3115/993079.993111>
- Bourigault, D., & Slodzian, M. (1999). Pour une terminologie textuelle. *Terminologies nouvelles*, 19(1999), 29–32.
- Bowker, L., & Pearson, J. (2002). *Working with specialized language: a practical guide to using corpora*. Routledge.
- Bréal, M. (1897). *Essai de sémantique*. Paris: Hachette. Recuperado de <https://archive.org/details/essaidesmantiq00bruoft/page/n11>
- Brooke, J., Tofiloski, M., & Taboada, M. (2009). Cross-linguistic sentiment analysis: From English to Spanish. *Proceedings of the international conference RANLP-2009*, 50–54.
- Brown, G. W., & Cliff, M. T. (2004). Investor sentiment and the near-term stock market. *Journal of empirical finance*, 11(1), 1–27.
- Brumby, J., & Verhoeven, M. (2010). Public expenditure after the global financial crisis. *AFTER*, 193.
- Bublitz, W. (1995). Semantic prosody and cohesive company: «somewhat predictable». *Leuvense Bijdragen*, 1-32.
- Bulher, K. (1934). *Sprachtheorie: Die Darstellungsfunktion der Sprache*. Jena: Verlag von Gustav Fischer.
- Bushee, B. J., Core, J. E., Guay, W., & Hamm, S. J. (2010). The role of the business press as an information intermediary. *Journal of Accounting Research*, 48(1), 1–19.
- Bybee, J. (2015). *Language Change: (1.ª ed.)*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139096768>
- Cabré Castellví, M. T. (2003). Theories of terminology: Their description, prescription and explanation. *Terminology*, 9(2), 163-199. <https://doi.org/10.1075/term.9.2.03cab>
- Cabré Castellví, M. T., Estopà Bagot, R., & Vivaldi Palatresi, J. (2001). Automatic term detection: A review of current systems. *Recent advances in computational terminology*, 2, 53-89.
- Cabré Castellví, M. T. (2000). Elements for a theory of terminology: Towards an alternative paradigm. *Terminology*, 6(1), 35-57. <https://doi.org/10.1075/term.6.1.03cab>
- Cabré Castellví, M. T. (1993). *La terminología: teoría, metodología, aplicaciones*. Barcelona: Antártida/Empúries.
- Cabré Castellví, M.T.. (1995). La terminología hoy: concepciones, tendencias y aplicaciones. *Ciência da Informação*, (24), 3.

- Cabré Castelví, María Teresa. (1998). Elementos para una teoría de la terminología: hacia un paradigma alternativo. *El Lenguaraz. Revista académica del Colegio de Traductores Públicos de la ciudad de Buenos Aires*, 1(1), 59-78.
- Cabré Castelví, María Teresa. (2002). Terminología y lingüística: la teoría de las puertas. *Estudios de lingüística del español*, (16), 3.
- Cadilhac, A., Benamara, F., & Aussenac-Gilles, N. (2010). Ontolexical resources for feature-based opinion mining: a case-study. *Proceedings of the 6th Workshop on Ontologies and Lexical Resources*, 77-86.
- Cambria, E., Schuller, B., Xia, Y., & Havasi, C. (2013). New avenues in opinion mining and sentiment analysis. *IEEE Intelligent systems*, 28(2), 15-21.
- Campo, Á. (2012). *The Reception of Eugen Wüster's Work and the Development of Terminology* (Université de Montréal). Recuperado de [https://papyrus.bib.umontreal.ca/xmlui/bitstream/handle/1866/9198/Campo\\_Angela\\_2012\\_these.pdf?sequence=2](https://papyrus.bib.umontreal.ca/xmlui/bitstream/handle/1866/9198/Campo_Angela_2012_these.pdf?sequence=2)
- Carassava, A. (2009, marzo 31). Greece: Citi Fined in Lehman Case. *The New York Times*. Recuperado de [https://www.nytimes.com/2009/03/31/business/global/31fobriefs-CITIFINEDINL\\_BRF.html](https://www.nytimes.com/2009/03/31/business/global/31fobriefs-CITIFINEDINL_BRF.html)
- Castiella, B. (2013, septiembre 14). Las causas del insoportable paro griego. *ABC*. Recuperado de <https://www.abc.es/economia/20130915/abci-grecia-paro-record-201309132020.html>
- Cavallin, K. (2012). Automatic extraction of potential examples of semantic change using lexical sets. *KONVENS*, 370-377. Recuperado de <http://www.academia.edu/download/30881859/proceedings.pdf#page=370>
- Chafe, W. L. (1986). Evidentiality in English Conversation and Academic Writing. En J. Nichols & W. L. Chafe (Eds.), *Evidentiality: the linguistic coding of epistemology* (pp. 261-272). Norwood: Ablex Pub. Corp.
- Chan, W. S. (2001). Stock Price Reaction to News and No-news: Drift and Reversal After Headlines. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.262452>
- Chan, S. W. K., & Chong, M. W. C. (2017). Sentiment analysis in financial texts. *Decision Support Systems*, 94, 53-64. <https://doi.org/10.1016/j.dss.2016.10.006>
- Chen, C. C., & Tseng, Y.-D. (2011). Quality evaluation of product reviews using an information quality framework. *Decision Support Systems*, 50(4), 755-768.
- Chen, M., Wang, S., Liang, P. P., Baltrušaitis, T., Zadeh, A., & Morency, L.-P. (2017). Multimodal sentiment analysis with word-level fusion and reinforcement learning. *Proceedings of the 19th ACM International Conference on Multimodal Interaction - ICMI 2017*, 163-171. <https://doi.org/10.1145/3136755.3136801>
- Cheung, C. M., & Thadani, D. R. (2012). The impact of electronic word-of-mouth communication: A literature analysis and integrative model. *Decision support systems*, 54(1), 461-470.
- Chomsky, N. (1957). *Syntactic structures*. La Haya: Mouton.
- Ciapuscio, G. E. (2010). *Textos especializados y terminología* (Vol. 6). Gerona: Documenta Universitaria.
- Ciapuscio, G.E. (1998). La terminología desde el punto de vista textual: selección, tratamiento y variación. *Organon*, 12(26).
- Collins Cobuild. (2019a). Debt definition and meaning. Recuperado 9 de febrero de 2019, de <https://www.collinsdictionary.com/es/diccionario/ingles/debt>
- Collins Cobuild. (2019b). Draconian definition and meaning. Recuperado 9 de febrero de 2019, de <https://www.collinsdictionary.com/dictionary/english/draconian>

- comScore (2012). ComScore. *Most Read Online Newspapers in the World: Mail Online, New York Times and The Guardian*. Recuperado de: <https://www.comscore.com/Insights/Infographics/Most-Read-Online-Newspapers-in-the-World-Mail-Online-New-York-Times-and-The-Guardian>
- Condamines, A. (1994). Terminologie et représentation des connaissances. *Didaskalia*, (5). <https://doi.org/10.4267/2042/23235>
- Conrad, S., & Biber, D. (2000). Adverbial marking of stance in speech and writing. En S. Hunston & G. Thompson (Eds.), *Evaluation in text: Authorial stance and the construction of discourse* (pp. 56–73). Oxford: Oxford University Press.
- Contexto y Acción. (2016, julio 27). El FMI admite su negligencia durante la crisis del euro. *Contexto y Acción*, (75). Recuperado de <http://ctxt.es/es/20160727/Politica/7520/Fondo-Monetario-internacional-informe-Grecia-tesis-euro-errores.htm>
- Cook, P., & Stevenson, S. (19-21). Automatically Identifying Changes in the Semantic Orientation of Words. En N. C. (Conference Chair), K. Choukri, B. Maegaard, J. Mariani, J. Odijk, S. Piperidis, ... D. Tapias (Eds.), *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*. Valletta, Malta: European Language Resources Association (ELRA).
- Cook, P., & Stevenson, S. (2010). Automatically identifying changes in the semantic orientation of words. *Proceedings of the 7th International Conference on Language Resources and Evaluation*, 28-34. Valletta, Malta.
- Coseriu, E. (1955). *Determinación y entorno: dos problemas de una lingüística del hablar*. Zürich: Romanisches Seminar.
- Coseriu, E. (1977). *Principios de semántica estructural*. Madrid: Gredos.
- Crowley, T., & Bower, C. (2010). *An Introduction to Historical Linguistics*. OUP USA.
- Das, S. R., & Chen, M. Y. (2007). Yahoo! for Amazon: Sentiment extraction from small talk on the web. *Management science*, 53(9), 1375–1388.
- Dave, K., Lawrence, S., & Pennock, D. M. (2003). Mining the peanut gallery: Opinion extraction and semantic classification of product reviews. *Proceedings of the 12th international conference on World Wide Web*, 519–528. ACM.
- Davis, A. (2011, mayo 31). News of the Financial Sector: Reporting on the City or to it? *openDemocracy*. Recuperado de <https://www.opendemocracy.net/en/opendemocracyuk/news-of-financial-sector-reporting-on-city-or-to-it/>
- Deng, S. (2014). *Three essays on opinion mining of social media texts*. Tesis doctoral. The University of Wisconsin-Milwaukee.
- Devitt, A., & Ahmad, K. (2007a). A lexicon for polarity: Affective content in financial news text. *Proceedings of the Conference on Language for Special Purposes*.
- Devitt, A., & Ahmad, K. (2007b). Sentiment polarity identification in financial news: A cohesion-based approach. *ACL*, 7, 1–8.
- Dewey, J. (1939). Theory of valuation. *International encyclopedia of unified science*.
- Ding, X., Liu, B., & Yu, P. S. (2008). A holistic lexicon-based approach to opinion mining. *Proceedings of the 2008 international conference on web search and data mining*, 231–240. ACM.
- Dörnyei, Z. (2007). *Research Methods in Applied Linguistics: Quantitative, Qualitative, and Mixed Methodologies*. Oxford: Oxford University Press.
- Dovring, K. (1954). Quantitative semantics in 18th century Sweden. *Public Opinion Quarterly*, 389–394.
- Ederington, L. H., & Lee, J. H. (1993). How markets process information: News releases and volatility. *The Journal of Finance*, 48(4), 1161–1191.



- Eggins, S. (2003). *Introducción a la lingüística sistémica*. Logroño: Universidad de La Rioja, Servicio de Publicaciones.
- Engelberg, J. (2008). *Costly information processing: Evidence from earnings announcements*.
- Engle, R. F., & Ng, V. K. (1993). Measuring and testing the impact of news on volatility. *The Journal of Finance*, 48(5), 1749–1778.
- Estefanía, J. (2018, septiembre 8). Diez años bastan. *El País - Babelia*. Recuperado de [https://elpais.com/cultura/2018/09/07/babelia/1536338430\\_931760.html](https://elpais.com/cultura/2018/09/07/babelia/1536338430_931760.html)
- Estopà, R. (2002). Extracción de terminología: elementos para la construcción de un extractor. *Tradterm*, 7, 225-250.
- Faber, P. (2009). The cognitive shift in terminology and specialized translation. *MonTI: Monografías de la traducción y la interpretación*, 1(1), 107-134.
- Faber, P. B. (Ed.). (2012). *A cognitive linguistics view of terminology and specialized language*. Berlin / Boston: De Gruyter Mouton.
- Faber, P., & Lopez-Rodriguez, C. (2012). *Terminology and Specialized Language*.
- Fairclough, N. (2015). *Language and power* (3ª edn.). Londres / Nueva York: Routledge, Taylor & Francis Group.
- Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, 25(2), 383–417.
- Farzindar, A., & Inkpen, D. (2015). Natural language processing for social media. *Synthesis Lectures on Human Language Technologies*, 8(2), 1–166.
- Fellbaum, C. (1998). *WordNet*. Wiley Online Library.
- Fernández-Cruz, J., & Ortega-Palomo, G. (2019, en prensa). Analyse de l'évolution du discours de la dette à travers les conférences de presse de la BCE : analyse du sentiment basé sur le lexique du domaine financier avec Lingmotif. En *Les discours de conjoncture économique en Europe*. París: L'Harmattan.
- Fernández-Cruz, J. (2014). *The Expression of Polarity in Financial Discourse and its Application in Sentiment Analysis*. Trabajo de fin de máster. Universidad de Málaga, Málaga.
- Fernández-Cruz, J. (2017). Can Sentiment in our Words be Quantified? An Introduction to Lingmotif, a Sentiment Analysis Software Tool and its Educational Application. *International Congress on the Didactics of the English Language*, 2(1).
- Fernández-Cruz, J., & Baixauli-Pérez, C. (2018). Análisis de sentimiento con Lingmotif: Aproximación metodológica al análisis de intervenciones en series temporales de la Valencia. En VV.AA. (Ed.), *La simulación en ingeniería: Trascendiendo fronteras* (1.ª ed., pp. 43-51). Esmeraldas, Ecuador: Editorial de la Pontificia Universidad Católica del Ecuador Sede Esmeraldas.
- Fernández-Lopez, M., Gomez-Perez, A., Sierra, J. P., & Sierra, A. P. (1999). Building a chemical ontology using Methontology and the Ontology Design Environment. *IEEE Intelligent Systems*, 14(1), 37-46. <https://doi.org/10.1109/5254.747904>
- Fillmore, C. J. (1985). Frames and the semantics of understanding. *Quaderni di semantica*, 6(2), 222–254.
- Fillmore, C. J., Baker, C. F., & Sato, H. (2002). The FrameNet Database and Software Tools. *LREC*.
- Finkel, J. R., Grenager, T., & Manning, C. (2005). Incorporating non-local information into information extraction systems by gibbs sampling. *Proceedings of the 43rd annual meeting on association for computational linguistics*, 363–370. Association for Computational Linguistics.

- Foucault, M. (1966). *Les mots et les choses: Une archéologie des sciences humaines*. Paris: Gallimard.
- Foucault, M. (1969). *La arqueología del saber [L'archéologie du savoir]* (2ª edn. rev.). México: Gallimard/Siglo XXI.
- Francis, G. (1986). *Anaphoric Nouns* (Discourse Analysis Monograph 11). English Language Research. Birmingham: University of Birmingham.
- Fратиани, M., & Marchionne, F. (2013). The banking bailout of the subprime crisis: Was the bang worth the buck? *Journal of International Financial Markets, Institutions and Money*, 23, 240–264.
- Freitas, L., & Vieira, R. (2013). *Comparing portuguese opinion lexicons in feature-based sentiment analysis*.
- Gablasova, D., Brezina, V., & McEnery, T. (2017). Collocations in corpus-based language learning research: Identifying, comparing, and interpreting the evidence. *Language learning*, 67(S1), 155–179.
- Gadea Boronat, L. (2017). *Estudio crítico lingüístico de publicaciones periodísticas sobre la crisis financiera en España*. Tesis doctoral. Valencia: Universitat Politècnica de València.
- Gallais-Hamonno, J. (1980). The Characteristics of English for Economists. *Fachsprache*, 2(2), 60–71.
- Gallais-Hamonno, J. (1982). Le lexique économique dans le discours scientifique anglo-saxon. *Les Cahiers de l'APLIUT. Pédagogie et recherche*, 4, 51-64.
- Gambier, Y. (2016). Des langues de spécialité aux documents multimodaux. *Pratiques*, (171-172). <https://doi.org/10.4000/pratiques.3183>
- Gamper, J., & Stock, O. (1998). Corpus-based terminology. *Terminology*, 5(2), 147-159. <https://doi.org/10.1075/term.5.2.05gam>
- García, D. (2013). Sentiment during Recessions. *The Journal of Finance*, 68(3), 1267–1300.
- Garofalo, G. (2018). Persiguiendo con imparcialidad “el total desprecio a la Constitución”: El léxico valorativo en la Querrela del Fiscal de Cataluña contra Carme Forcadell i Lluís. *Quaderns de Filologia - Estudis Lingüístics*, 22(22), 79. <https://doi.org/10.7203/qf.22.11302>
- Gautier, L. (2014). Des langues de spécialité à la communication spécialisée: un nouveau paradigme de recherche à l'intersection entre sciences du langage, info-com et sciences cognitives? *Etudes Interdisciplinaires en Sciences humaines*, 1, 225-245.
- Giachanou, A., & Crestani, F. (2016). *Tracking Sentiment by Time Series Analysis*. 1037-1040. <https://doi.org/10.1145/2911451.2914702>
- Gillmor, D. (2011, noviembre 22). Occupy language: the struggle over meaning. *The Guardian*. Recuperado de <https://www.theguardian.com/commentisfree/cifamerica/2011/nov/22/occupy-language-struggle-over-meaning>
- Gläser, R. (1988). Neologismen im englischen Wortschatz der Ökonomie und Probleme der Übersetzung. *Sprache und Information*, 172–182.
- Glassner, V. (2010). *The public sector in the crisis* (European Trade Union Institution). Bruselas.
- Godbole, N., Srinivasaiyah, M., & Skiena, S. (2007). Large-Scale Sentiment Analysis for News and Blogs. *Icwsn*, 7(21), 219–222.
- Godin, R. (2016, ago). Grecia: El FMI reconoce sus errores, pero todo sigue igual (A. Andrades, Trad.). *Contexto y Acción*, (76). Recuperado de

- <http://ctxt.es/es/20160803/Politica/7600/Grecia-deuda-FMI-acreedores-ajuste.htm>
- Gordon, R. J. (2019). When will the longest expansion end? *Business Economics*, 54(1), 69-78. <https://doi.org/10.1057/s11369-018-00108-1>
- Gouws, S., Metzler, D., Cai, C., & Hovy, E. (2011). Contextual Bearing on Linguistic Variation in Social Media. *Proceedings of the Workshop on Languages in Social Media*, 20–29. Recuperado de <http://dl.acm.org/citation.cfm?id=2021109.2021113>
- Goźdz-Roszkowski, S., & Hunston, S. (2016). Corpora and beyond – investigating evaluation in discourse: Introduction to the special issue on corpus approaches to evaluation. *Corpora*, 11(2), 131-141. <https://doi.org/10.3366/cor.2016.0089>
- Gruber, T. R. (1995). Toward principles for the design of ontologies used for knowledge sharing? *International Journal of Human-Computer Studies*, 43(5-6), 907-928. <https://doi.org/10.1006/ijhc.1995.1081>
- Giachanou, A., & Crestani, F. (2016). *Tracking Sentiment by Time Series Analysis*. 1037-1040. <https://doi.org/10.1145/2911451.2914702>
- Giachanou, A., Mele, I., & Crestani, F. (2016). Explaining sentiment spikes in twitter. *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*, 2263–2268. ACM.
- Haas, M., & Versley, Y. (2015). Subsentential sentiment on a shoestring: A crosslingual analysis of compositional classification. *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 694–704.
- Haase, F.-A. (2010). The Linguistic Representation of Economic Breakdowns in the Mass Media Language as Inverted Rhetoric of Vivity. *Argumentum*, 6, 37–50.
- Halliday, M. A. K. (1975). *Learning How to Mean: Explorations in the Development of Language*. Londres: Edward Arnold.
- Halliday, M. A. K., & Matthiessen, C. M. I. M. (2014). *Halliday's introduction to functional grammar* (Fourth Edition). Milton Park, Abingdon, Oxon: Routledge.
- Hamilton, W. L., Clark, K., Leskovec, J., & Jurafsky, D. (2016). Inducing domain-specific sentiment lexicons from unlabeled corpora. *Proceedings of the Conference on Empirical Methods in Natural Language Processing. Conference on Empirical Methods in Natural Language Processing, 2016*, 595. NIH Public Access.
- Hamilton, W. L., Leskovec, J., & Jurafsky, D. (2016a). Cultural shift or linguistic drift? comparing two computational measures of semantic change. *Proceedings of the Conference on Empirical Methods in Natural Language Processing. Conference on Empirical Methods in Natural Language Processing, 2016*, 2116. NIH Public Access.
- Hamilton, W. L., Leskovec, J., & Jurafsky, D. (2016b). Diachronic word embeddings reveal statistical laws of semantic change. *arXiv preprint arXiv:1605.09096*.
- Hampp, T., & Lang, A. (2005). Semantic search in websphere information integrator omnifind edition: The case for semantic search. *IBM Developer Works*.
- Harris, Z. (1982). Discourse and Sublanguage. En R. Kittredge & J. Lehrberger (Eds.), *Sublanguage: studies of language in restricted semantic domains* (pp. 231-236). Berlín/Nueva York: Walter de Gruyter.
- Harvey, D. I., Leybourne, S. J., & Newbold, P. (2002). Seasonal unit root tests with seasonal mean shifts. *Economics Letters*, 76(2), 295-302. [https://doi.org/10.1016/S0165-1765\(02\)00057-5](https://doi.org/10.1016/S0165-1765(02)00057-5)

- Hatzivassiloglou, V., & McKeown, K. R. (1997). Predicting the semantic orientation of adjectives. *Proceedings of the 35th annual meeting of the association for computational linguistics and eighth conference of the european chapter of the association for computational linguistics*, 174–181. Association for Computational Linguistics.
- Havránek, B. (1964). The functional differentiation of the standard language. *Chapters from the history of Czech functional linguistics*.
- Hearst, M. A. (1992). Direction-based text interpretation as an information access refinement. *Text-based intelligent systems: current research and practice in information extraction and retrieval*, 257–274.
- Hoey, M. (2000). Persuasive Rhetoric in Linguistics: A Stylistic Study of Some Features of the language of Noam Chomsky. En S. Hunston & G. Thompson (Eds.), *Evaluation in text: Authorial stance and the construction of discourse* (pp. 28-37). Oxford: Oxford University Press.
- Hoey, M. (2005). *Lexical priming a new theory of words and language*. Londres/Nueva York: Routledge/AHRB.
- Hoffmann, L. (1998). *Els llenguatges d'especialitat: selecció de textos*. Barcelona: Institut Universitari de Lingüística Aplicada.
- Holborow, M. (2015). *Language and neoliberalism*. Nueva York: Routledge.
- Hu, M., & Liu, B. (2004). Mining and summarizing customer reviews. *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, 168–177. ACM.
- Huang, A. H., Zang, A. Y., & Zheng, R. (2014). Evidence on the information content of text in analyst reports. *The Accounting Review*, 89(6), 2151–2180.
- Hunston, S. (2014). Flavours of Corpus Linguistics. *Paper given at Charles University, Prague 2012 and at Corpus Linguistics 2011*. Presentado en Birmingham. Birmingham.
- Hunston, S., & Francis, G. (2000). *Pattern grammar: A corpus-driven approach to the lexical grammar of English* (Vol. 4). Amsterdam/Filadelfia: John Benjamins.
- Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, 9(3), 90–95. <https://doi.org/10.1109/MCSE.2007.55>
- Iedema, R., Feez, S., & White, P. R. (1994). Media literacy. *Sydney: Disadvantaged schools program, NSW Department of school education*.
- Israel, M. (2011). *The Grammar of Polarity: Pragmatics, Sensitivity, and the Logic of Scales*. <https://doi.org/10.1017/CBO9780511975288>
- Jakubíček, M., Kilgarriff, A., Kovář, V., Rychlý, P., & Suchomel, V. (2013). The TenTen Corpus Family. *7th International Corpus Linguistics Conference CL*, 125-127.
- Jakubíček, M., & Smerk, P. (2016). Large Scale Keyword Extraction using a Finite State Backend. *RASLAN*, 143–146.
- Jakubíček, Miloš, Kilgarriff, A., Kovář, V., Rychlý, P., & Suchomel, V. (2014). Finding terms in corpora for many languages with the Sketch Engine. *Proceedings of the Demonstrations at the 14th Conference of the European Chapter of the Association for Computational Linguistics*, 53–56.
- Janot, P. (2012). Histoire d'une crise, histoire d'un terme : de quelques stratégies discursives autour du terme subprime dans la presse généraliste française. En L. Gautier (Ed.), *Les discours de la bourse et de la finance* (pp. 47-62). Berlín: Frank und Timme.
- Jiménez Gómez, I. (2015). Los 7 «mejores» eufemismos económicos de la Troika - El salmón contracorriente. Recuperado 15 de abril de 2018, de El Salmón



- Contracorriente website: <http://www.elsalmoncontracorriente.es/?Los-7-mejores-eufemismos-642>
- Jindal, N., & Liu, B. (2006). Identifying comparative sentences in text documents. *Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval*, 244–251. ACM.
- Jordà, Ò., Schularick, M., & Taylor, A. M. (2016). The great mortgaging: housing finance, crises and business cycles. *Economic Policy*, 31(85), 107–152.
- Jorrín, J. G. (2018, septiembre 16). Por qué EEUU recuperó en 2013 el dinero del rescate y España perderá 60.000 millones. *El Confidencial*. Recuperado de [https://www.elconfidencial.com/economia/2018-09-15/por-que-eeuu-recupero-dinero-rescate-espana-perdera\\_1616325/](https://www.elconfidencial.com/economia/2018-09-15/por-que-eeuu-recupero-dinero-rescate-espana-perdera_1616325/)
- Jucker, A. H. (1992). *Social stylistics: Syntactic variation in British newspapers* (Vol. 6). Berlín/Nueva York: Walter de Gruyter.
- Jullian, P. M. (2009). *An exploration of strategies to convey evaluation in the "NoteBook" texts* (PhD Thesis). University of Birmingham.
- Kabatek, J. (2005). Tradiciones discursivas y cambio lingüístico. *Lexis*, 29(2), 151–177.
- Kageura, K. (2002). *The dynamics of terminology: A descriptive theory of term formation and terminological growth*. Amsterdam ; Filadelfia: John Benjamins.
- Kamp, H., & Reyle, U. (1993). *From discourse to logic*. Dordrecht: Kluwer.
- Kamps, J., Marx, M., Mokken, R. J., De Rijke, M., & others. (2004). Using WordNet to measure semantic orientations of adjectives. *LREC*, 4, 1115–1118. Citeseer.
- Kant, I. (1785). *Groundwork for the metaphysics of Morals* (trad. A. Wood, 2002). New Haven: Yale University Press.
- Kaplan, N. (2004). Nuevos desarrollos en el estudio de la evaluación en el lenguaje: La Teoría de la Valoración. *Boletín de Lingüística*, (22), 52-78.
- Kaplan, N. (2007). *La construcción discursiva del evento conflictivo en las noticias por televisión*. Universidad Central de Venezuela, Caracas.
- Kennedy, A., & Inkpen, D. (2006). Sentiment classification of movie reviews using contextual valence shifters. *Computational intelligence*, 22(2), 110–125.
- Kensil, S., & Margraf, K. (2015). *The Advantage of Failing First: Bear Stearns V. Lehman Brothers*. Recuperado de <https://papers.ssrn.com/abstract=2689667>
- Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., ... Suchomel, V. (2014). The Sketch Engine: ten years on. *Lexicography*, 1(1), 7–36.
- Kilgarriff, A., Rychly, P., Smrz, P., & Tugwell, D. (2004). Itri-04-08 the sketch engine. *Information Technology*, 105, 116.
- Kilgarriff, A., & Yallop, C. (2000). What's in a Thesaurus? *LREC*, 1371–1379.
- Kim, S.-M., & Hovy, E. (2004). Determining the sentiment of opinions. *Proceedings of the 20th international conference on Computational Linguistics*, 1367. Association for Computational Linguistics.
- Klein, F. C., & Prestbo, J. A. (1974). *News and the Market*. Washington: H. Regnery Co.
- Koch, P. (1997). Diskurstraditionen: Zu ihrem sprachtheoretischen Status und ihrer Dynamik. En Haye, Thomas, D. Tophinke, & B. Frank (Eds.), *Gattungen mittelalterlicher Schriftlichkeit*. (pp. 43-79). Tübingen: Narr.
- Koppel, M., & Shtrimberg, I. (2006). Good news or bad news? let the market decide. En *Computing attitude and affect in text: Theory and applications* (pp. 297–301). Chicago: Springer.
- Kövecses, Z. (2003). *Metaphor and emotion: Language, culture, and body in human feeling*. Cambridge: Cambridge University Press.

- Krishnamoorthy, S. (2017). Sentiment analysis of financial news articles using performance indicators. *Knowledge and Information Systems*, 1-22. <https://doi.org/10.1007/s10115-017-1134-1>
- Kristeva, J. (1986). Word, Dialogue and Novel. En T. Moi (Ed.), *The Kristeva Reader*. Oxford: Basil Blackwell.
- Krugman, P. (2012, julio 29). Crash of the Bumblebee. The New York Times. Recuperado de <https://www.nytimes.com/2012/07/30/opinion/krugman-crash-of-the-bumblebee.html>
- Krumm, J., Davies, N., & Narayanaswami, C. (2008). User-generated content. *IEEE Pervasive Computing*, (4), 10–11.
- Labov, W. (1972). *Sociolinguistic patterns*. Filadelfia: University of Pennsylvania Press.
- Labov, W. (1984). Intensity. En D. Schrifin (Ed.), *Meaning, form, and use in context: Linguistic applications* (Vol. 70, pp. 48-79). Washington D.C.: Georgetown University Press.
- Lainé-Cruzel, S. (2004). Documents, ressources, données : Les avatars de l'information numérique. *Revue I3 - Information Interaction Intelligence*, 4(1). Recuperado de [https://archivesic.ccsd.cnrs.fr/sic\\_00001018/document](https://archivesic.ccsd.cnrs.fr/sic_00001018/document)
- Lakoff, G. (1988). Cognitive semantics. *Meaning and mental representations*, 119, 154.
- Lau, R. Y., Li, C., & Liao, S. S. (2014). Social analytics: Learning fuzzy product ontologies for aspect-oriented sentiment analysis. *Decision Support Systems*, 65, 80–94.
- Leech, G. (1992). Corpora and theories of linguistic performance. En J. Svartvik (Ed.), *Directions in Corpus Linguistics* (pp. 105–22). Berlin: Mouton de Gruyter.
- Lemke, J. L. (1998). Resources for attitudinal meaning: Evaluative orientations in text semantics. *Functions of Language*, 5(1), 33-56. <https://doi.org/10.1075/fo1.5.1.03lem>
- León-Araúz, P., Martin, A. S., & Reimerink, A. (2018). The EcoLexicon English Corpus as an Open Corpus in Sketch Engine. *arXiv preprint arXiv:1807.05797*.
- Lerat, P. (1995). *Les langues spécialisées*. Paris: PUF.
- Li, F. (2010). The information content of forward-looking statements in corporate filings—A naïve Bayesian machine learning approach. *Journal of Accounting Research*, 48(5), 1049–1102.
- Li, Q., Wang, T., Li, P., Liu, L., Gong, Q., & Chen, Y. (2014). The effect of news and public mood on stock movements. *Information Sciences*, 278, 826–840.
- Li, X., Xie, H., Chen, L., Wang, J., & Deng, X. (2014). News impact on stock price return via sentiment analysis. *Knowledge-Based Systems*, 69, 14–23.
- Liu, B. (2007). *Web data mining: exploring hyperlinks, contents, and usage data*. Chicago: Springer.
- Liu, B. (2010). Sentiment Analysis and Subjectivity. *Handbook of natural language processing*, 2, 627–666.
- Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies*, 5(1), 1–167.
- Liu, B. (2015). *Sentiment analysis: Mining opinions, sentiments, and emotions*. Nueva York: Cambridge University Press.
- López Manjón, P. (2013, abril 25). El paro llega a 6.202.700 personas y la tasa sube al récord del 27,16% en el primer trimestre de 2013. Recuperado 19 de abril de 2019, de RTVE.es website: <http://www.rtve.es/noticias/20130425/paro-llega-6202700-personas-tasa-2716-primer-trimestre-2013/649182.shtml>

- López Muñoz, J. M., & Verine, B. (2014). L'événement : Du syntagme au discours. *Cahiers de praxématique*, (63). Recuperado de <http://journals.openedition.org/praxematique/2336>
- López Noguero, F. (2002). El análisis de contenido como método de investigación. *XXI, Revista de Educación*, 4, 167-179.
- López Rodríguez, C. I., Buendía Castro, M., & García Aragón, A. (2010). User needs to the test: evaluating a terminological knowledge base on the environment by trainee translators. *Journal of Specialised Translation*, (18).
- López Rodríguez, C. I., Faber, P., León Arauz, P., Prieto Velasco, J. A., & Tercedor Sánchez, M. (2010). La Terminología basada en marcos y su aplicación a las Ciencias Ambientales: los proyectos Marcocosta y Ecosistema. *Arena Romanística*, 7(10), 52-74.
- Loughran, T., & McDonald, B. (2011). When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *The Journal of Finance*, 66(1), 35-65.
- Loughran, T., & McDonald, B. (2014). Measuring readability in financial disclosures. *The Journal of Finance*, 69(4), 1643-1671.
- Loughran, T., & McDonald, B. (2015). The use of word lists in textual analysis. *Journal of Behavioral Finance*, 16(1), 1-11.
- Loughran, T., & McDonald, B. (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research*, 54(4), 1187-1230.
- Louw, B. (1993). Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies. En M. Baker, F. Gill, & Tognini-Bonelli, E. (Eds.), *Text and technology: In honour of John Sinclair*. 157-177. Amsterdam: John Benjamins.
- Louwrens, L. J. (1993). Semantic change in loan words. *South African Journal of African Languages*, 13(1), 8-16. <https://doi.org/10.1080/02572117.1993.10586958>
- Luna-Alonso, A. (2016). La "refundación del capitalismo": Estudio terminológico en la prensa francesa y española. *Filología e Lingüística Portuguesa*, 18(1), 145-174.
- Lusztig, P., & Schwab, B. (2014). *Managerial finance in a Canadian setting*. Butterworth-Heinemann.
- Lyons, J. (1977). *Semantics (vols I & II)*. Cambridge: Cambridge University Press.
- Macken-Horarik, M., & Isaac, A. (2014). Appraising appraisal. *Evaluation in context*, 67-92.
- MacKinnon, K. (2012). User-Generated Content vs. Advertising: Do Consumers Trust the Word of Others over Advertisers? *The Elon Journal of Undergraduate Research in Communications*, 3(1), 14-22.
- Malo, P., Sinha, A., Korhonen, P., Wallenius, J., & Takala, P. (2014). Good debt or bad debt: Detecting semantic orientations in economic texts. *Journal of the Association for Information Science and Technology*, 65(4), 782-796. <https://doi.org/10.1002/asi.23062>
- Marcus, M., Santorini, B., & Marcinkiewicz, M. A. (1993). *Building a large annotated corpus of English: The Penn Treebank*.
- Marsh, D. (2016). *Europe's Deadlock: How the Euro Crisis Could be Solved-and why it Still Won't Happen*. New Haven: Yale University Press.
- Martin, James R. (1992). *English text: System and structure*. John Benjamins Publishing.
- Martin, J.R., & White, P. (2005). *The Language of Evaluation: Appraisal in English*. Basingstoke, Hampshire: Palgrave Macmillan.
- Martin, K. C. (2012, septiembre 14). The lexical legacy of Occupy Wall Street. Recuperado 3 de mayo de 2016, de OxfordWords blog website: <http://blog.oxforddictionaries.com/2012/09/occupy-wall-street/>

- Martín-Valdivia, M.-T., Martínez-Cámara, E., Perea-Ortega, J.-M., & Ureña-López, L. A. (2013). Sentiment polarity detection in Spanish reviews combining supervised and unsupervised approaches. *Expert Systems with Applications*, 40(10), 3934–3942.
- Massey, D. (2013). Vocabularies of the economy. *Soundings*, 54(54), 9-22. <https://doi.org/10.3898/136266213807299023>
- Masterman, M. (1957). The thesaurus in syntax and semantics. *Mechanical Translation*, 4, 1–2.
- McEnery, T., & Hardie, A. (2011). *Corpus linguistics: Method, theory and practice*. Cambridge: Cambridge University Press.
- McEnery, T., Xiao, R., & Tono, Y. (2006). *Corpus-based Language Studies: An Advanced Resource Book*. Londres/Nueva York: Taylor & Francis.
- McKinney, W. (2010). Data Structures for Statistical Computing in Python. En S. van der Walt & J. Millman (Eds.), *Proceedings of the 9th Python in Science Conference* (pp. 51-56).
- Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 5(4), 1093–1113.
- Melvin, M., & Yin, X. (2000). Public information arrival, exchange rate volatility, and quote frequency. *The Economic Journal*, 110(465), 644–661.
- Meyer, C. F. (2002). *English Corpus Linguistics: An Introduction*. Cambridge: Cambridge University Press.
- Meyer, I., & Mackintosh, K. (1996). The corpus from a terminographer's viewpoint. *International Journal of Corpus Linguistics*, 1(2), 257–285.
- Miles, M. B., & Huberman, A. M. (1994). *An expanded sourcebook qualitative data analysis* (2nd edn.). Londres: Sage Publications.
- Miller, G. A. (1995). WordNet: a lexical database for English. *Communications of the ACM*, 38(11), 39–41.
- Mohammad, S., Dunne, C., & Dorr, B. (2009). Generating high-coverage semantic orientation lexicons from overtly marked words and a thesaurus. *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 2-Volume 2*, 599–608. Association for Computational Linguistics.
- Mohammad, S., Kiritchenko, S., & Zhu, X. (2013). NRC-Canada: Building the State-of-the-Art in Sentiment Analysis of Tweets. *Second Joint Conference on Lexical and Computational Semantics (\*SEM), Volume 2: Proceedings of the Seventh International Workshop on Semantic Evaluation (SemEval 2013)*, 2, 321–327.
- Mohammad, S. M., & Turney, P. D. (2010). Emotions evoked by common words and phrases: Using mechanical turk to create an emotion lexicon. *Proceedings of the NAACL HLT 2010 workshop on computational approaches to analysis and generation of emotion in text*, 26–34. Association for Computational Linguistics.
- Moirand, S. (2007). *Les Discours de la Presse Quotidienne: Observer, Analyser, Comprendre*. Paris: Presses Universitaires de France - PUF.
- Moirand, S. (2016). “La construction de l'événement dans la presse entre sémantique discursive, hétérogénéités énonciatives et inscription de l'émotion”. Recuperado de <https://hal-univ-paris3.archives-ouvertes.fr/hal-01476098>
- Mongin, P. (2006). Value Judgments and Value Neutrality in Economics. *Economica*, 73(290), 257-286. <https://doi.org/10.1111/j.1468-0335.2006.00501.x>
- Montague, R. (1974). *Formal Philosophy: Selected Papers of Richard Montague*. Ed. and with an Introd. by Richmond H. Thomason. Yale University Press.



- Montero Martínez, S. (2003). Estructuración conceptual y formalización terminográfica de frases en el subdominio de la oncología. *Estudios de Lingüística del Español (ELiEs)*, 19.
- Moreno Ortiz, A., & Perez Hernandez, C. (2014). Form and function in evaluative language: The use of corpora to identify context valence shifters in a linguistically-motivated sentiment analysis system. En M. de los Á. Gómez González, F. J. Ruiz de Mendoza Ibáñez, F. González García, & A. Downing (Eds.), *Pragmatics & Beyond New Series* (Vol. 247, pp. 87-110). Recuperado de <https://benjamins.com/catalog/pbns.247.05mor>
- Moreno-Ortiz, A. (2016, abril 5). *Extracting Domain-Specific Features for Sentiment Analysis Using Simple NLP Techniques: Running Shoes Reviews*. Presentado en Congreso Internacional de Lingüística de Corpus 2016, Málaga.
- Moreno-Ortiz, A. (2017a). Lingmotif: A User-focused Sentiment Analysis Tool. *Procesamiento de Lenguaje Natural*, 58.
- Moreno-Ortiz, A. (2017b). *Lingmotif: Sentiment Analysis for the Digital Humanities*. Presentado en Proceedings of the EACL, Valencia, Spain.
- Moreno-Ortiz, A. (2019). Mi opinión cuenta: la expresión del sentimiento en la Red. En S. Robles Avila & A. Moreno-Ortiz (Eds.), *Comunicación mediada por ordenador: la lengua, el discurso y la imagen* (1a edn.). Madrid: Cátedra.
- Moreno-Ortiz, A., & Fernández-Cruz, J. (2015). Identifying Polarity in Financial Texts for Sentiment Analysis: A Corpus-based Approach. *Procedia - Social and Behavioral Sciences*, 198, 330-338. <http://dx.doi.org/10.1016/j.sbspro.2015.07.451>
- Moreno-Ortiz, A., & Pérez-Hernández, C. (2012). Lexicon-Based Sentiment Analysis of Twitter Messages in Spanish. *Procesamiento del Lenguaje Natural*, 50(0), 93-100.
- Moreno-Ortiz, A., & Pérez-Hernández, C. (2013). Managing Multiword Expressions in a Lexicon-Based Sentiment Analysis System for Spanish. *Proceedings of the 9th Workshop on Multiword Expressions*, 1-10. Atlanta, USA: Association for Computational Linguistics.
- Moreno-Ortiz, A., & Pérez-Hernández, C. (2017). Tecnolengua Lingmotif at TASS 2017: Spanish Twitter Dataset Classification Combining Wide-coverage Lexical Resources and Text Features. *TASS 2017: Workshop on Semantic Analysis at SEPLN*, 35-42. Recuperado de [http://ceur-ws.org/Vol-1896/p3\\_tecnolengua\\_tass2017.pdf](http://ceur-ws.org/Vol-1896/p3_tecnolengua_tass2017.pdf)
- Moreno-Ortiz, A., & Pérez-Hernández, C. (2018). Lingmotif-lex: a Wide-coverage, State-of-the-art Lexicon for Sentiment Analysis. *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC-2018)*.
- Moreno-Ortiz, A., Pérez-Hernández, C., & Hidalgo-García, R. (2011). Domain-neutral, Linguistically-motivated Sentiment Analysis: a performance evaluation. *Actas del XXVII Congreso de la SEPLN*, 847-856. Huelva, Spain.
- Moreno-Ortiz, A., Pérez-Pozo, Á., & Torres-Sánchez, S. (2010). Sentitext: sistema de análisis de sentimiento para el español. *Procesamiento del Lenguaje Natural*, 45, 297-298.
- Moreno-Ortiz, Antonio, Pineda Castillo, F., & Hidalgo García, R. (2010). Análisis de Valoraciones de Usuario de Hoteles con Sentitext: Un sistema de análisis de sentimiento independiente del dominio. *Procesamiento del lenguaje natural*, 45, 31-39.

- Morley, J., & Partington, A. (2009). A few Frequently Asked Questions about semantic—or evaluative—prosody. *International Journal of Corpus Linguistics*, 14(2), 139–158.
- Mueller, R. (2019, abril 19). Diez años desde la caída de Lehman Brothers. Recuperado 19 de abril de 2019, de Funds People website: <https://es.fundspeople.com/news/diez-anos-desde-la-caida-de-lehman-brothers>
- Murison-Bowie, S. (1996). Linguistic corpora and language teaching. *Annual review of applied linguistics*, 16, 182–199.
- Musat, C., & Trausan-Matu, S. (2010). The impact of valence shifters on mining implicit economic opinions. *International Conference on Artificial Intelligence: Methodology, Systems, and Applications*, 131–140. Springer.
- Nafra, I. (2017). *La reinención de The New York Times: Cómo la dama gris del periodismo se está adaptando (con éxito) a la era móvil*. Austin: Knight Center for Journalism in the Americas The University of Texas at Austin.
- Nasukawa, T., & Yi, J. (2003). Sentiment analysis: Capturing favorability using natural language processing. *Proceedings of the 2nd international conference on Knowledge capture*, 70–77. ACM.
- Navigli, R., & Ponzetto, S. P. (2012). BabelNet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network. *Artificial Intelligence*, 193, 217-250. <https://doi.org/10.1016/j.artint.2012.07.001>
- Nelson, E. (2017, julio 26). Five years ago today, Mario Draghi saved the euro. *Quartz*. Recuperado de <https://qz.com/1038954/whatever-it-takes-five-years-ago-today-mario-draghi-saved-the-euro-with-a-momentous-speech/>
- Niederhoffer, V. (1971). The analysis of world events and stock prices. *The Journal of Business*, 44(2), 193–219.
- Nguyen Phuoc, V. C. (2011). Lexical And Semantic Features of a Financial, Accounting, and Banking English Text. *Ngôn Ngữ, Văn Hóa & Dịch Thuật*. Recuperado de <https://nguyenphuocvinhco.com/2011/07/21/lexical-and-semantic-features-of-a-financial-accounting-and-banking-english-text/>
- Noeth, B., Sengupta, R., & otros. (2012). Global European banks and the financial crisis. *Federal Reserve Bank of St. Louis Review*, 94(6), 457–79.
- Noy, N. F., McGuinness, D. L., & otros. (2001). *Ontology development 101: A guide to creating your first ontology*. Stanford knowledge systems laboratory technical report KSL-01-05 and ....
- Observatorio Metropolitano. (2011). *Crisis y revolución en Europa = People of Europe, rise up!* Madrid: Traficantes de Sueños.
- Ocampo, J. A. (2008). O colapso do sistema financeiro mundial. *Estudos Avançados, São Paulo*, 22(64).
- O'Halloran, K., & Coffin, C. (2004). Checking over-interpretation and under-interpretation: help from corpora in critical linguistics. *Text and texture: systemic functional viewpoints on the nature and structure of text*, 275–297.
- O'Hare, N., Davy, M., Bermingham, A., Ferguson, P., Sheridan, P., Gurrin, C., & Smeaton, A. F. (2009). Topic-dependent sentiment analysis of financial blogs. *Proceedings of the 1st international CIKM workshop on Topic-sentiment analysis for mass opinion*, 9–16. ACM.
- O'Leary, D. (2011). Blog Mining-Review and Extensions. *Decision Support Systems*, 821-830.
- O'Neil, C. (2017). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Portland: Broadway Books.

- Osgood, C. E., & Richards, M. M. (1973). From yang and yin to and or but. *Language*, 380–412.
- Österreicher, W. (1997). Zur fundierung von diskurstraditionen. En Haye, Thomas & D. Tophinke (Eds.), *Gattungen mittelalterlicher Schriftlichkeit* (pp. 19-41). Tübingen: Narr.
- Oxford English Dictionary. (2019). credit | Definition of credit in English. Recuperado 9 de febrero de 2019, de Oxford Dictionaries: <https://en.oxforddictionaries.com/definicion/credit>
- Paiva Jacobini, M. L. (2008). Economic journalism and the conception of market: a content analysis of the economy sections of Folha de S. Paulo and O Estado de S. Paulo. *Brazilian Journalism Research*, 4(2), 176–194.
- Palmer, H. E. (1933). *Second interim report on English collocations*. Presentado en Tokyo, Institute for Research in English Teaching, 1933, Tokyo, Institute for Research in English Teaching, 1933.
- Palmieri, R., & Palmieri, C. (2012). Text types, activity types and the genre system of financial communication. En L. Gautier (Ed.), *Les discours de la bourse et de la finance* (pp. 85-105). Berlín: Frank und Timme.
- Pang, B., & Lee, L. (2005). Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales. *Proceedings of the 43rd annual meeting on association for computational linguistics*, 115–124. Association for Computational Linguistics.
- Pang, B., Lee, L., & others. (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval*, 2(1–2), 1–135.
- Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up?: sentiment classification using machine learning techniques. *Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10*, 79–86. Association for Computational Linguistics.
- Partington, A. (1998). *Patterns and meanings: Using corpora for English language research and teaching* (Vol. 2). John Benjamins Publishing.
- Paúls, B. G. (2017). Pseudopolítica en la red: indicadores discursivos de desideologización en Twitter. *Pragmalingüística*, (25), 189–210.
- Pawar, A. B., Jawale, M. A., & Kyatanavar, D. N. (2016). Fundamentals of Sentiment Analysis: Concepts and Methodology. En W. Pedrycz & S.-M. Chen (Eds.), *Sentiment Analysis and Ontology Engineering* (Vol. 639, pp. 25-48). Cham: Springer International Publishing.
- Pazienza, M. T., Pennacchiotti, M., & Zanzotto, F. M. (2005). Terminology Extraction: An Analysis of Linguistic and Statistical Approaches. *Knowledge Mining*, 255-279. [https://doi.org/10.1007/3-540-32394-5\\_20](https://doi.org/10.1007/3-540-32394-5_20)
- Pédauque, R. T. (2005). *Le texte en jeu Permanence et transformations du document*. Recuperado de [https://archivesic.ccsd.cnrs.fr/sic\\_00001401/document](https://archivesic.ccsd.cnrs.fr/sic_00001401/document)
- Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). *The development and psychometric properties of LIWC2015*.
- Pérez, C. (2012, julio 26). “El BCE hará lo necesario para sostener el euro”. Recuperado 2 de junio de 2019, de EL PAÍS website: [https://elpais.com/economia/2012/07/26/actualidad/1343298454\\_662169.html](https://elpais.com/economia/2012/07/26/actualidad/1343298454_662169.html)
- Pérez, C. (2018, septiembre 9). Recesión a lo grande: Crónica de los 10 años de crisis que cambiaron el mundo. Recuperado 19 de abril de 2019, de EL PAÍS website: [https://elpais.com/economia/2018/09/07/actualidad/1536333092\\_303809.html](https://elpais.com/economia/2018/09/07/actualidad/1536333092_303809.html)

- Pérez Hernández, C. (2002). Explotación de los corpórea textuales informatizados para la creación de bases de datos terminológicas basadas en el conocimiento. *Estudios de Lingüística del Español (ELiEs)*, 18.
- Pérez-Hernández, C., Moreno-Ortiz, A., & Butler, C. (2004). Corpus based approaches to language and language research. En F. Pineda-Castillo (Ed.), *An Integrating Approach to Applied Linguistics: Academic and Professional Insights* (pp. 255-289). Albolote: Comares.
- Periñán-Pascual, C. (2015). The underpinnings of a composite measure for automatic term extraction: the case of SRC. *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication*, 21(2), 151–179.
- Piketty, T. (2013). *Le capital au XXIe siècle*. Paris: Éditions du Seuil.
- Polanyi, L., & Zaenen, A. (2006). Contextual Valence Shifters. En J. G. Shanahan, Y. Qu, & J. Wiebe (Eds.), *Computing Attitude and Affect in Text: Theory and Applications* (pp. 1-10). [https://doi.org/10.1007/1-4020-4102-0\\_1](https://doi.org/10.1007/1-4020-4102-0_1)
- Poon, S.-H., & Granger, C. (2005). Practical issues in forecasting volatility. *Financial analysts journal*, 61(1), 45–56.
- Popescu, A.-M., & Etzioni, O. (2007). Extracting product features and opinions from reviews. En *Natural language processing and text mining* (pp. 9–28). Springer.
- Psara, M., & Lazaro, A. (2018, 08). Moscovici reconoce que «hubo errores» en el rescate griego. *euronews*. Recuperado de <https://es.euronews.com/2018/08/20/moscovici-reconoce-que-hubo-errores-en-el-rescate-griego>
- Pustejovsky, J., Ingria, B., Sauri, R., Castano, J., Littman, J., Gaizauskas, R., ... Mani, I. (2005). The specification language TimeML. *The language of time: A reader*, 545–557.
- Qiu, G., He, X., Zhang, F., Shi, Y., Bu, J., & Chen, C. (2010). DASA: dissatisfaction-oriented advertising based on sentiment analysis. *Expert Systems with Applications*, 37(9), 6182–6191.
- Quenet, G. (2010). Fléaux de Dieu ou catastrophes naturelles?. Les tremblements de terre en France à l' époque moderne. *Terrain. Anthropologie & sciences humaines*, (54), 10–25.
- Quine, W. V. (1951). Ontology and ideology. *Philosophical Studies*, 2(1), 11-15. <https://doi.org/10.1007/BF02198233>
- Ricoy Lorenzo, C. (2006). Contribución sobre los paradigmas de investigación. *Educação*, 31(1), 11-22.
- Riloff, E., & Wiebe, J. (2003). *Learning extraction patterns for subjective expressions*. 10, 105-112. <https://doi.org/10.3115/1119355.1119369>
- Ríos, B. (2018, ago). Moscovici: «Es el fin de un ciclo para Grecia pero también el fin de la crisis de la Eurozona». *El Mundo*. Recuperado de <https://www.elmundo.es/economia/2018/08/20/5b7aa9d346163f8c368b4587.html>
- Rivas Flores, J. I. (1990). *Investigación naturalista en educación: una revisión crítica*. Valencia: Promolibro.
- Robinson, P. C. (1980). *EPS (English for specific Purposes): the present position*. Oxford: Pergamon Press.
- Rocci, A., Palmieri, R., & Gautier, L. (2015). Introduction to thematic section on text and discourse analysis in financial communication. *Studies in Communication Sciences*, (1).
- Rondeau, G. (1983). *Introduction à la terminologie*. Quebec: Gaëtan Morin.



- Rosch, E., & Lloyd, B. B. (1978). *Cognition and categorization*.
- Rossi, M. (2015). *In rure alieno: métaphores et termes nomades dans les langues de spécialité*. Berna: Peter Lang.
- Roubini, N., & Rosa, B. (2018). The makings of a 2020 recession and financial crisis. *Project Syndicate: September, 13*.
- RTÉ News. (2014, febrero 14). Greek recession easing faster than expected. *RTÉ News*. Recuperado de <https://www.rte.ie/news/business/2014/0214/504345-greek-economy/>
- Ruppenhofer, J., & Rehbein, I. (2012). Yes we can!?! Annotating the senses of English modal verbs. *LREC 2012, Eighth International Conference on Language Resources and Evaluation*, 1538-1545. Paris: European Language Resources Association (ELRA).
- Sag, I. A., Baldwin, T., Bond, F., Copestake, A., & Flickinger, D. (2002). Multiword expressions: A pain in the neck for NLP. *Proceedings of the 3rd International Conference on Intelligent Text Processing and Computational Linguistics (CICLing-2002)*, pp. 1–15. Ciudad de México.
- Sager, J. C. (1990). *A Practical Course in Terminology Processing*. John Benjamins Publishing.
- Salas-Zarate, M. del P., Paredes-Valverde, M. A., Limon, J., Tlapa, D. A., & Báez, Y. A. (2016). Sentiment Classification of Spanish Reviews: An Approach based on Feature Selection and Machine Learning Methods. *J. UCS*, 22(5), 691–708.
- Sampson, G., & Bentham, Y. (1992). Probabilistic parsing. En J. Svartvik (Ed.), *Directions in corpus linguistics* (pp. 425–447). Berlin: Mouton de Gruyter.
- Samson, C. (2006). «Is Different From»: A Corpus-Based Study of Evaluative Adjectives in Economics Discourse. *IEEE transactions on professional communication*, 49(3), 236–245.
- Sanfilippo, A., y otros. (1999). *Preliminary Recommendations on Lexical Semantic Encoding-Final Report*. Recuperado de <http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=4212302885B116F8379074E38E1B8333?doi=10.1.1.27.5838&rep=rep1&type=pdf>
- Santamaría Pérez, I. (2009). *La terminología: definición, funciones y aplicaciones*. Alicante: Liceus, Servicios de Gestió.
- Sarkozy, N. (2010, noviembre) *Discours d'ouverture du Président Sarkozy à Davos*. Presentado en 40° Foro Económico Mundial, Davos. Recuperado de <https://uk.ambafrance.org/Discours-d-ouverture-du-President,16757>
- Saurí, R., & Pustejovsky, J. (2009). FactBank: a corpus annotated with event factuality. *Language resources and evaluation*, 43(3), 227.
- Sawyer, S. (2018). VI—The Importance of Concepts. *Proceedings of the Aristotelian Society*, 118(2), 127-147. <https://doi.org/10.1093/arisoc/aoy008>
- Shiller, R. J. (2000). *Irrational Exuberance*. Princeton: Princeton University Press.
- Schumaker, R. P., & Chen, H. (2009). Textual analysis of stock market prediction using breaking financial news: The AZFin text system. *ACM Transactions on Information Systems (TOIS)*, 27(2), 12.
- Sen, A. (1970). *Collective Choice and Social Welfare*. San Francisco: Holden Day.
- Serrano, P. (2010). *Traficantes de información: La historia oculta de los grupos de comunicación españoles*. Ediciones AKAL.
- Serrano, E. (2014). El 15M como medio: autoorganización y comunicación distribuida. *Calleja-López, A., Monterde, A. Serrano, E. y Toret, J.(Eds.)(2014). 15MP2P. Una mirada transdisciplinar del Movimiento 15-M, 15, 120–133.*

- Shambaugh, J., Reis, R., & Rey, H. (2012). The Euro's Three Crises. *Brookings Papers on Economic Activity*, 157.
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Sinclair, J. (1987). *Looking up: An account of the COBUILD project in lexical computing and the development of the Collins COBUILD English language dictionary*. Collins Elt.
- Slodzian, M. (2000). L'émergence d'une terminologie textuelle et le retour du sens. *Le sens en terminologie*, 61–85 Lyon: Presses universitaires de Lyon.
- Spillner, B. (2007). Discours économique: Variétés intralinguistiques et différences contrastives. En I. Behr, D. Hentschell, M. Kauffmann, & A. Kern (Eds.), *Langue, économie, entreprise : le travail des mots* (pp. 117–145). Paris: Presse Sorbonne Nouvelle.
- Stern, G. (1931). *Meaning and Change of Meaning*. Bloomington: Indiana University Press.
- Stewart, D. (2010). *Semantic prosody: A critical evaluation*. Londres: Routledge.
- Stiglitz, J. (2009). Crisis mundial, protección social y empleo. *Revista Internacional Del Trabajo*, 128(1-2), 1-15. <https://doi.org/10.1111/j.1564-9148.2009.00046.x>
- Stone, P. J., & Hunt, E. B. (1963). A computer approach to content analysis: studies using the General Inquirer system. *Proceedings of the May 21-23, 1963, Spring Joint Computer Conference*, 241–256. ACM.
- Streicher-Arseneault, V. (2010). Las políticas lingüísticas en Quebec y Cataluña: un reflejo de las dinámicas sociales. *Tinkuy*, 5, 25-40.
- Stubbs, M. (1995). Collocations and semantic profiles: On the cause of the trouble with quantitative studies. *Functions of language*, 2(1), 23–55.
- Stubbs, M. (1996). *Text and corpus analysis: Computer-assisted studies of language and culture*. Oxford, R.U.; Cambridge, EE.UU.: Blackwell Publishers.
- Stubbs, M. (2001). *Words and phrases: Corpus studies of lexical semantics*. Oxford: Blackwell Publishers.
- Swales, J. (1985). *Episodes in ESP: A source and reference book on the development of English for science and technology* (Vol. 1). Pergamon.
- Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-based methods for sentiment analysis. *Computational linguistics*, 37(2), 267–307.
- Taboada, M., & Grieve, J. (2004). Analyzing appraisal automatically. *Proceedings of AAAI Spring Symposium on Exploring Attitude and Affect in Text (AAAI Technical Report)*, Stanford University, CA, 158-161. AAAI Press.
- Tao, X., Zhou, X., Zhang, J., & Yong, J. (2016). Sentiment Analysis for Depression Detection on Social Networks. *Advanced Data Mining and Applications*, 807-810. [https://doi.org/10.1007/978-3-319-49586-6\\_59](https://doi.org/10.1007/978-3-319-49586-6_59)
- Teather, D. (2008, septiembre 29). Wall Street crisis spreads through Europe's banks. The Guardian. Recuperado de <http://www.theguardian.com/business/2008/sep/30/banking.europeanbanks>
- Temmerman, R. (2000). *Towards new ways of terminology description: The sociocognitive-approach*. Amsterdam: John Benjamins.
- Ten Hacken, P. (2010). The tension between definition and reality in terminology. *Proceedings of Euralex*.
- Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of finance*, 62(3), 1139–1168.
- Tetlock, P. C., Saar-Tsechansky, M., & Macskassy, S. (2008). More than words: Quantifying language to measure firms' fundamentals. *The Journal of Finance*, 63(3), 1437–1467.

- Thomson, E. A., & White, P. (2008). The news story as rhetoric: Linguistic approaches to the analysis of journalistic discourse. *Communicating Conflict: Multilingual Case Studies of the News Media*. Londres: Continuum, 1–24.
- Toret Medina, J. (2015). *Tecnopolítica y 15M: La potencia de las multitudes conectadas. Un estudio sobre la gestación y explosión del 15 M*. Barcelona: Editorial UOC.
- Traugott, E. C., & Dasher, R. B. (2001). *Regularity in Semantic Change*. Cambridge: Cambridge University Press.
- Tumitan, D., & Becker, K. (2014). Sentiment-based features for predicting election polls: a case study on the brazilian scenario. *Proceedings of the 2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)-Volume 02*, 126–133. IEEE Computer Society.
- Turney, P. D. (2002). Thumbs up or thumbs down?: Semantic orientation applied to unsupervised classification of reviews. *Proceedings of the 40th annual meeting on association for computational linguistics*, 417–424. Recuperado de <http://dl.acm.org/citation.cfm?id=1073153>
- Turney, P. D., & Littman, M. L. (2003). Measuring praise and criticism: Inference of semantic orientation from association. *ACM Transactions on Information Systems (TOIS)*, 21(4), 315–346.
- Ullman, S. (1957). *Principles of Semantics* (2ª edn.), Oxford: Blackwell.
- Ullman, S. (1962). *Semantics: An introduction to the science of meaning*, Oxford: Blackwell.
- Ullman, S. (1970). “Change of meaning” en Hungerford, H., Robinson, J & Sledd, J. *English linguistics*. Illinois: Scott, Foresman & Co.
- Valero, Ó. (2018, agosto 21). Grecia: Fin del rescate que salvó a un país y arruinó a una nación. *El Confidencial*. Recuperado de [https://www.elconfidencial.com/mundo/2018-08-20/grecia-tercer-rescate-austeridad\\_1606096/](https://www.elconfidencial.com/mundo/2018-08-20/grecia-tercer-rescate-austeridad_1606096/)
- Valiante, D. (2017). The ‘Visible Hand’ of the ECB’s first quantitative easing. *International Economics and Economic Policy*, 14(4), 601–624.
- Van de Kauter, M., Breesch, D., & Hoste, V. (2015). Fine-grained analysis of explicit and implicit sentiment in financial news articles. *Expert Systems with applications*, 42(11), 4999–5010.
- Velardi, P., Fabriani, P., & Missikoff, M. (2001). Using Text Processing Techniques to Automatically Enrich a Domain Ontology. *Proceedings of the International Conference on Formal Ontology in Information Systems - Volume 2001*, 270–284. <https://doi.org/10.1145/505168.505194>
- Verdú, J. P. (2011). El Movimiento 15-M. Un nuevo actor sociopolítico frente a la “dictadura de los mercados”. *Sociedad y Utopía. Revista de Ciencias Sociales*, (38), 109–124.
- Vincent, B. (2011, octubre 14). Biggest cost of «Occupy Wall Street» protests: Semantics. *The Examiner*. Recuperado de <http://www.examiner.com/article/biggest-cost-of-occupy-wall-street-protests-semantics>
- Vincze, V., Szarvas, G., Farkas, R., Móra, G., & Csirik, J. (2008). The BioScope corpus: biomedical texts annotated for uncertainty, negation and their scopes. *BMC bioinformatics*, 9(11), S9.
- Vohra, S., & Teraiya, J. (2013). A comparative study of sentiment analysis techniques. *Journal JIKRCE*, 2(2), 313–317.
- Waloszek, A., & Waloszek, W. (2017). SACAM - A Model for Describing and Classifying Sentiment Analysis Methods: *Proceedings of the 9th International*

- Conference on Agents and Artificial Intelligence*, 196-206. <https://doi.org/10.5220/0006199901960206>
- Wang, H., Can, D., Kazemzadeh, A., Bar, F., & Narayanan, S. (2012). A system for real-time twitter sentiment analysis of 2012 us presidential election cycle. *Proceedings of the ACL 2012 System Demonstrations*, 115–120. Association for Computational Linguistics.
- Weissenhofer, P. (1995). *Conceptology in terminology theory, semantics and word-formation: A morpho-conceptually based approach to classification as exemplified by the English baseball terminology*. Viena: TermNet.
- Wen, M., & Wu, Y. (2011). Mining the Sentiment expectation of nouns using Bootstrapping method. *Proceedings of 5th International Joint Conference on Natural Language Processing*, 1423–1427.
- White, P.R.R. (2003). Un recorrido por la teoría de la valoración (Teoría de la valoración). Traducido por Elsa Ghio. Recuperado de [www.grammatics.com/appraisal/SpanishTranslation-AppraisalOutline.doc](http://www.grammatics.com/appraisal/SpanishTranslation-AppraisalOutline.doc)
- White, P. R. R. (2015). Appraisal Theory. En *The International Encyclopedia of Language and Social Interaction*. <https://doi.org/10.1002/9781118611463.wbielsi041>
- Whitelaw, C., Garg, N., & Argamon, S. (2005). Using appraisal groups for sentiment analysis. *Proceedings of the 14th ACM international conference on Information and knowledge management*, 625–631. ACM.
- Wiebe, J. M. (2000). Learning subjective adjectives from corpora. *AAAI/IAAI*, 20(0), 0.
- Wiebe, J. M. (1994). Tracking point of view in narrative. *Computational Linguistics*, 20(2), 233–287.
- Wiebe, J. M., Bruce, R. F., & O'Hara, T. P. (1999). Development and use of a gold-standard data set for subjectivity classifications. *Proceedings of the 37th annual meeting of the Association for Computational Linguistics on Computational Linguistics*, 246–253. Association for Computational Linguistics.
- Wiggins, R., Piontek, T., & Metrick, A. (2014). The Lehman brothers bankruptcy a: overview. *Yale Program on Financial Stability Case Study*.
- Wilson, T., Wiebe, J., & Hoffmann, P. (2005). *Recognizing contextual polarity in phrase-level sentiment analysis*. 347-354. <https://doi.org/10.3115/1220575.1220619>
- Wright, S. E., & Budin, G. (1997). *Handbook of Terminology Management: Basic aspects of terminology management*. Amsterdam: John Benjamins Publishing.
- Wüster, E. (1998). *Introducción a la teoría general de la terminología y a la lexicografía terminológica*. Barcelona: Inst. Univ. de Lingüística Aplicada.
- Xiao, R., & Hu, X. (2015). Corpus-Based Translation Studies: An Evolving Paradigm. En *Corpus-Based Studies of Translational Chinese in English-Chinese Translation*. Springer.
- Ynduráin Pardo de Santayana, C. (2008). Individuo. En J. A. Pascual (Ed.), *Nomen exempli et exemplum vitae: studia in honorem sapientissimi Iohannis Didaci Atauriensis* (pp. 263–272). Madrid: Sesgo.
- Yu, Y., Duan, W., & Cao, Q. (2013). The impact of social and conventional media on firm equity value: A sentiment analysis approach. *Decision Support Systems*, 55(4), 919–926.
- Zamora Maass, M. L. (2016). *Use of semantic change for sentiment analysis*. Seminar in Semantics course: Artificial Neural Networks. New York University.
- Zhang, W., Xu, H., & Wan, W. (2012). Weakness Finder: Find product weakness from Chinese reviews by using aspects based sentiment analysis. *Expert Systems with Applications*, 39(11), 10283–10291.

- 
- Zhang, Y., Lai, G., Zhang, M., Zhang, Y., Liu, Y., & Ma, S. (2014). Explicit factor models for explainable recommendation based on phrase-level sentiment analysis. *Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval*, 83–92. ACM.
- Zhao, Z., & Ahmad, K. (2015). Qualitative and Quantitative Sentiment Proxies: Interaction Between Markets. En K. Jackowski, R. Burduk, K. Walkowiak, M. Wozniak, & H. Yin (Eds.), *Intelligent Data Engineering and Automated Learning – IDEAL 2015* (Vol. 9375, pp. 466-474). [https://doi.org/10.1007/978-3-319-24834-9\\_54](https://doi.org/10.1007/978-3-319-24834-9_54)
- Žižek, S. (2012). The Spectre of Ideology. En S. Žižek (Ed.), *Mapping ideology*. Londres / Nueva York: Verso.



## Anexo 1. Resumen de la tesis

### El lenguaje económico en los tiempos de la crisis global: un estudio longitudinal de análisis de sentimiento

Javier Fernández Cruz

Programa de doctorado en Lingüística, Literatura y Traducción

Universidad de Málaga

Tesis dirigida por Dr. Antonio Moreno-Ortiz

¿Han influido las consecuencias de la crisis económica mundial en la forma en que usamos el lenguaje? La variación en la orientación semántica es indudablemente difícil de rastrear exclusivamente a través de la introspección. Tal observación se ve ahora facilitada por la disponibilidad actual de grandes cantidades de texto y el desarrollo de soluciones de procesamiento del lenguaje natural. Se parte de que existe una serie de *palabras evento* en el dominio de la economía y las finanzas cuyo sentimiento varió rápidamente tras el estallido de la crisis global en el año 2008. Los lenguajes especializados tienen una terminología, expresiones y construcciones propias que distan de ser inalterables, y que además reflejan el espíritu de los tiempos. Todos estos factores salen a la luz en la manera en la que se expresan las evaluaciones y opiniones en los textos. Para ello este trabajo analiza longitudinalmente el sentimiento del contexto de una serie de términos en un corpus de prensa económica durante el intervalo 2007-2015.

El análisis de sentimiento se ocupa del tratamiento computacional de la opinión y la subjetividad en los textos. A través de distintos enfoques y algoritmos es posible calcular con gran precisión la polaridad u orientación semántica de elementos léxicos, oraciones y textos. Básicamente, una herramienta de análisis de sentimiento clasifica un texto como positivo o negativo, a través de una clasificación binaria o una puntuación dentro de una escala. El análisis de sentimiento depende profundamente del contexto y, además, del dominio específico en cuestión. Por ello, es necesario recopilar un lexicón *plug-in* en lengua inglesa para el dominio de las noticias económicas con alta precisión y exhaustividad.

El objetivo de esta tesis es el siguiente:

**Objetivo general.** Analizar las fluctuaciones en la orientación semántica de determinados términos económicos dentro del período 2007-2015 a través de la caracterización del impacto de los eventos de mayor orden en las variaciones semánticas de las unidades léxicas.

Dados el objetivo general antes mencionado, los siguientes objetivos específicos están en orden:

**OE1.** Determinar un lexicón de sentimiento de dominio específico en lengua inglesa para el dominio de las noticias económicas de alta precisión y exhaustividad a partir de un corpus de texto especializado.

**OE2.** Definir un subconjunto de datos longitudinal en forma de oraciones que contienen los términos de estudio y que serán el input del análisis de sentimiento.

**OE3.** Para los términos económicos-financieros objetos de estudio, identificar los factores (sociales, políticos, económicos, etc.) que han acompañado a cambios en el uso, desde la perspectiva de sentimiento, durante el período 2007-2015.

**OE4.** Analizar cualitativamente las posibles variaciones en la prosodia semántica de los términos de estudio a través de los años.

El segundo capítulo plantea un marco contextualizador de la tesis que permite mejorar la comprensión de los factores sociales que influyeron en el planteamiento del estudio. Se divide en tres partes: una contextualización histórica, una contextualización de la influencia de la crisis y la prensa en el uso del lenguaje económico y, por último, una breve exposición sobre la influencia y posibilidades de los contenidos en red y la opinión en la actualidad. La primera parte revisa cronológicamente los eventos principales de la crisis global: desde el estallido de la burbuja inmobiliaria que dio lugar en 2008 a la crisis de las hipotecas *subprime* y la crisis crediticia. Posteriormente, se analiza la crisis de la deuda, centrada principalmente en la Zona Euro. Posteriormente examina el impacto de las consecuencias de la *Gran Recesión* en la vida cotidiana de la ciudadanía. La segunda parte revisa la influencia de la prensa como mediadora del lenguaje económico durante la crisis. Se exploran los motivos por los que los términos financieros y económicos se pusieron en primer plano y se debate sobre los usos ideológicos del lenguaje en relación con los posibles cambios en la actitud del público hacia el discurso económico.

Los capítulos 3, 4 y 5 constituyen el marco teórico de este trabajo, y revisan el estado del arte de los tres temas principales de la tesis: el lenguaje evaluativo, los lenguajes especializados y el análisis de sentimiento.

Así, el tercer capítulo recopila los principales enfoques teóricos que permiten acercarse a un entendimiento estructurado del funcionamiento del lenguaje evaluativo. Tras una caracterización de la evaluación en distintas disciplinas, se realiza un recorrido por las últimas cuatro décadas de desarrollo teórico, el cual coincide con el desarrollo de la lingüística de corpus y el cambio de paradigma textual de la lingüística. La evaluación se entiende como una función realizada por el texto a través de una acción explícita o implícita a través del uso de una amplia variedad de recursos lingüísticos. El estudio de la evaluación, afecto o valoración, entre otros términos, es una cuestión multidisciplinar que implica, además del estudio del lenguaje, conocimientos de psicología o sociología. La relevancia que ha tomado este campo de trabajo es indiscutible, ya que las teorías de la evaluación han servido de base teórica de cualquier estudio computacional de la opinión. Además, se realiza un recorrido general del desarrollo de la teoría de valoración de Martin y White (2005), el cual cubre de manera altamente estructurada todos los aspectos relacionados con el uso evaluativo del lenguaje. Desde una perspectiva predominantemente interpersonal, se observan las acciones dentro del discurso: expresar actitud, posicionamiento y alineamiento con respecto a una audiencia determinada o un interlocutor, así como sus distintos niveles de gradación.

El discurso y los fenómenos sintácticos, textuales e intertextuales juegan un rol crucial en la resolución del problema del análisis de sentimiento. Por ello, los distintos factores estudiados deben incorporarse a los sistemas de análisis de sentimiento. Ya que el análisis de sentimiento no es necesariamente una meta en sí misma, sino una herramienta que permite realizar estudios de mayor calibre, se considera aquí que el estudio textual de la evaluación puede combinarse con distintos métodos estadísticos. Así se logra obtener una mayor cantidad de datos que aporten nuevas perspectivas para analizar la realidad discursiva de una manera más eficiente.

El cuarto capítulo propone una panorámica de los lenguajes especializados enfocada al tratamiento sistematizado de la terminología. Primero se revisaron las fronteras, muchas veces difusas, entre el dominio general y las lenguas de especialidad. Si bien existen consideraciones que imponen mayores distancias entre dominios, desde hace dos décadas existe un consenso en que la diferencia está, principalmente, en los usos comunicativos en el campo de trabajo de cada disciplina.

La terminología sigue siendo un factor muy importante en la discusión, pero esta debe tenerse en cuenta como un factor más del estudio, junto con factores textuales y contextuales. Por ello se revisan los modelos teóricos más importantes. El modelo terminológico propuesto por la *teoría general de la terminología* de la Escuela de Viena propone un punto de vista conceptual jerárquico en el que las definiciones son altamente normativas, unívocas y sincrónicas: solo existirá una denominación para cada concepto y su definición será eminentemente neutral, sin ninguna consideración a su evolución a lo largo de los tiempos. El significado evaluativo o cualquier relación con el sentimiento quedan fuera del ámbito de estudio. Surgieron nuevos modelos en los albores del siglo XXI, destacando aquí la teoría comunicativa de Cabré (1998) y la socioterminología de Temmerman (2000), que supusieron un giro copernicano en el tratamiento de lenguajes especializados. Estos trabajos ofrecen nuevas posibilidades teóricas en los sistemas de transmisión del conocimiento y se anticipan a las distintas aplicaciones en el terreno del procesamiento del lenguaje natural. Según estos enfoques las unidades terminológicas compartirían propiedades lingüísticas con el léxico general, diferenciándose en el modo de significación y en sus aspectos pragmáticos, sociales y cognitivos dependientes ya no exclusivamente de la ontología sino del discurso. Esto es interesante para este trabajo ya que se considera que los términos contienen más significados que sus denominaciones (p.ej., significado afectivo) y que este no es sincrónico. Se considera entonces la posibilidad de que el sentimiento que evoque un término tendrá relación con el momento histórico y con el uso que tenga dentro del dominio.

Con el fin de aportar una infraestructura teórica sólida a la construcción del lexicón, se han revisado los puntos más importantes de la literatura sobre la terminografía. La terminografía se ocupa exclusivamente de compilar colecciones estructuradas de los términos pertenecientes a los lenguajes especializados. El desarrollo del tesauro, las ontologías y sus técnicas relacionadas han trascendido el dominio lingüístico en cuanto a la capacidad de estructuración compleja de la lengua y el conocimiento. Esto no sería posible sin el salto de gigante que ha supuesto la universalización de la informática en el tratamiento terminológico a partir de los años 90 y la revolución que supone en la actualidad la aplicación de las tecnologías relacionadas con el aprendizaje automático. Por último, se trata la descripción de los factores subyacentes correspondientes al léxico, organización textual y contextual de los dominios específicos.

En el quinto capítulo se revisa exhaustivamente el desarrollo del análisis de sentimiento como disciplina, así como sus principales enfoques de trabajo. Existen tres niveles de análisis de sentimiento según la entrada (documento, oración y entidad) que son enfocados de manera distinta durante su proceso de asignación automática de orientación semántica. En cuanto a la granularidad de los resultados, existen dos enfoques principales: el primero llamado *Thumbs up/Thumbs down* en la literatura que ofrece resultados binarios o ternarios (positivo, negativo y, dado el caso, neutro) y el segundo tipo de sistemas ofrece resultados positivos dentro de una escala (p. ej., de 0 a 10). La tercera clasificación de los sistemas es el método de análisis. Aquí concurren dos enfoques principales: uno basado en el *aprendizaje automático* basado en modelos estadísticos, que, en el caso del aprendizaje supervisado, depende de conjuntos de datos de entrenamiento etiquetados previamente y el rigor de sus resultados dependerá, entre otros factores, del volumen de estos. El *enfoque basado en léxico*, adoptado en este trabajo, se fundamenta en recopilaciones lexicográficas de entradas y reglas contextuales etiquetadas según el sentimiento que denotan. Es relativamente frecuente encontrar sistemas *híbridos* de análisis que utilizan técnicas de ambos enfoques.

No existe una solución definitiva para el problema del análisis de sentimiento. Los sistemas de análisis de sentimiento son, en gran parte, sistemas diseñados para resolver



problemas ad-hoc y no hay formas universales de evaluar el rendimiento de estos mediante una comparación universal entre los distintos trabajos. Es por esto que distintos enfoques ofrecerán mejores o peores resultados según la información de entrada. Sin embargo, desde una perspectiva lingüística, los enfoques basados en léxico permitirán sacar mejor partido al estudio de las distintas teorías sobre el lenguaje evaluativo. Gracias a estos, es posible implementar aplicaciones prácticas y estudiarlas desde una nueva perspectiva.

El sentimiento es especialmente sensible al dominio, pues las valencias del análisis variarán dependiendo del tipo de lenguaje especializado (o no) que se use. Por ello, se aborda este problema a partir del estudio del dominio específico económico-financiero y se revisan los principales trabajos realizados sobre el dominio. Se trata de un viaje de ida y vuelta entre dos disciplinas pues, en relación con este apartado, se observarán las ventajas y desventajas de la aplicación del enfoque léxico en este dominio, un trabajo eminentemente terminográfico. Además, se analizan los modelos teóricos en que los significados experimentan cambios en sus connotaciones y resultar el germen de nuevos sentidos de las palabras de manera que estos adquieren nuevos aspectos de significado y otros caen en desuso. El eje *mejora/peyoración* dentro de los estudios sobre el cambio semántico es fundamental para este trabajo. En estos procesos de cambio, el sentido de una palabra cambia para volverse más positivo o negativo, respectivamente a lo largo de un intervalo de tiempo determinado.

En el quinto capítulo se describen la metodología, las fuentes de datos y las herramientas necesarias para cumplir con los objetivos establecidos. Está dividido en cinco partes: una revisión de los paradigmas y enfoques metodológicos relacionados con el trabajo, el diseño metodológico de la tesis, el diseño y construcción de un corpus de dominio específico, la descripción de las herramientas externas utilizadas y la construcción de *LexiEcon*, un lexicón de dominio específico en lengua inglesa para *Lingmotif*.

Los dos primeros apartados del capítulo describen un diseño metodológico mixto en cuatro pasos que permite extraer y clasificar la información sobre lenguajes especializados desde una perspectiva que generalmente no recibe mucha atención en la literatura: la expresión de la opinión y las emociones, un aspecto del lenguaje que, como parece evidente, empapa la totalidad del uso del lenguaje. El diseño metodológico se puede resumir en cuatro pasos:

1. Recopilación de un lexicón de dominio específico para el software de análisis de sentimiento:
  - 1.1. Diseño y construcción de un corpus de dominio específico.
  - 1.2. Extracción terminológica a través de distintas técnicas de lingüística de corpus.
2. Mediante series temporales, analizar la evolución longitudinal del sentimiento de una serie términos encontrados en oraciones procedentes del corpus
3. Determinar los factores extralingüísticos (sociales, políticos, económicos, etc.) que influyen los cambios.
4. Interpretar cualitativamente, con detalle, las variaciones sociolingüísticas diacrónicas de las condiciones de uso de los términos seleccionados (frecuencia, eventos, colocaciones, usos del lenguaje).

La tercera parte del capítulo describe el diseño y recopilación de la fuente de datos: los textos que formaron el *Esmeraldas Corpus of Business English* de +24M de *tokens*. El corpus está recopilado según los principios de equilibrio y representatividad tradicionales de la disciplina de lingüística de corpus. La colección de textos clasifica de manera cronológica todos los textos de la sección *Business* de los periódicos The Guardian y New

York Times a lo largo de los años en los que tuvo lugar la crisis (2007 y 2015). Para su extracción se utilizaron *scrapers* web, y los textos fueron clasificados de manera semiautomática. El producto final fue procesado y empaquetado en *The Sketch Engine* y está disponible bajo demanda.

La siguiente sección describe las dos herramientas externas utilizadas para cumplir con los objetivos: la suite de corpus *The Sketch Engine* y la aplicación de análisis de sentimiento *Lingmotif*. La primera es un administrador de corpus y un software de análisis de texto *online* que ofrece múltiples posibilidades (concordancia, extractor de terminología, *word sketches*, entre otras) y además ofrece una gran variedad de corpus de múltiples dominios y lenguas. Por otro lado, *Lingmotif* es una aplicación que analiza textos desde la perspectiva del análisis de sentimiento. En resumidas cuentas, permite determinar la orientación semántica (si es positivo o negativo y en qué grado) de un texto mediante la detección de expresiones lingüísticas que indican una determinada polaridad. Incluye un lexicón muy detallado en inglés y en español y, a facilita el análisis de sentimiento en lenguajes especializados.

En la quinta parte del capítulo lleva a cabo el primer objetivo específico, describiéndose la recopilación de un lexicón de sentimiento de dominio económico-financiero de gran tamaño ajustado a las especificaciones de *Lingmotif*. El lexicón se recopiló de acuerdo con la base teórica revisada. Para ello, se combinaron técnicas cuantitativas y cualitativas de extracción terminológica. Los candidatos a término se obtuvieron principalmente a partir de técnicas de extracción basada en corpus en las que, a través de la herramienta *KeyWords* de *The Sketch Engine*, se extrajeron del ECBE 30.000 candidatos a término. Con la ayuda de un experto en el dominio, cada uno de los candidatos fue comparado con los valores de las entradas del lexicón de sentimiento de dominio general. Se seleccionaron exclusivamente aquellas entradas cuyo sentimiento en el dominio específico difiriera del léxico base y a cada una de ellas se le asignó la valencia correspondiente al dominio específico.

Como resultado se presentó LexiEcon, un lexicón plug-in de dominio específico para la lengua inglesa que en la actualidad incluye 6.843 entradas. Los resultados de su evaluación fueron muy satisfactorios. Tras analizar un *gold standard* que contiene 400 oraciones aleatorias de dominio específico extraídas del corpus ECBE a través de dos barridos (primero aplicando tan solo con el lexicón general y, posteriormente, en conjunto con el *plug-in*), se observan excelentes niveles de precisión en el dominio. Las cifras de rendimiento son notables pues, utilizándose en conjunto con el lexicón de *Lingmotif*, obtiene un F1 medio de 0,735. Esta cifra supone alrededor de un 20% más que los resultados que ofrece *Lingmotif* sin plug-in cuando clasifica los textos del dominio económico-financiero.

La naturaleza especializada de los textos facilita el análisis de sentimiento basado en léxico y ofrece mejores resultados que en el dominio general. A pesar de que el uso de los términos no es íntegramente unívoco, los buenos resultados se justifican por la naturaleza del lenguaje especializado ya que tiende a un bajo nivel de ambigüedad. Si bien en el discurso de la prensa semiespecializada se da una falta de precisión estricta, estos niveles de precisión en el uso del lenguaje son mucho mayores en comparación con el uso del lenguaje en el dominio general (p. ej., el uso de ironías o anáforas). Además, la redacción de los artículos económicos favorece la claridad de las enunciaciones con el fin de facilitar al público la comprensión de una información con un alto nivel de abstracción.

En el Capítulo 6 se realiza el análisis de sentimiento de diez palabras que fueron clave durante la crisis: “credit”, “debt”, “inflation”, “housing”, “markets”, “investor”, “public”, “reform”, “spending” y “company”. Para ello se extrajeron del corpus las oraciones que incluían los términos clave seleccionados y se procesaron con la suite de

análisis de sentimiento *Lingmotif* aplicando el *plug-in* de dominio específico *LexiEcon*. El análisis consta de tres partes: (a) una tabla de resultados estadísticos descriptivos longitudinales relativos a las puntuaciones de sentimiento, (b) una tabla anual de colocaciones y (c) una discusión sobre los hallazgos en el corpus a partir de la observación de rankings anuales de colocaciones, con la intención de triangular los datos obtenidos.

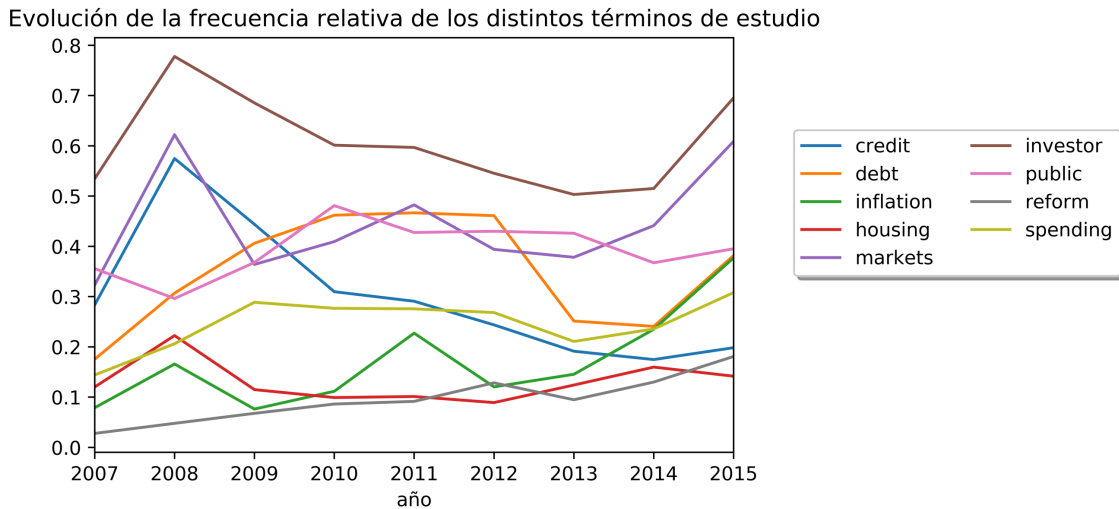
El Capítulo 7 resume los principales resultados tras el procesado y posterior análisis, y discute estos resultados en frente al marco teórico. Se destacan dos hechos:

1. Los términos se convierten en *palabras evento* dado el enorme aumento de su frecuencia de uso debido a los eventos clave de la crisis. Generalmente, a partir de este aumento se producen cambios significativos en el uso y el tipo de colocaciones varía, entrando en juego nuevas unidades léxicas con una orientación semántica diferente y que, en general, suelen ser menos especializadas que las fijaciones tradicionales relacionadas con términos.
2. Las medias anuales de la orientación semántica de un término contextualizado permiten observar con claridad fluctuaciones importantes en el sentimiento embebido en el discurso. Una triangulación de los datos cuantitativos con sus colocaciones más significativas y los eventos relacionados con la *Gran Recesión* permite concluir que la orientación semántica de los términos del dominio económico-financiero es muy susceptible de variar a medida que se desarrollaron los hechos de la crisis financiera.

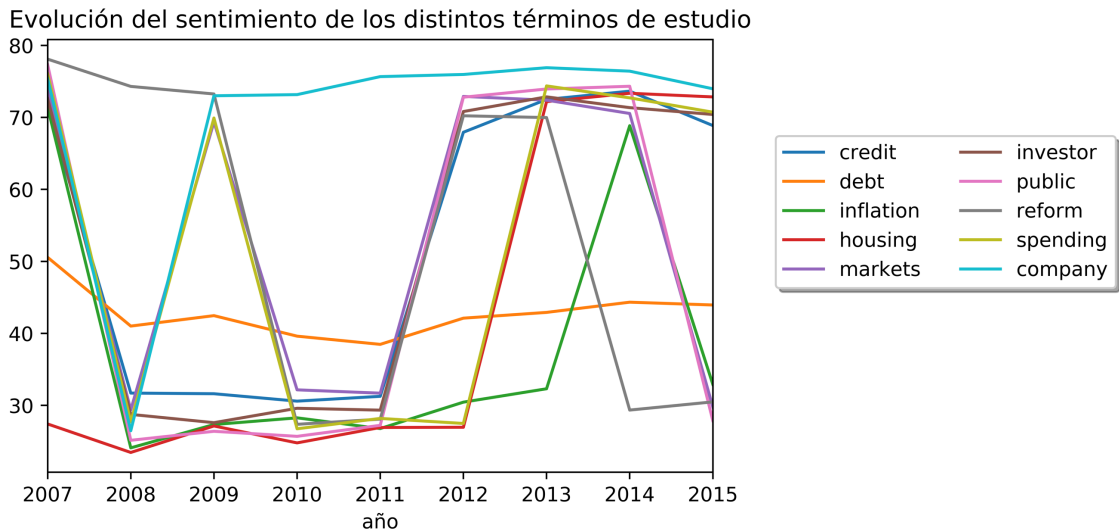
El periodismo económico es un discurso mediado, en el que los profesionales interactúan con diversas categorías de expertos (políticos, economistas, usuarios, etc.). La cadena de transmisión experto-periodista-público pone sobre la mesa la relevancia y el uso de las unidades léxicas especializadas, convirtiendo a los términos en eventos. En perspectiva, es esta circulación la que hace que un hecho mundano se transmute en un evento, asignándole un significado social memorable a sus términos destacados, con unas connotaciones determinadas a lo largo de un intervalo de tiempo.

La tarea de análisis realizada trata de identificar brechas en los usos lingüísticos que constituyen los eventos y establecer las causas. La observación cuantitativa de la frecuencia relativa del uso de los distintos términos (Figura 1) permite observar que en la mayoría de los casos el estallido de la crisis en el año 2008 fue muy relevante, generándose un pico en el uso de los términos en cuestión. En muchos casos, además, esto ha ido acompañado de un cambio radical en sus connotaciones de orientación semántica.

En la mayoría de los casos se pasa de una media positiva de sentimiento previa a la crisis (año 2007) a una caída de las medias al umbral de la negatividad, como se puede observar en la Figura 2. La separación de la crisis del crédito de la crisis de la deuda es también evidente, tanto en las frecuencias como en el sentimiento, y se observan patrones similares en la mayoría de los casos. De hecho, en el año 2009 se produce un aumento puntual en la media de los datos de sentimiento para, posteriormente, volver a caer en el año 2010 al aumentar la deuda y, consecuentemente, el deterioro en las condiciones de vida de la ciudadanía. Posiblemente a partir de una mejora leve en las condiciones macroeconómicas de los países del norte de la Zona Euro y en gran medida con el espíritu y giro discursivo que supuso el lanzamiento de políticas de alivio económico, en especial la *expansión cuantitativa* por el Banco Central Europeo iconizada en el famoso discurso “Whatever it takes” de su presidente, Mario Draghi, se observó un aumento significativo en las puntuaciones de sentimiento, que en la mayoría de los casos se prolongó desde 2011 o 2012 hasta el final del estudio.



**Figura 1.** Evolución de la frecuencia relativa de los distintos términos de estudio



**Figura 2** Evolución del sentimiento de los distintos términos de estudio

A partir de lo observado en las colocaciones de los términos se constata que la prosodia semántica es una propiedad del léxico que no es estática, sino variable en el tiempo. De acuerdo con Bublitz (1996) o Morley et al. (2009), se evidencia que un conjunto de colocaciones evoluciona para adquirir un matiz de orientación semántica positiva o negativa lo largo de un intervalo temporal. Los cambios cuantitativos en el sentimiento surgidos a partir de la crisis se correlacionan con la aparición de colocaciones con sentimiento positivo o negativo según la fluctuación de las medias de sentimiento de los contextos de los términos y coincidiendo con los eventos económicos. Para ilustrar esto, en el caso de “public”, en plena crisis del crédito se observa que la mayoría de las oraciones engloban un sentimiento negativo. Además, las colocaciones que se encuentran en el vecindario de los distintos conjuntos también suelen serlo. Por ejemplo, el término coocurre puntualmente con el sustantivo “cuts”, y a su vez aparecen frecuentemente colocaciones que mantienen la armonía del sentimiento, por ejemplo: “rig”, “conspiring” o “miserable”. Esto evidencia el uso de los términos en contextos más negativos. Se produce a todos los efectos un “contagio” inmediato del significado pragmático o actitudinal del término a causa de un evento repentino y violento como es la crisis, que permaneció constante mientras duraban sus efectos. Como consecuencia, las unidades



léxicas tienden a adquirir connotaciones más positivas o negativas debido a que su prosodia semántica varía a lo largo de un intervalo temporal determinado.

A nivel teórico, se pueden extraer tres conclusiones principales:

1. Los resultados del análisis de rendimiento de *LexiEcon* y el posterior análisis de los textos evidencian las enormes posibilidades que ofrece el estudio de los lenguajes especializados desde una perspectiva escasamente explorada como es la expresión del sentimiento.
2. Existen términos del dominio económico-financiero que adquieren gran relevancia durante eventos de gran calado social, como es el caso de la Gran Recesión. En relación con estos términos, se pudo observar que este sentimiento inherente no es en ninguno de los casos un valor estático, sino que en mayor o menor medida este varía con las condiciones socioeconómicas.
3. La observación de las puntuaciones de sentimiento, los rankings de colocaciones y su evolución a lo largo del intervalo de estudio no es sino una puesta en escena de los planteamientos de la lingüística sistémica donde la lengua es un sistema de elecciones por el cual el hablante elige de forma explícita qué expresión utilizará en detrimento de sus alternativas. De hecho, el buen rendimiento de *LexiEcon* en el análisis de los textos muestra que el lenguaje especializado no es una excepción y está organizado de una forma no arbitraria y enteramente funcional, en consonancia con su uso. Desde este punto de vista, de acuerdo con la perspectiva sistémico-funcional, el lenguaje es un recurso semiótico social que permite a los hablantes opciones para construir significados, destacando aquí la expresión de la evaluación que no escapa de la gramática.

A través del análisis de los términos, se ha podido evidenciar que su uso evaluativo está presente en todos los niveles lingüísticos. De hecho, el uso evaluativo no es ninguna excepción para el lenguaje especializado. Su puesta en escena cumple con la definición de evaluación vista en el Capítulo 3 pues desde este dominio es posible la expresión estructurada de los sentimientos sobre las entidades de las que habla, con vistas a influir a una audiencia de forma que se refleja el conjunto de valores de la comunidad de hablantes.

Es necesario resaltar que uso divulgativo del lenguaje por parte de la prensa exige niveles menores de especialización que la precisión requerida por los textos dirigidos a expertos (p. ej., en divulgaciones corporativas). Por otro lado, se observa que el aspecto evaluativo juega un papel fundamental. A través de la opinión y la discusión se pueden expresar percepciones subjetivas sobre el estado de la actualidad económica y aún así cumplir con las condiciones suficientes y necesarias para describir la realidad de la crisis sin perder rigor, y teniendo en cuenta la dimensión dialógica que señale las fricciones entre distintos marcos teóricos e ideológicos.

La observación de las colocaciones, cotexto y contexto de los términos evidencia fluctuaciones en el sentimiento de manera casi inmediata como reacción a los eventos económicos. A partir de una perspectiva computacional con base lingüística, la observación del sentimiento de las unidades léxicas especializadas permite también avanzar desde una terminología conceptual eminentemente estructural a la consideración de estas como unidades también comunicativas y emocionales. De acuerdo con las nociones contemporáneas, se propone realizar algunas aportaciones a los planteamientos clásicos de la terminología, de forma que se supere la visión aséptica y estática del término para desplazarse a un conjunto teórico que refleje una naturaleza dinámica y poliédrica de la terminología que incluya significados que vayan más allá de la linealidad concepto/definición. Los resultados coinciden con las aportaciones críticas con la



consideración de los términos como compartimentos estancos de carácter síncrono (p. ej., Temmerman, 2000), que no tienen en cuenta ningún tipo de evolución del término en el transcurso del tiempo.

El uso cotidiano de la terminología no es del todo *unívoco*. De hecho, la puesta en uso de los términos corresponde a dos factores principales en el ámbito del dominio. Por un lado, la estricta y necesaria expresión científico-técnica en el dominio económico y, especialmente aquí, los usos prototípicos que emergen del discurso divulgativo cotidiano del periodista especializado como mediador entre el experto y el público. Por poner un ejemplo relacionado con el uso, este sería el caso de “bad debt” de forma que quedase clara la naturaleza negativa de la deuda. A través de ejemplos como estos se confirma que el aspecto evaluativo juega un papel fundamental en este tipo de discurso.

El Capítulo 9 expone las conclusiones y se divide en dos apartados. Revisa la consecución del objetivo principal y los cuatro objetivos específicos propuestos y da respuesta de manera concisa a las preguntas realizadas en la introducción de la tesis. Esta labor se realiza como recopilación de las discusiones, reflexiones y hallazgos adelantadas a lo largo de las conclusiones de cada uno de los capítulos de esta tesis. De esta forma, se pretende dar cierre al trabajo realizado y abrir la puerta a las nuevas preguntas que emergen a partir de aquí. Los objetivos marcados para la presente tesis se cumplen ampliamente. Se considera que el desarrollo metodológico longitudinal cumple con su doble propósito de describir patrones de cambio y explicarlos a través de su relación con los eventos sociopolíticos de la crisis.

La última sección describe los límites y nuevos desafíos que ofrece el fin de esta investigación. Estos se pueden dividir en tres campos: ampliación de fuentes de datos, ampliación del objeto de estudio y ampliación de la cuestión técnica. En cuanto al primer aspecto, se considera que el diseño de este trabajo ha contemplado el corpus creado como un bloque único. Sin embargo, en la elaboración del corpus se ha contemplado la clasificación por publicación. Por un lado, sería interesante el estudio comparativo del uso del lenguaje evaluativo entre *The Guardian* y *The New York Times* y descubrir los factores lingüísticos, periodísticos, sociales e ideológicos que influyen en las posibles diferencias. Por otra parte, se podría repetir el experimento y contrastar los resultados con periódicos de distinta ideología.

Por otra parte, está la cuestión del estudio de la expresión del lenguaje económico desde distintos niveles de especialización. Por ejemplo, a través del análisis de discusiones del público en blogs, comentarios en periódicos o redes sociales u observar las variaciones en el uso del lenguaje altamente especializado. De esta forma, se podrían observar las divergencias léxicas, cognitivas y afectivas entre distintos niveles de especialización.

Se plantean varios desafíos en torno al desarrollo técnico de soluciones para el análisis de sentimiento y las aplicaciones prácticas de *LexiEcon*. Por ejemplo, adaptar el sistema de trabajo para la lengua castellana de forma que esta solución sea multilingüe. Finalmente, se sugiere realizar ensayos en el estudio de la relación entre el discurso y los rendimientos económicos con vistas a establecer una metodología de trabajo que alcance a determinar el impacto del discurso en la economía para así realizar análisis y predicciones que beneficien al bienestar comunitario.

## Anexo 2. Thesis summary

### **Economic language in the times of the global crisis: a longitudinal sentiment analysis study**

Javier Fernández Cruz

Ph.D. Program in Linguistics, Literature and Translation

Universidad de Málaga

Thesis directed by Dr. Antonio Moreno-Ortiz

Have the consequences of the global economic crisis influenced the way we use language? The variation in semantic orientation is undoubtedly difficult to trace only through introspection, such observation is now facilitated by the current availability of large amounts of text and the development of natural language processing solutions. It is assumed that there is a series of event words in the field of economics and finance whose sentiment varied rapidly after the outbreak of the global crisis in 2008. Specialized languages contain terminology, expressions and own constructions that are far from being unalterable, and they also reflect the spirit of the times. All these factors emerge in the manner that evaluations and opinions are expressed in texts. This work analyzed longitudinally the sentiment of the context of a series of terms in an economic press corpus during the 2007-2015 period.

Sentiment analysis deals with the computational treatment of opinion and subjectivity in texts. Through different approaches and algorithms, it is possible to calculate with great precision the polarity or semantic orientation of lexical elements, sentences, and texts. Basically, a sentiment analysis tool classifies a text as positive or negative, through a binary classification or a score within a scale. Sentiment analysis is highly dependent on the specific domain in question. To do this, it is necessary to tune up the SA system for the requirements of different specialized languages.

This work aims to resolve a general objective and four specific objectives, as follows:

**General objective.** To analyze the fluctuations in the semantic orientation of certain economic terms within the period 2007-2015 through the characterization of the impact of higher order events on the semantic variations of the lexical units.

Given the aforementioned general objective, the following specific objectives are in order:

**SO1.** To build a specific domain sentiment lexicon in English for the domain of economic news of high precision and completeness from a specialized text corpus.

**SO2.** To define a subset of longitudinal data in the form of sentences that contain the terms of study that will be the input of the sentiment analysis.

**SO3.** For the economic-financial terms of study, identify the factors (social, political, economic, etc.) that have accompanied changes in the use, from the perspective of feeling, during the period 2007-2015.

**SO4.** To analyze qualitatively the possible variations in the semantic prosody of the study terms over the years.

The second chapter proposes a contextualizing framework of the thesis which aims to improve the understanding of the economic and social factors that influenced the period of the study. It is divided into three parts: a historical contextualization, a contextualization of the influence of the crisis and the press on the use of economic language and, finally, a brief presentation on the influence and possibilities of content on the web and opinion. The first part reviewed the main chronological events of the global



crisis: from the bursting of the real estate bubble that gave rise in 2008 to the subprime mortgage crisis and the credit crunch. As a consequence of the breakdown of the banking system, a second debt crisis emerged in 2010, which had severe social and political consequences in the Euro Zone. It is followed by a review of the impact of the consequences of the Great Recession on the daily life of citizens and how new social movements appeared around the world as a reaction to austerity and shock policies. The second part reviews the influence of the press as the mediator of the economic language during the crisis and the reasons financial and economic terms were highly exposed, along with a discussion on the ideological uses of language in relation to possible changes in public attitudes towards economic discourse. The third part deals with the importance of user-generated contents in the expression of opinion.

Chapters 3, 4 and 5 make up the theoretical framework and review the state-of-the-art of the three main themes of the thesis: evaluative language, specialized languages, and sentiment analysis.

The third chapter compiles the main theoretical foundations of the functioning of evaluative language. A characterization of evaluation in different disciplines is presented and along with a review of the last four decades of its theoretical development, which coincides with the development of corpus linguistics and the change of paradigm to textual linguistics. Evaluation is understood as a function performed by the text through an explicit or implicit action and the use of a wide variety of linguistic resources. The study of evaluation, stance or affect, among other terms, is a multidisciplinary issue that implies, besides the study of language, an understanding of psychology or sociology. The relevance gained by this fieldwork is indisputable since theories of evaluation have served as the theoretical basis of any computational study of opinion. In addition, there is a general overview of the development of Martin and White's (2005) Appraisal Theory which covers all aspects related to the evaluative use of language in a highly structured manner. From a predominantly interpersonal perspective, the actions within the discourse are observed: expressing attitude, engagement (positioning and alignment) to a given audience, as well as their different levels of gradation.

Discourse and the syntactic, textual and inter-textual phenomena play a crucial role in sentiment analysis. Since sentiment analysis is not necessarily a goal itself, but a tool that allows performing larger-scale tasks. As the textual study of evaluation can be combined with different statistical methods, it is possible to obtain a greater amount of data in order to provide new perspectives that allow analyzing the discursive reality in a more efficient way.

The fourth chapter proposes an overview of specialized languages focused on the systematic management of terminology. First, the fuzzy limits between the general domain and the specialized languages were revised. Although there are opinions that impose greater distances between domains, from the last two decades there is a consensus that considers that the main difference lies in the communicative uses in the field of each discipline.

While discussions on terminology are still on the agenda, many experts consider that they should just be taken into account as one more factor of analysis, along with textual and contextual factors. Subsequently, the most important theoretical models were reviewed. The terminological model proposed by the General Theory of Terminology of the Vienna School proposes a hierarchical conceptual point of view in which the definitions are highly normative, univocal and synchronous. According to this, there will only be one denomination for each concept and its definition will be eminently neutral, with no consideration of its evolution. Evaluative meaning or any relationship with sentiment is outside this theory. New models emerged at the dawn of the 21st century,

highlighting Cabré's Communicative Theory of Terminology (1998) and Temmerman's Socioterminology (2000) that represented a Copernican turn in the treatment of specialized languages. These systems offer new theoretical possibilities in the knowledge transmission systems and anticipate the different applications in the field of Natural Language Processing. While differing in the way of meaning and its pragmatic, social and cognitive aspects; terminological units share linguistic properties with the general lexicon. In addition, they are no longer exclusively dependent on ontology but on discourse. These approaches are interesting for this task as it is considered that the terms contain more meanings than their definition (e.g. affective meaning) and that these meanings are not synchronous. Now the possibility of terms containing sentiment meaning is considered, which can also be related to the historical momentum and to specialized language use.

In order to provide a solid theoretical framework for the construction of the lexicon, the most important aspects of terminography literature have been reviewed. Terminography deals exclusively with the compilation of structured collections of terms belonging to the specialized languages. The development of the thesaurus, the ontologies, and their related techniques have transcended the linguistic domain in terms of the capacity for complex structuring of language and knowledge. This would not be possible without the giant leap forward that has provided the universalization of information technology in the terminological management since the 1990s and the revolution that currently is taking place around the application of technologies related to machine learning. Finally, we describe the underlying factors corresponding to the lexical, textual and contextual features specific to the domains of journalism and economy in English.

In the fifth chapter, we reviewed the history of the development of sentiment analysis (SA) as a discipline, as well as its main work approaches. There are three levels of sentiment analysis according to the input (document, sentence, and entity) that are focused differently during their process of automatic assignment of semantic orientation. Regarding the granularity of the results, there are two main approaches: the first called *Thumbs up/Thumbs down* in the literature that offers binary or ternary results (positive, negative and, if necessary, neutral) and *Seeing stars*, the second type of systems offers positive results within a scale (e.g., from 0 to 10). The third classification of SA systems is the method of analysis. Here they attend two main approaches: one based on machine learning algorithms (supervised, semi-supervised, unsupervised and reinforced), which depend heavily on previously labeled training data sets. The rigor of their results will depend, among other factors, on the volume of these. The lexicon-based approach will be the one that will be used in this project and its operation is based on lexicographical compilations of entries and contextual rules labeled according to the sentiment they denote. A sentiment lexicon includes a compendium of lexical units that have been assigned sentiment. Manual compilation of the lexicons is a time-consuming task, this is the main reason why different terminographic techniques are used to extract relevant entries from a corpus. In addition, it is relatively common to find hybrid analysis systems that use techniques from both approaches.

There is no ultimate solution to the problem of sentiment analysis. Sentiment analysis systems are, in large part, systems designed to solve ad hoc problems and there are no universal methods to test their performance by means of a universal comparison between the different tasks. This is why some approaches will have better or worse performance according to the input information. However, from a linguistic perspective, lexicon-based approaches will make possible to take better advantage of the study of the different theories about evaluative language. Thanks to them, it is possible to implement practical applications and study them from a new perspective.

Sentiment is especially sensitive to the domain since the valences of the analysis will vary depending on the type of specialized language (or not) that is used. For this reason, the domain problem is addressed from the study of the specific economic-financial domain and the main works carried out on the domain are reviewed. It is a round trip between two disciplines because, in relation to this section, we observed the advantages and disadvantages of applying the lexicon-based approach in this domain, which is an eminently terminographic work. In addition, we analyze theoretical models in which meanings undergo changes in their connotations and result in the germ of new word meanings so that they acquire new aspects of meaning while others fall into disuse.

The literature on semantic change was also revised. It is a phenomenon of transformation in language related to the evolution of the use of the word at different levels. Each word has a variety of meanings and connotations, which can be added, eliminated or varied over time. Change of meaning, then, concerns a shift of relationships between words on the one hand and integral senses on the other hand. Behind each change of meaning there is a chain of causality that can be observed at various levels (changes in material, social or psychological conditions, etc.). In relation to sentiment, semasiological changes are relevant, highlighting the concepts of amelioration and pejoration here. In these processes of change, the meaning of a word changes to become more positive or negative, respectively.

Chapter 5 describes the methodological design, the data sources and the necessary tools to comply with the established objectives. It is divided into five parts: a review of the paradigms and methodological approaches related to the work, the methodological design of the thesis, the design and construction of a specific domain corpus, the description of the external tools used (Lingmotif, the sentiment analysis software, and The Sketch Engine corpus suite) and the construction of LexiEcon a large domain-specific lexicon in English for Lingmotif.

The first two sections of the chapter propose a mixed methodological design in four steps in order to extract and classify information from on specialized domain texts. It provides a perspective that rarely gets much attention in the literature: the expression of opinion and emotions, an aspect of language that, as it seems clear, permeate the entire use of language. The design of this research can be summarized in the development of a model that involves four steps:

1. Compilation of a specific domain lexicon for the sentiment analysis software:
  - 1.1. Design and construction of a specific domain corpus.
  - 1.2. Terminology extraction through different corpus linguistics techniques.
2. Through time series, analyze the longitudinal evolution of the feeling of a series of terms found in sentences from the corpus
3. To determine the extra-linguistic factors (social, political, economic, etc.) that influence the changes.
4. To interpret qualitatively, in detail, the diachronic sociolinguistic variations of the conditions of use of the selected terms (frequency, events, placements, uses of the language).

The third part of the chapter dealt with the design and collection of data source: the texts that formed the Esmeraldas Corpus of Business English of + 24M tokens. The corpus is compiled according to the traditional balance and representativeness principles of the corpus linguistics. The collection of texts classifies chronologically all the texts of the Business section of the newspapers The Guardian and New York Times throughout the years in which the crisis took place (2007-2015). For their extraction web scrapers

were used, and the texts were classified in a semi-automatic way. The final product was processed and packaged in The Sketch Engine and is available on demand.

The next section describes the two external tools used to meet the objectives: a corpus suite, The Sketch Engine, and sentiment analysis application, Lingmotif. The first is a corpus manager and text analysis software online that offers multiple possibilities (concordance, terminology extractor, word sketches, among others) and also offers a wide variety of corpora of multiple domains and languages. On the other hand, Lingmotif is an application that analyzes texts from the perspective of sentiment analysis. In short, it allows determining the semantic orientation (if positive or negative and to what degree) of a text by detecting linguistic expressions that show a certain polarity. It includes a detailed English and Spanish lexicon and plug-ins that can easily adapt the sentiment analysis to the various specialized languages.

The fifth section fulfilled the first specific objective: compiling a large sentiment lexicon of the economic-financial domain that conforms to the specifications of Lingmotif. The lexicon was compiled according to the theoretical base reviewed. For this, quantitative and qualitative terminology extraction techniques have been combined. The term candidates were obtained mainly from corpus-based terminology extraction techniques in which, through KeyWords of The Sketch Engine, 30,000 term candidates were extracted from the ECBE. With the help of a domain expert, each of the candidates was compared to the values of the general domain sentiment lexicon entries. Only those entries whose sentiment in the specific domain differed from the base lexicon were selected and each one of them was assigned the valence corresponding to the specific domain.

As a result, we present LexiEcon, a domain-specific plug-in lexicon for the English language that currently includes 6,843 entries. The results of his evaluation were very satisfactory. After analyzing a gold standard that contains 400 random specific domain sentences extracted from the ECBE corpus through two sweeps (first applying only to the general lexicon and, later, in conjunction with the plug-in), excellent levels of accuracy in the domain are observed. Performance figures are remarkable as, in conjunction with the Lingmotif's core lexicon, an F1 average of 0.735 was obtained. This figure performs around 20% better than the results that Lingmotif offers without plug-in at the time of classifying the texts of the economic-financial domain. In addition, given the technical specifications and the open data philosophy, LexiEcon is a final product that offers great versatility as it permits its expansion and adaptation according to the specific needs of other systems and research models.

The specialized nature of the texts facilitates the sentiment analysis based on vocabulary and offers better results than in the general domain. Although the use of the terms is not entirely univocal, the good results are justified by the nature of the specialized language since it tends to a low level of ambiguity. While there is a lack of strict precision in the semi-specialized press discourse, these levels of accuracy in the use of language are much greater compared to the use of language in the general domain (e.g., the use of ironies or anaphora). In addition, the written composition of the economic articles favors clearer enunciations in order to ease the understanding of information with a high level of abstraction.

In Chapter 6, a sentiment analysis of ten key words during the crisis has been applied. These words are the following: “credit”, “debt”, “inflation”, “housing”, “markets”, “investor”, “public”, “reform”, “spending” and “company”. The analyzed data were sentences extracted from the corpus that included the selected key terms. These data were classified by date and collected in a structured data CSV file so that each line included a sentence according to the technical requirements. These files were processed

with Lingmotif that applied the LexiEcon specific domain plug-in. The analysis consists of three parts: (a) a table of longitudinal descriptive statistical results related to sentiment scores, (b) an annual table of collocations and (c) a discussion of the findings in the corpus from the observation of annual rankings of collocations intending to triangulate the data obtained.

Chapter 7 sums up the main findings on the processing and subsequent data analysis. It discusses these results in the theoretical framework. Around the results, two facts stand out:

1. Terms become event words given the enormous increase in their frequency of use, in relation to the key events of the crisis. Generally, as a result of this increase, significant changes occur in the use and the type of collocations varies, new lexical units with a different semantic orientation coming into play. In general, they tend to be less specialized than traditional fixations related to terms.
2. The average annual semantic orientation of contextualized terms permits provide a crystal-clear observation of important fluctuations in the sentiment embedded in the speech. A triangulation of the quantitative data with its most significant collocations and the historical events related to the Great Recession allow us to conclude that the semantic orientation of the terms of the economic-financial domain is very susceptible to change as the facts of the financial crisis developed.

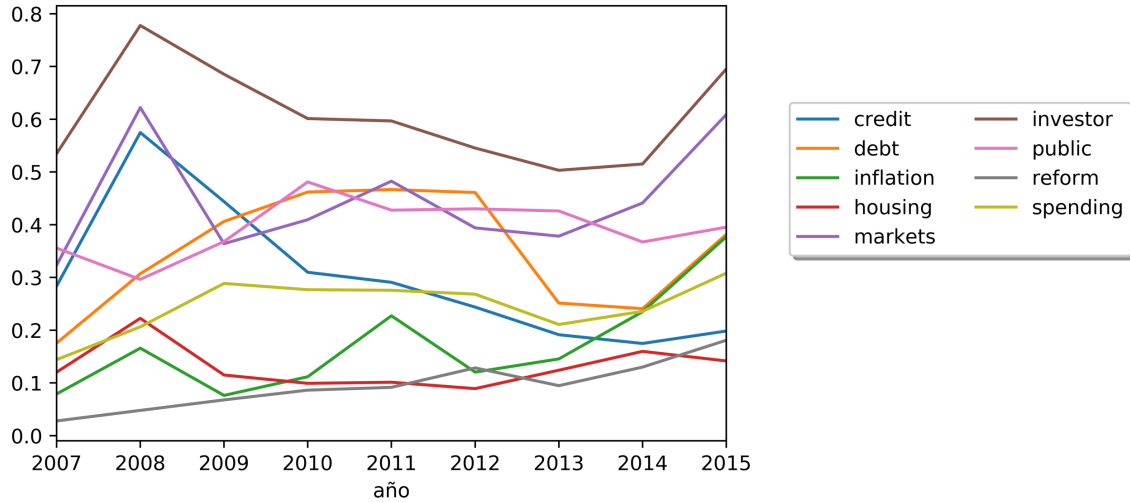
Economic journalism is a mediated discourse, in which journalists interact with various categories of experts (politicians, economists, users, etc.). The expert-journalist-public transmission chain highlights the relevance of the use of certain specialized lexical units, turning them into events. In perspective, it is this circulation that makes a mundane event transmute into an event, assigning a memorable social meaning to its outstanding terms, with certain connotations over a period of time.

The task of analysis is to identify gaps in the linguistic uses that make up the events and establish the causes. The quantitative observation of the relative frequency of the use of the different terms (Figure A) shows that in most cases the outbreak of the crisis in 2008 was very significant, generating a peak in the use of the terms in question. In many cases, in addition, this has been accompanied by a radical change in its connotations of semantic orientation.

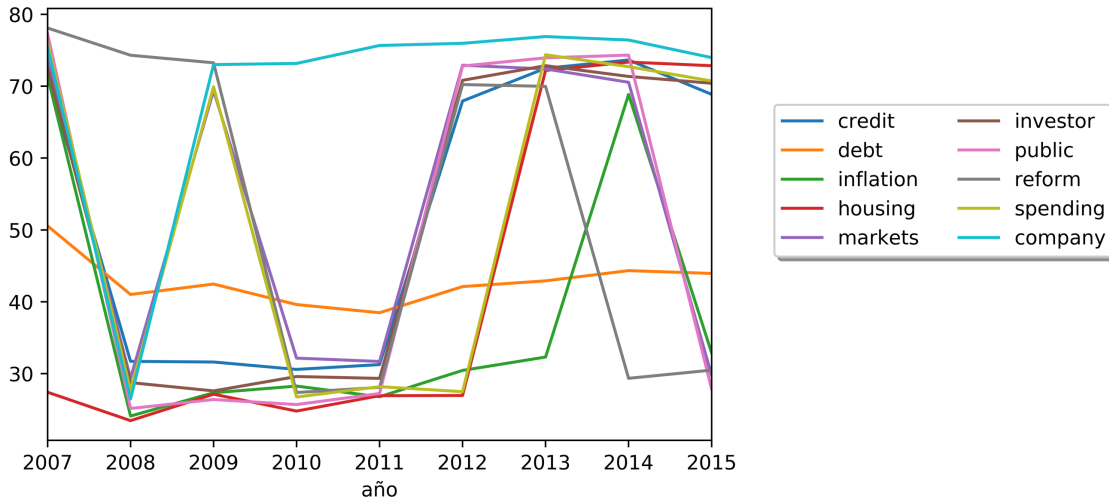
In most cases, figures move from a positive pre-crisis average sentiment in 2007 to a fall in the average threshold of negativity, as can be seen in Figure B. The separation of the credit crisis with the debt crisis is also evident in both frequencies and sentiment and similar patterns are observed in most cases. In fact, in 2009 there was a punctual increase in the average sentiment data to fall again in 2010 as the debt increased and, consequently, the deterioration in the living conditions of the citizens. It is possibly caused by a slight improvement in the macroeconomic conditions of the countries in the North of the Euro Zone. To a large extent, the launching of policies that supposed a relative economic relief, especially the quantitative easing caused a turn in spirits and discourse by the European Central Bank iconized in the famous speech “Whatever it takes” by its president, Mario Draghi. From that time on, there was a significant increase in sentiment scores, which in most cases lasted from 2011 or 2012 until the end of the study.



Evolución de la frecuencia relativa de los distintos términos de estudio

**Figure 1.** Evolution of the relative frequency of the different study terms

Evolución del sentimiento de los distintos términos de estudio

**Figure 2.** Evolution of the sentiment of the different terms of study

From what is observed in the term's collocations, it is found that semantic prosody is a non-static property of the lexicon, but variable in time. According to Bublitz (1996) or Morley et al. (2009), a set of collocations evolves to acquire a tinge of semantic orientation positive or negative to over a time interval. The quantitative changes in sentiment arising from the crisis have gone hand in hand with the appearance of collocations with positive or negative sentiment. In most cases, the collocations coincide with the fluctuation of the sentiment averages of the term's context analysis and also with economic events. To illustrate this, in the case of "public", in the middle of the credit crisis in 2008 it is observed that most sentences include a negative sentiment. In addition, the collocations that are in the neighborhood of the different sets also tend to be it. For example, the term co-occurs punctually with the noun "cuts", and at the same time there are frequent collocations that maintain the harmony of sentiment, for example: "rig", "conspiring" or "miserable". These evidences the use of the terms in more negative contexts. It occurs for all purposes an immediate "contagion" of the pragmatic or attitudinal meaning of the term because of a sudden and violent event such as the crisis, which remained constant as long as its effects lasted. As a result, lexical units tend to

acquire more positive or negative connotations because their semantic prosody varies over a given time interval. The connotative meaning of the words is a *living entity* that is intimately related to the mental state and the opinion of the community of speakers. The results make explicit that, in addition to the technical transmission of information, opinion is inscribed in the translation of the abstract macroeconomic figures into linguistic statements.

At the theoretical level, three main conclusions can be drawn:

1. The results of the analysis of *LexiEcon*'s performance and the subsequent analysis of the texts show the enormous possibilities offered by the study of specialized languages from a scarcely explored perspective such as the expression of sentiment.
2. There are terms of the economic-financial domain that acquire great relevance during events of great social importance, as is the case of the Great Recession. In relation to these terms, it could be observed that this inherent sentiment is not in any of the cases a static value, but that to a greater or lesser extent this varies with the socioeconomic conditions.
3. The observation of the sentiment scores, the rankings of collocations and their evolution throughout the study interval is nothing but a staging of the approaches of systemic linguistics where language is a system of choices by which The speaker choose in an explicit way what expression (words, terms, etc.) will be used to the detriment of all of its alternatives. In fact, the good performance of LexiEcon in the analysis of the texts shows that the specialized language is not an exception and is organized in a non-arbitrary and entirely functional way, in consonance with its use. In this selection of lexicogrammatical features, it is important to remark that the expression of the evaluation does not escape grammar. From this point of view and according to the systemic functional perspective, language is a social semiotic resource that allows speakers to construct meaning from a range of choices.

Through the analysis of the terms, it has been possible to evidence that the evaluative use of terms is present in all linguistic levels. In fact, evaluative use is no exception for specialized language. Its staging complies entirely with the definition of evaluation seen in Chapter 3 because from this domain it is possible to observe the structured expression of sentiments on the entities that are dealt with, with a view to influencing an audience in a way that reflects the set of values of the community of speakers.

Daily use of terminology is not entirely univocal. It is necessary to emphasize that the informative use of language by the press requires lower levels of specialization than the precision required by the texts addressed to experts (e.g. in corporate disclosures). On the other hand, it is observed that the evaluative aspect plays a role fundamental. Through opinion and discussion, subjective perceptions about the state of economic news can be expressed and still meet the necessary and sufficient conditions to describe the reality of the crisis without losing rigor and taking into account the dialogical dimension that points out the frictions between different theoretical and ideological frameworks. In fact, the use of the terms corresponds to two main factors in the domain's scope. On the one hand, the strict and necessary scientific-technical expression in the economic domain and, especially here, the prototypical uses that emerge from the daily informative discourse of the specialized journalist as a mediator between the expert and the public. To give an example related to the use, this would be the case of "bad debt" which uses clear sentiment adjectives so that the negative nature of the debt becomes clear. Through examples like



these, it is confirmed that the evaluative aspect plays a role fundamental in this type of discourse.

The observation of collocations, co-text and context of the terms evidences fluctuations in the sentiment almost immediately as a reaction to economic events. From a linguistic based computational perspective, the observation of the sentiment of the specialized lexical units also allows to advance from conceptual terminology, prominently structural, to their consideration as units that are also communicative and emotional. Sentiment meanings go beyond the linearity concept/definition. According to contemporary notions, we propose some contributions to classical approaches terminology so that the aseptic and static view of terms is overcome and transferred to a theoretical set that reflects the polyhedral and dynamic nature of terminology. The results coincide with the criticism on the consideration of the terms as “tight” synchronic unit (e.g., Temmerman, 2000) which do not take into account any type of evolution of the term through time.

Chapter 9 presents the conclusions and is divided into two sections. Review the achievement of the main objective and the four specific objectives proposed and give a concise answer to the questions asked in the thesis's introduction. This work is carried out as a compilation of the discussions, reflections and findings advanced throughout the conclusions of each of the chapters of this thesis. In this way, it is intended to close the work done and open the door to the new questions that emerge from here. The general and specific objectives set for this thesis are widely met. It is considered that the longitudinal methodological development fulfills its double purpose of describing patterns of change and explaining them through its relationship with the sociopolitical events of the crisis.

The last section describes the limits and new challenges offered by the end of this research. These can be divided into three fields: expansion of data sources, the extension of the object of study and the extension of the technical question. In relation to the first aspect, it is considered that the design of this work has contemplated the corpus created as a single block. However, in the corpus composition, the classification by publication has been contemplated. On the one hand, it would be an interesting comparative study of the use of evaluative language between *The Guardian* and *The New York Times* and discover the linguistic, journalistic, social and ideological factors that influence on the possible differences. On the other hand, the experiment could be repeated with different publications, and the results compared with newspapers of a different ideology.

On the other hand, there is the question of the study of the expression of economic language from different levels of specialization. For example, through the analysis of public discussions on blogs, comments in newspapers or social networks or observe variations in the use of highly specialized language. In this way, the lexical, cognitive and affective divergences between different levels of specialization could be observed.

Several challenges are posed around the technical development of solutions for sentiment analysis and practical applications of LexiEcon. For example, to adapt the terminographic pipeline for the Spanish language in order to provide a multilingual. Finally, it is suggested to conduct trials in the study of the correlations between discourse and economic returns in order to establish a working methodology that is able to determine the impact of discourse on the economy in order to carry out analyses and predictions that benefit community welfare.

## Anexo 3. Conclusions

### Achievement of Objectives

This section aims to recapitulate the achievement of the main objective and the four specific objectives proposed and give a concise answer to the four formal questions made in the introduction of the thesis. This work is carried out as a compilation of the discussions, reflections and findings advanced throughout the conclusions of each of the chapters of this thesis. In this way, it is intended to bring closure to the work done and open the door to the new questions that emerge from here.

The present work is of a highly multidisciplinary nature. It combines a discussion of the eminently linguistic role of the study of specialized languages and the change in semantic orientation with the socioeconomic factors of the greatest economic crisis of the last eight decades; the dynamism of the economic news and the rigor of the latest technical developments in sentiment analysis, one of the most cutting-edge disciplines that natural language processing encompasses. In turn, the development of sentiment analysis as a discipline harmonizes theory-based terminographic techniques in conjunction with the latest developments in mathematical algorithms and data science.

Thanks to this perspective, observing changes in language becomes a fascinating task, either from more classical structural approaches such as variations in morphology, syntax or more linked to the study of meaning, mainly through the window that the study of semantic prosody opens. Any change in the different aspects of human nature responds to a historical motivation constructed in a collective way that highlights the latent hyperdynamic relationship between the communicative use of language and the cultural devices that surround us. Discursive acts reveal the cognitive nature and position of the speaker with the world, especially showing the subjective perception of the state of things. In short, the observation of the contextual relations of ordinary language through a time series allows to know in depth the structural relationships between the discourse and the spirit of the times and to observe with simplicity the fluidity of the feeling that emanates from them. Our longitudinal methodology fulfills its double purpose of describing sentiment change patterns and explaining them through its correlation with crisis-related sociopolitical events.

The general and specific objectives set for this thesis are widely met. With respect to the general objective, “To analyze fluctuations in the semantic orientation of certain financial terms within the period 2007-2015 through sentiment analysis”, this could be done thanks to the speed and efficiency offered by Lingmotif. Throughout Chapter 7 it was possible to analyze the quantitative data related to the evolution of the semantic orientation and the relative frequency of ten financial terms within the 2007-2015 period. Thanks to this quantitative study, it is observed that the sentiment of the terms is not static over time intervals, but there are rapid and pronounced variations (pejorations and ameliorations) of the semantic orientation that emerge in the use of the community of speakers as an immediate reaction to drastic changes in the socioeconomic state-of-affairs.

Sentiment depends deeply on context and, in addition, on the specific domain in question. When we open the communicative door of the use of specialized language, we can see the prosodic framework inscribed in the use. This becomes an excellent instrument to observe the projection of sentiment along short intervals, so that we can get a still picture of the contemporary use of language. This analysis reflects that the living connotative meaning of words and intimately related to the mental state and the opinion

of the community of speakers. Given the informative need of the population, the economic journalist must perform a mediating work between the expert and the public. The results make explicit that, in addition to the technical transmission of information, opinion is in any case inscribed in the translation of the abstract macroeconomic figures into linguistic statements.

In addition to opening new discussion horizons in the linguistic domain, it offers practical applications in economics. In short, the study of the evolution of sentiment also serves as a thermometer of experts' opinion on economic markers. As stated in the introduction to this thesis, sentiment analysis offers another perspective of economic policy observation that, although it is not yet possible to dominate, it begins to glimpse its potential and opens the door to new challenges in multidisciplinary research.

The first specific objective is "To build a specific domain sentiment lexicon in English for the domain of economic news of high precision and completeness from a specialized text corpus". The development of LexiEcon provides one of the most complete economic-financial domain lexicons in English currently available. As a result, our compilation, LexiEcon is a domain-specific plug-in lexicon for the English language that currently featuring 6,843 entries (words and multi-word expressions). As it was thoroughly reviewed in Chapter 6, the performance after its evaluation through a gold standard returned excellent precision and recall results. In addition, given the technical specifications and the open data philosophy, LexiEcon is a product that offers great simplicity of expansion and adaptation to the specific needs of other systems and research models.

In the process of compiling the *LexiEcon plug-in*, a corpus of economic-financial press texts in English structured monthly for the 2007-2015 period was compiled and made available on demand. Section 6.3 deals with the design and compilation of the *Esmeraldas Corpus of Business English* (ECBE), the main data source of this work featuring +24M tokens. The corpus is compiled according to the traditional balance and representativity principles of the corpus linguistics discipline: electronic, with authentic texts, balanced and representative of a language or variety of language. The collection classifies chronologically all the online texts of the Business section of the The Guardian and New York Times headers throughout the years in which the Global Crisis took place until the beginning of this study (2007 and 2015). Web scrapers were used to extract and classify texts semi-automatically. The final product was tokenized, lemmatized and packaged in The Sketch Engine and is available on demand. After its use, we reaffirm that the online general press is, without a doubt, one of the best possible resources. News faithfully reflect the discourse of an era and becomes an easy instrument to use to obtain accurate information about the language in use. Unlike financial reports or the complexity of the study of sentiment in the oral language or social media (loaded with irony and colloquial uses), the study of the press provides an interpretation of the political-economic situation with a degree of impartiality, but still loaded with opinion, which serves as a bridge between the opinion of the financial market and the opinion in the debates at street level.

Likewise, the researcher in search of a specialized discourse analysis methodology could find interesting the adaptation to his domain of specialty of the set of methodologies and procedures that were used to extract the linguistic and terminological information with the aim of analyzing sentiment and its representation in the specific domain lexicon from a specialized corpus. As detailed in section 6.5, quantitative and qualitative techniques were combined to extract terminology. An extensive list of 30,000 term candidates (15,000 words and multi-word expressions) was extracted from ECBE thanks to the KeyWords terminology extractor. With the help of an expert in the domain,

each of the candidates was compared with the values of the entries in the general domain lexicon, selecting only those entries whose sentiment in the specific domain differed. Subsequently, for each of them a valence corresponding to the specific domain was assigned. From the corpus, a series of directional elements was collected (e.g., “rise”, “fall”, “decline”, etc.) in conjunction with a list of potentially positive or negative economic indicators (e.g. “debt”, “spending”) to which the corresponding sentiment was assigned according to Lingmotif’s format. This way, the entry “<rise> \_3\_debt” allows to cover cases that use the same pattern and share sentiment such as “rise in bad debts”, “rise in the national debt” or “rising level of debt”.

The second specific objective is “To define a subset of longitudinal data in the form of sentences that contain the terms of study that will be the input of the sentiment analysis”. This objective was met through the sentence extraction tool offered in the Sketch Engine’s concordance search, which compiled all the sentences into a plain text file for each of the key terms, labelled individually by date. The data set meets Lingmotif technical requirements and is compiled under open standards. Like the corpus, these sets are easily modified and reusable, facilitate the automated study of the economic-financial domain and are available on demand.

The third specific objective is “For the economic-financial terms of study, identify the factors (social, political, economic, etc.) that have accompanied changes in the use, from the perspective of feeling, during the period 2007-2015”. In addition to the particularities of each term revised extensively in Chapter 7, in Chapter 8 we discuss that the relative frequency of the use of the different terms allows concluding that in most cases the outbreak of the crisis in 2008 caused a rise of frequency of terms related to events (e.g. the collapse of Lehman Brothers). This phenomenon, moreover, has been accompanied by a decrease in the threshold of negativity of the means of feeling for sentences that include these terms. In fact, in almost all cases, the starting point was a positive average of sentiment in 2007, which is followed in 2008 by a sharp drop in sentiment average falling under the threshold of negativity.

Another general factor is the separation of the “credit crunch” years, of a more global nature, with the debt crisis in the Euro Zone. Although was influential abroad, it can be considered as an eminently European phenomenon. For most terms, relative frequency figures and the sentiment averages offer similar patterns. In fact, in 2009 there was a specific increase in the average sentiment data to subsequently fall again in 2010 when the debt increased and, as a result of the austerity policies that sought to mitigate its effects caused the deterioration in the conditions of life of citizenship. In the 2010-2011 biennium, there was an increase in the average of sentiment scores towards the threshold of positivity in most cases, produced by a relative improvement in macroeconomic conditions, mainly in the northern countries of the Euro Zone. It is necessary to highlight the spirit of relief inflicted by the policies of quantitative expansion carried out by the European Central Bank and focused eminently on capital. It is gathered in the discussion that the perlocutive power of the spirit of the “Whatever it takes” speech by Mario Draghi in the press conference of the European Central Bank. This speech, according to the consensus of the experts, “saved the euro” and was also reflected in the semantic orientation average of related terms. The spirit of relief policies translated, in many cases, into a significant increase in sentiment scores, so it lasted until the end of the study.

The fourth and last specific objective is “To analyze qualitatively the possible variations in the semantic prosody of the study terms over the years”. Accompanying the quantitative study of sentiment, our research design also proposes a qualitative analysis of language use, so it serves as a triangulation of the results and provides greater robustness to our conclusions. Specialized languages are rich in terminology, expressions

and their own constructions that are far from being unalterable, and that also reflect the spirit of the times. In the vast majority of cases, the terms have gone on to have annual placements of a more negative or positive nature. All these factors come to light in the way in which evaluations and opinions are expressed in the texts. According to the statements reviewed throughout the project, e.g., Moesner (2003) or Morley et al. (2009), by which the semantic content of words is modified as they are associated with more or less favorable concepts than they used to do. As observed through the annual rankings of placements and the subsequent observation of the examples, the semantic orientation of the context of the terms evolves when their positivity or negativity is intensified, and it is the case that their polarity is reversed completely throughout a period of time. Furthermore, it is the case that while these units have been considered *even words*, they are usually related to lexical units of a non-specialized nature. To illustrate this, in 2007, during the outbreak of the housing bubble and prior to the collapse of subprime mortgages, the term “housing” was related to negative units, i.e., “downturn”, “slump”, “weaken”, “bubble”, “slowdown”, “slow” or “trouble”. With a recovery of the real estate market in the Anglo-Saxon sphere in 2013, the semantic orientation of placements is totally opposite i.e., “affordable”, “boom”, “recovery”, “allowance”.

In other cases, the use of collocations was not so evident. For this reason, a complete revision was made through the examples obtained from the corpus, which in the less evident chaos was enlightening, since neutral collocations prevailed in the rankings. The qualitative observation of the sentiment of texts at sentence level, and subsequently as a whole evidenced the prevalence of a determined semantic orientation that coincided with the values of sentiment average.

In addition to the presentation of the main objectives, two formal research questions were asked, which constituted the starting point and the main thread of the thesis. In a synthetic way, an answer will be given to each of them according to what has been discussed throughout the eight previous chapters. To facilitate the revision work, these questions will be reproduced in a textual manner.

Both questions can be answered together: “Do lexical units vary as they are associated with more or less favorable concepts than before?”, “Do higher historical events, e.g. crises, lead to rapid changes in the semantic orientation of the lexical units?” As we have seen throughout Chapter 7, there is a relationship between the increase in relative frequency and a change in the mean semantic orientation. As Bybee (2015) states, once a word enters fully into the daily speech of its speakers, in a greater or lesser degree it will undergo changes of a semantic nature. There is a tension between the need for stability in meaning (so that the language community uses these terms normally) and the need to adapt the terms to new uses. Although correlation does not imply causation, it has been observed here that changes in the sentiment of words are often accompanied by a sudden increase in their frequency of use.

There are event words, according to Moirand (2007), which started being highly specialized, that is, they initially have a superficially aseptic and technical point of view. As a consequence of a remarkable event e.g. a political turn, the use of these terms begins to evolve by becoming a means of transmitting specialized communication events for the general public. Finally, they end up acquiring lexical characteristics which become permanent to a greater or lesser degree from the events since their repercussions remain in the collective memory.

Almost in all cases, negative sentiment scores averages have accompanied unfavorable economic situations related to the factors that imply the use of a certain term and, in turn, the appearance of more negative collocations, and vice versa. The connotations of certain lexical units evolve from contexts with different sentiment values



that are observable through their different lexical relations: discursive reformulations, collocations, colligations, patterns, etc. As a consequence, lexical units tend to acquire more positive or negative connotations because their semantic prosody varies over a given time interval. It can be concluded that through the term “aura”, that is, semantic prosody, in Louw’s (1993) terms, the connotative meaning of the term fluctuates in relation to the course of events and their socioeconomic factors. To illustrate this, in the case of “public” during the credit crisis, it is observed that most sentences carry negative sentiment. In addition, the collocations surrounding different sets also tend to do the same. For example, the aforementioned term co-occurs punctually with the noun “cuts”, and at the same time there are frequent collocations that maintain the harmony of sentiment, for example: “rig”, “conspiring” or “miserable”.

We have worked exhaustively in the design and analysis of the Esmeraldas Corpus of Business English that compiles balanced economic press texts rich in expression of opinion along with factual information of the domain. As we have seen throughout Chapter 3, evaluative language permeates through the text (word, phrase, sentence, text and context). It is cumulative and manifests itself in very diverse ways. The study of grammatical patterns by means of corpus (e.g. Hunston et al., 2000) facilitates the detection of a series of relational structures (e.g. collocations, colligations or semantic prosody) susceptible of including evaluative segments. However, it is necessary to remark the difficulty that involves the differentiation between evaluative and factual enunciations. In the case of the economic domain, especially, there are statements that are apparently factual and yet are sentiment laden. For example, this is the case of “Analysts predict wages will keep rising” which, however, will carry positive sentiment from a social point of view.

Sentiment is domain-dependent, as can be observed from the results offered by LexiEcon’s tests. Sentiment analysis systems usually provide a smaller precision figures, as there are evaluative patterns particularities in each domain. The results offered by the analysis of an economic domain gold standard, are evident. There are deep differences between the application of Lingmotif+Lexiecon which return an average F-1 score of 0.735, a 20% more than the results returned by Lingmotif without plug-in.

Depending on the domain, the lexical units that through polysemy are present in different domains can carry different sentiment depending on the context. This is the main reason why it is necessary to adjust sentiment analysis systems to the domain. For example, while “launder” would carry neutral sentiment in the general domain, the observation of its use economy would refer to money laundering, which is a criminal economic activity. Thus, lexicon-based sentiment analysis system needs extensive lexicographical work on the target domain. In the case of economy, it is essential to extract a broad list of terminology, idioms and jargon that includes the three main elements: performance indicators, directionality of performance and other domain-specific terms.

On the other hand, the specialized nature of the texts facilitates the analysis of sentiment based on lexicon and offers better results than in the general domain. Although the use of the terms is not entirely univocal, the good results obtained are justified by the nature of the specialized language since it tends to low ambiguity levels. Although there is a lack of strict precision in the semi-specialized press discourse, these levels of accuracy are higher compared to the use of language in the general domain (e.g., the use of irony or anaphora). In addition, language use in economic articles favors the clarity of the enunciations in order to make it easier for the public to understand information with a high level of abstraction.

In closing, this work has successfully described the development of a methodology for observing the evolution of sentiment of economic-financial language during a major event: the greatest economic and social crisis of the last eight decades. After a complete review of the state of the art in the fields of evaluative language, specialized languages and sentiment analysis, this challenge has been carried out through an evidence-based methodological design, which combines the sentiment analysis with qualitative techniques. In addition to the scientific contributions, this work provides added value for the community through the technical development and release of a English-language sentiment for the domain of economy with a high level of precision and recall.

### **Limitations and future perspectives**

The limits and new challenges offered by the end of this research can be divided into three areas: expansion of data sources, extension of the object of study and extension of the technical question.

Regarding the chosen text sources, the design of this work has contemplated the corpus as a single block, even if there are two different sources in the variety of English. On the one hand, a comparative study of the use of evaluative language between both newspapers and further research on the linguistic, journalistic, social and ideological factors that influence the possible differences would be interesting. Given that both sources share, largely, a similar ideological spectrum, these results could also be contrasted with other publications of different ideology.

On the other hand, there is the question of the study of the expression of economic language from different levels of specialization. It would also be noteworthy to study the extent of the use of economic language in order to observe its changes in the discourse of the population from less formal sources than the journalistic headers such as the user-generated content: blogs, comments in newspapers or social networks. In fact, in relation to the wave of anti-austerity social movements of this decade, one could study if there has been a *rebellion* in the lexicon of the average speaker during the crisis at the time of acquiring the specialized lexicon. It would be pertinent to observe the possible variations in the use of highly specialized language from documents issued for specialists, as is the case of internal corporate disclosures. In this way, the lexical, cognitive and affective divergences between different levels of specialization could be observed: specialist discourse, journalistic / informative language and public consumer of the terms.

Finally, several challenges arise regarding the technical development of sentiment analysis solutions and the practical applications of *LexiEcon*. After weighing the advantages and disadvantages of the terminology extraction techniques described here, they could be applied in a Spanish language corpus in order to obtain a multilingual solution. A translatology application could be designed to take into account the conservation or variation of the semantic prosody of specific terms considered as translations over a given period. On the other hand, it would be interesting to extend the tests carried out (Fernández-Cruz and Ortega Palomo, 2019 in press) in the study of the relationship between discourse and economic returns, so that a work methodology can be established in order to determine the impact of discourse in the economy and carry out further analyses and predictions that may benefit community well-being.