



**El proceso de localización web: estudio  
contrastivo de un corpus comparable del  
género *sitio web corporativo***

Tesis Doctoral presentada por Miguel A. Jiménez Crespo  
y dirigida por la Dra. Maribel Tercedor Sánchez

Doctorado en Procesos de Traducción  
Departamento de Traducción e Interpretación  
Universidad de Granada  
Abril de 2008





## AGRADECIMIENTOS

He de dedicar mi primer agradecimiento a la Dra. Maribel Tercedor Sánchez por haber sido una maestra excepcional durante la larga y emocionante aventura que ha supuesto la presente tesis. Su inestimable dedicación, apoyo y pasión por la localización han hecho posible, a pesar de la distancia entre Carolina del Norte y Granada, que el camino iniciado hace cinco años haya llegado a buen puerto con la rigurosidad, calidad y buen hacer que tal empresa se merece.

Me gustaría asimismo dar las gracias a la Dra. Clara Inés López por haberme sabiamente introducido al fascinante mundo de la Lingüística de Corpus y sus aplicaciones a la Traductología. Igualmente, he de agradecer al Dr. Ricardo Muñoz el haber sido un ejemplo de rigurosidad en mis primeros pasos en este programa de doctorado.

Cómo no, he de agradecer la ayuda que me han brindado mis colegas en la Universidad de Carolina del Norte en Wilmington por sus inestimables consejos y por haberse convertido en el segundo grupo de investigadores con los que compartir mis dudas, problemas y debates dialécticos. En especial, he de dar las gracias a la Dra. Denise DiPuccio, directora del Departamento de Lenguas Extranjeras, y a la Dra. María Camí Vela que han contribuido a que mis años en UNCW sean los más productivos de mi vida.

No podría concluir este apartado sin una mención expresa a mis padres, Miguel y Juana M<sup>a</sup>, a mi prima Conchi, a mis hermanos, cuñados y sobrinos que han convertido mis estancias en España trabajando en mi investigación en el mejor entorno imaginable. Y por último, esta tesis no hubiera sido posible sin el apoyo de mis mejores amigos en EEUU, en especial Patrick, Maria, Max, Asheley y Tanya, que con su cariño y afecto se convirtieron en mi familia adoptiva durante esta importante singladura.



## SIGLAS Y ABREVIATURAS

CMS – *Content Management System*

I14n – *Internationalization*

IFE – Ingles para Fines Específicos

G11n – *Globalization*

LISA - *Localization Industry Standard Association*

L10N – *Localization*

LO – *Locale* de Origen

LM – *Locale* Meta

OSCAR - *Open Standards for Container/Content Allowing Reuse*

QA – Control de Calidad o *Quality Assurance*

ROI – *Return on Investment* – Rendimiento sobre la Inversión

TM – Texto Meta

TMX – *Translation Memory eXchange*

TO – Texto Origen

XLIFF – *Xml Localization Interchange File Format*

XML – *eXtensive Markup Language*



---

## ÍNDICE GENERAL

### INTRODUCCIÓN

<b>1. INTRODUCCIÓN .....</b>	<b>1</b>
1.1. Presentación.....	1
1.2. Motivación .....	2
1.3. Localización y traducción de sitios web.....	6
1.3.1. Investigación previa sobre la localización web (Jiménez 2005a).....	11
1.3.2. La localización web y su importancia en Traductología .....	14
1.3.3. Los sitios web corporativos como objeto de estudio .....	17
1.4. Hipótesis de trabajo .....	19
1.5. Objetivos.....	20
1.6. Estructura de la tesis.....	21

### PARTE I: MARCO TEÓRICO

<b>2. LA LOCALIZACIÓN COMO OBJETO DE ESTUDIO.....</b>	<b>27</b>
2.1. El origen de la localización .....	29
2.2. G.I.L.T.....	30
2.2.1. Globalización.....	34
2.2.2. Internacionalización .....	35
2.3. Definición de localización y <i>locale</i> .....	37
2.4. Agentes en el proceso de localización .....	45
2.5. La adaptación técnica .....	48
2.6. Componente cultural.....	50
2.7. Diferentes grados de localización .....	54
2.8. La localización de software y la localización web .....	59



2.9. Revisión de estudios sobre localización.....	63
2.9.1. Estudios desde la perspectiva traductológica .....	65
2.9.1.1. Las investigaciones de Alicia Bolaños (2002, 2003; Bolaños <i>et al.</i> 2005, 2007) .....	69
2.9.2. Estudios desde el sector de la localización.....	72
2.10. Recapitulación.....	73
<b>3. EL TEXTO EN LOCALIZACIÓN .....</b>	<b>77</b>
3.1. La noción de texto.....	79
3.2. Elementos de textualidad: la coherencia en localización .....	86
3.2.1. La coherencia como parámetro estático en el producto .....	88
3.2.2. La coherencia como parámetro dinámico en el proceso .....	89
3.3. El hipertexto.....	90
3.3.1. Clases de hipertextos.....	94
3.3.2. Características específicas de los hipertextos .....	97
3.3.3. La coherencia en hipertextos .....	100
3.4. El sitio web como texto unitario: hacia una definición.....	103
3.4.1. Definición de texto en localización .....	106
<b>4. LAS TIPOLOGÍAS TEXTUALES EN TRADUCTOLOGÍA.....</b>	<b>109</b>
4.1. Las clasificaciones textuales y la confusión terminológica y conceptual .....	111
4.2. Tipologías textuales .....	115
4.2.1. Tipos textuales en Traductología .....	117
4.2.1.1. La propuesta funcionalista de Katrina Reiss (1971, 1976, 1981) y Reiss y Vermeer (1984) .....	119
4.2.1.2. La propuesta de Juliane House (1977).....	123
4.2.1.3. La propuesta de Hatim y Mason (1990, 1997).....	126
4.2.1.3.1. El campo, el modo y el tenor.....	131
4.2.1.4. La propuesta de Nord (1997a, 1997b) .....	133
4.2.2. Tipologías textuales en Internet y Localización.....	140
4.2.2.1. Tipologías textuales en Internet (Biber 2004; Santini 2005a, 2005b) .....	141
4.2.2.2. Tipologías textuales en Localización (Austermühl 2006; Reinke 2005) .....	144
4.2.3. Los tipos textuales en el estudio comparativo de sitios web .....	148

---

<b>5. GÉNERO TEXTUAL EN LOCALIZACIÓN .....</b>	<b>151</b>
5.1. El concepto de género .....	152
5.1.1. Definición de género .....	154
5.1.2. Las clasificaciones por géneros: los subgéneros y supragéneros.....	160
5.1.2.1. Propuesta de situación del género objeto de estudio .....	163
5.2. El género y la noción de prototipo .....	165
5.3. El género como estructura textual.....	166
5.3.1. La superestructura textual en los géneros web .....	170
5.3.1.1. El cuerpo de la página web: el texto de «contenido» frente al texto «interfaz».....	173
5.3.1.2. Los menús de navegación .....	175
5.4. Estudios contrastivos empíricos basados en géneros .....	177
5.5. El modelo de género de Silvia Gamero (1998, 2001).....	179
5.5.1. Rasgos convencionales.....	180
5.5.1.1. Convención.....	180
5.5.1.2. Convenciones de género .....	183
5.5.1.3. Convenciones y normas.....	189
5.5.2. La función textual .....	190
5.5.3. Elementos de la situación comunicativa.....	192
5.5.4. Influencia del contexto sociocultural .....	195
5.5.5. Elementos intratextuales .....	197
5.5.6. Funcionalidad.....	199
5.5.6.1. La funcionalidad como criterio clasificador de géneros digitales en Shepherd y Watters (1998).....	202
5.6. Los géneros web: una constante evolución .....	205
5.7. Esquema del modelo de caracterización de género .....	207
5.8. Propuesta de caracterización general del género <i>sitio web corporativo</i> .....	209
5.9. Los géneros digitales desde la perspectiva prescriptiva de las guías de estilo web .....	210
5.9.1. Consideraciones generales.....	212
5.9.2. Nivel léxico-semántico .....	213
5.9.3. Nivel morfosintáctico .....	215
5.9.4. Nivel estilístico.....	216

5.9.5. Nivel tipográfico.....	219
5.9.6. Nivel estructural.....	221
5.9.7. Nivel icónico-visual .....	222
5.9.8. Edición del texto .....	224
<b>6. LA LINGÜÍSTICA DE CORPUS EN LA INVESTIGACIÓN EMPÍRICA EN TRADUCCIÓN .....</b>	<b>226</b>
6.1. La Lingüística de Corpus en Traductología .....	227
6.2. El concepto de corpus.....	230
6.3. Tipología de corpus.....	232
6.3.1. Tipología de corpus de Sara Laviosa (2002).....	234
6.3.2. Corpus Monolingües.....	236
6.3.3. Corpus comparables.....	238
6.3.4. Corpus paralelos.....	242
6.3.5. Corpus comparable frente a corpus paralelo en la investigación traductológica .....	243

## **PARTE II: MARCO METODOLÓGICO**

<b>7. METODOLOGÍA DE LA INVESTIGACIÓN EMPÍRICA.....</b>	<b>247</b>
7.1. Metodología de análisis .....	247
7.1.1. Análisis extratextual e intratextual .....	248
7.1.2. Metodología de análisis de la superestructura.....	250
7.2. Recopilación del corpus comparable .....	256
7.2.1. Criterios en la recopilación del corpus .....	258
7.2.2. Comparabilidad y representatividad en el proceso de recopilación .....	264
7.2.3. Criterios para asegurar que se trata de un texto original o localizado .....	268
7.3. Descarga del corpus .....	271
7.4. Composición final del corpus comparable .....	272
7.5. Limitaciones de la recopilación.....	275
7.5.1. Los marcos y la arquitectura servidor-usuario .....	276
7.5.2. Macromedia Flash .....	278

7.5.3. Sitios web eliminados por distorsionar las estadísticas.....	280
7.6. El análisis textual y la elaboración de subcorpus.....	281
7.7. Herramientas de análisis.....	283
<b>8. ANÁLISIS EXTRATEXTUAL DEL GÉNERO <i>SITIO WEB CORPORATIVO</i></b> .....	<b>289</b>
8.1. Criterios de análisis extratextual.....	289
8.1.1. Factores pragmáticos.....	290
8.1.1.1. Análisis de los factores pragmáticos mediante formas verbales .....	294
8.1.2. Factores comunicativos .....	298
8.1.2.1. El carácter multilingüe de los sitios web corporativos .....	300
8.1.2.2. Estadísticas de localizaciones de los sitios web españoles.....	304
8.1.2.3. Análisis del tenor .....	307
8.1.3. Análisis del contexto sociocultural .....	315
8.1.3.1. Marco normativo.....	316
8.1.3.2. Marco jurídico .....	319
8.1.4. Marco socioprofesional.....	322
8.1.4.1. El redactor técnico o redactor web .....	322
8.1.4.2. El localizador de textos web .....	326
8.1.4.2.1. La competencia traductora en localización .....	330

## PARTE III: RESULTADOS Y DISCUSIÓN

<b>9. ANÁLISIS INTRATEXTUAL CONTRASTIVO .....</b>	<b>334</b>
9.1. La superestructura prototípica del sitio web corporativo español.....	336
9.1.1. Análisis contrastivo de la superestructura.....	343
9.1.2. Análisis descriptivo-contrastivo de los bloques y secciones comunicativas.....	352
9.1.2.1. Análisis del bloque «página de inicio» (A).....	352
9.1.2.1.1. Análisis contrastivo de títulos en bloque «página de inicio» .....	353
9.1.2.1.2. Las páginas web de presentación.....	356
9.1.2.2. Análisis del bloque «contacto» (B) .....	357
9.1.2.2.1. Análisis contrastivo de títulos en el bloque «contacto» .....	359

9.1.2.3. Análisis del bloque «empresa» (C).....	362
9.1.2.3.1. Análisis contrastivo de títulos en el bloque «empresa».....	365
9.1.2.4. Análisis del bloque «noticias» (D).....	375
9.1.2.4.1. Análisis contrastivo de títulos en bloque «noticias».....	379
9.1.2.5. Análisis del bloque «productos/servicios» (E).....	381
9.1.2.5.1. Análisis contrastivo de títulos en el bloque «Productos-Servicios».....	383
9.1.2.6. Análisis del bloque «legal» (F).....	386
9.1.2.6.1. Análisis contrastivo de títulos en bloque «legal».....	391
9.1.2.7. Análisis del bloque «zonas especiales» (G).....	397
9.1.2.7.1. Análisis contrastivo de títulos en el bloque «Área de clientes» (G).....	401
9.1.2.8. Análisis del bloque «interactividad con el sitio web» (H).....	405
9.1.2.8.1. Análisis contrastivo de títulos en el bloque «Interactividad».....	409
9.1.3. Discusión sobre el grado de convencionalización terminológica de terminología en menús de navegación.....	415
9.2. Análisis intratextual de elementos en la estructura interna de las páginas web.....	420
9.2.1. Los títulos de las páginas html.....	421
9.2.2. Las descripciones en el atributo ALT y la accesibilidad.....	426
9.2.2.1. Elementos visuales y multimedia.....	430
9.2.3. Segmentos textuales en etiquetas <meta>.....	432
9.2.4. Los segmentos textuales en eventos OnMouseover.....	435
9.2.5. Los segmentos textuales en scripts, <Select> y <Input>.....	435

## **10. ANÁLISIS CONTRASTIVO DE FORMAS LINGÜÍSTICAS**

### **CONVENCIONALES EN BLOQUES CONTACTO (B) Y LEGAL (D)..... 438**

10.1. Metodología de análisis de formas lingüísticas convencionales.....	439
10.2. Análisis adicional del bloque «contacto» (B).....	441
10.2.1. Análisis de la unidad significativa de ofrecimiento de información de contacto (B-I) .....	443
10.2.2. Análisis de la unidad significativa formulario de contacto (B-III).....	450
10.2.2.1. Subunidad significativa «nombre» (B-III-a).....	457
10.2.2.2. Subunidad significativa «información personal adicional» (B-III-b).....	459
10.2.2.3. Subunidad significativa «empresa» (B-III-c).....	460
10.2.2.4. Subunidad significativa «dirección» (B-III-d).....	460
10.2.2.5. Subunidad significativa «medios de contacto» (B-III-e).....	461

10.2.2.6. Subunidad significativa «comentarios» (B-III-g).....	463
10.2.2.7. Subunidad significativa «envío y validación de formulario» (B-III-i) .....	464
10.3. Análisis adicional del bloque legal (F) .....	467
10.3.3. Análisis contrastivo de unidades y subunidades significativas en bloque legal (F) ....	471
10.3.4. Análisis léxico comparativo del subcorpus legal.....	476
10.3.5. Análisis de unidades y subunidades significativas en la sección avisos legales (F1) ...	480
10.3.5.1. Unidad significativa de presentación de avisos legales (F-1-I).....	481
10.3.5.2. Unidad significativa de aceptación de los términos legales (F-1-II) .....	482
10.3.5.3. Unidad significativa de registro legal (F-1-III).....	486
10.3.5.4. Unidad significativa sobre la legislación aplicable y jurisdicción (F-I-IV).....	488
10.3.5.5. Unidad significativa sobre los derechos de propiedad intelectual (F-I-V) .....	494
10.3.5.5.1. Análisis de subunidad significativa sobre materiales protegidos por los derechos de propiedad intelectual (F-I-VI-a).....	498
10.3.5.6. Unidad significativa sobre el titular del sitio web (F-I-VII).....	499
10.3.6. Análisis de unidades y subunidades significativas en la sección de privacidad (F2)...	501
10.3.6.1. Unidad significativa sobre el cumplimiento de la ley española, LOPD, del 15/1999 (F-2-I-a) .....	502
10.3.6.2. Unidad significativa sobre enlaces hacia y desde el sitio web (F-2-IV).....	503
10.3.6.3. Unidad significativa de contacto por cuestiones de protección de datos (F-2-VI) .....	507
10.3.7. Análisis de unidades y subunidades significativas en la sección de condiciones generales (F3) .....	508
10.3.7.1.1. Subunidad significativa sobre la limitación de responsabilidades por virus (F- 3-I-a).....	512
10.3.7.2. Unidad significativa sobre modificaciones al contenido del sitio (F-3-II).....	518
10.3.8. Resumen contrastivo de formas lingüísticas convencionales asociadas a la sección legal (F) .....	522
10.4. Resumen y discusión sobre análisis adicional de los bloques «contacto» y «legal».....	529
<b>11. LOS ERRORES DE LOCALIZACIÓN Y LAS CONVENCIONES DE GÉNERO EN LOS TEXTOS LOCALIZADOS.....</b>	<b>536</b>
11.1. El proceso de localización y los errores de traducción.....	539
11.2. Las formas lingüísticas convencionales y los menús de navegación.....	541

11.3. Errores de traducción en los menús de navegación.....	542
11.3.1. Errores de coherencia terminológica.....	542
11.3.2. Errores de expresión en la lengua de llegada en menús de navegación .....	545
11.4. Otros errores de expresión en la lengua de llegada en los sitios web localizados.....	549
11.4.1. Nivel léxico-semántico .....	549
11.4.2. Nivel morfosintáctico .....	553
11.4.3. Nivel estilístico.....	555
11.4.4. Errores ortotipográficos.....	557
11.5. Reflexiones finales sobre los errores de traducción en textos localizados.....	560

## **PARTE IV: CONCLUSIONES**

12. CONCLUSIONES .....	567
12.1. Futuras vías de investigación .....	583

## **PARTE V: BIBLIOGRAFÍA Y APÉNDICES**

13. BIBLIOGRAFÍA .....	591
14. APÉNDICES.....	622
14.1. Apéndice I. Composición del corpus comparable .....	622

## ÍNDICE DE FIGURAS

Figura 1. Predicción en el volumen de negocio en industrias de la lengua para el 2007. Fuente IDC Research (cf. Schäler 2005).....	7
Figura 2. Evolución del uso de Internet en España. Fuente: Asociación de Usuarios de Internet. ....	9
Figura 3. Frecuencia de uso en corpus original y localizado. De Jiménez (2005a).....	13
Figura 4. Interdependencia de los procesos GILT y desarrollo de la concienciación sobre las interdependencias de los GILT. Adaptado de Dunne (2006: 5). ....	31
Figura 5. Esquema del ciclo de desarrollo de un producto global según LISA (2007:19).....	32
Figura 6. Proceso global de globalización según Microsoft (Dr. International 2003). ....	33
Figura 7. Propuesta de interdependencia de los procesos GILT en los que se enmarca la traducción. ....	33
Figura 8. Adaptación cultural a la percepción según Singh y Pereira (2005). Las flechas apuntan a la dirección de lectura sea a la derecha en hebreo o la izquierda en inglés. www.ikea.com. ....	53
Figura 9. Diferentes grados de localización en los productos de software (Dr. International 2003). ....	55
Figura 10. Diferentes grados de localización en Ikea.com. El sitio de Holanda es similar al de Francia. Igualmente el sitio turco y saudita presenta contenidos similares aunque adaptados a la percepción con su lectura de izquierda a derecha o al contrario. ....	57
Figura 11. Disposición líquida en pantalla con posibilidad de variar el tamaño de la letra. Icono de www.ono.es. ....	61
Figura 12. Producto de software completo. Elaboración propia.....	63
Figura 13. Modelo de texto lineal. Adaptado de Fritz (1997: 222). ....	91
Figura 14. Modelo de la estructura de un hipertexto en el contexto de la WWW. Adaptado de Janoschka (2003: 173). ....	93
Figura 15. Esquema de <i>Global Content Management System</i> . ....	98
Figura 16. Aviso de abandono de un sitio web global como texto web unitario. ....	104
Figura 17. Tipología textual de Reiss (1971, 1976, 1977). Adaptado de Hurtado Albir (2001: 475). ....	121
Figura 18. Tipología textual propuesta por Hatim y Mason (1990: 158).....	129
Figura 19. Esquema de las tres dimensiones del contexto (Hatim y Mason 1990: 58).....	132
Figura 20. Tipología funcional de traducciones de Nord (1997 a: 52).....	138
Figura 21. Tipología textual en localización de Austermühl (2006: 80).....	147
Figura 22. Relación de entre géneros web complejos y secundarios. ....	162



Figura 23. Propuesta de situación del género objeto de estudio, el sitio web corporativo tradicional. .....	164
Figura 24. Superestructura de un sitio web representada en el mapa del sitio web jerárquico (Engebresten 2008). <a href="http://www.lifescan.es">www.lifescan.es</a> .....	169
Figura 25. Distribución textual en página web. De <a href="http://www.microrent.es/index.html">www.microrent.es/index.html</a> . ....	174
Figura 26. Menú de navegación superior. <a href="http://www.www.mahou-sanmiguel.com/inicio.html">www.www.mahou-sanmiguel.com/inicio.html</a> .....	175
Figura 27. Pie de página. <a href="http://www.azud.com/es/index.html">www.azud.com/es/index.html</a> .....	176
Figura 28. Modelo de comunicación de masas interactivo. Adaptado de Janoschka (2003). ....	193
Figura 29. Resumen de contenido forma y funcionalidad de los seis cibergéneros descritos por Shepherd y Watters (1999).....	201
Figura 30. Los cibergéneros. Adaptado de Shepherd y Watters, (1998), y ubicación del objetivo de estudio, combinado con Kennedy y Shepherd (2005). ....	203
Figura 31. Clasificación de géneros según la empresa Nielsen-Netratings. ....	207
Figura 32. Esquema de modelo de caracterización de géneros digitales, adaptado de Gamero (2001: 60). ....	208
Figura 33. Estilo periodístico textual en forma de pirámide invertida. Adaptado de Jeney (2006: 148). ....	215
Figura 34. Diferencia entre la fuente <i>sans-serif</i> frente a la <i>serif</i> . ....	219
Figura 35. Inclusión de inglés en formulario bilingüe. <a href="http://www.fcasas.es/form-2.html">www.fcasas.es/form-2.html</a> .....	238
Figura 36. Bloque «contacto» incluido en la página de inicio. ....	253
Figura 37. Superestructura representada en el mapa web. De <a href="http://www.arteriego.com">www.arteriego.com</a> . ....	254
Figura 38. Extracto de lista global de palabras en corpus original que demuestra la presencia del inglés. ....	301
Figura 39. Página de inicio multilingüe del sitio web <a href="http://www.fabregas.com">www.fabregas.com</a> .....	302
Figura 40. Idioma de la página de destino en menú de navegación. De <a href="http://www1.euro.dell.com">www1.euro.dell.com</a> . ....	304
Figura 41. Porcentaje de localizaciones en español de las mayores empresas norteamericanas. ....	305
Figura 42. Unidades léxicas y formas léxicas más frecuentes en la selección de <i>locale</i> . ....	307
Figura 43. Análisis contrastivo del tenor digital a través del modo imperativo según frecuencias globales en ambas secciones del corpus. ....	309
Figura 44. Análisis contrastivo del tenor digital por formas verbales en modo imperativo en textos globales. Estadísticas extraídas de la frecuencia en la lista global de palabras de ambas secciones del corpus. ....	310
Figura 45. Análisis contrastivo del tenor digital a través de las formas de cortesía. ....	311
Figura 46. El tenor en los menús de navegación. Extracto de bloque «Contacto».....	312

Figura 47. El tenor en los menús de navegación. Formas verbales en tú frente a usted «Contacto».	313
Figura 48. El tenor digital en la sección «Registro».	313
Figura 49. El tenor digital en la sección «Registro», formas verbales en modo imperativo formales e informales.	314
Figura 50. Tenor digital en la sección «Trabajo». Sustantivo o forma verbal.	315
Figura 51. Porcentajes de uso de codificación en ambos corpus.	319
Figura 52. Grado de prototipicidad de bloques comunicativos en el género web sitio web corporativo español.	338
Figura 53. Variables de análisis contrastivo de la superestructura ( <i>content features</i> ) y la forma ( <i>design features</i> ) de Robbin y Stylianou, (2003: 207).	342
Figura 54. Bloques y secciones con mayor índice diferencial de frecuencia.	350
Figura 55. Uso de iconos anglosajones en página web de inicio española. <a href="http://www.caditesting.es">www.caditesting.es</a> .	354
Figura 56. Análisis contrastivo de términos empleados en el bloque página de inicio (A).	355
Figura 57. Pantallas animadas <i>Splash</i> en formato Macromedia Flash.	357
Figura 58. Formulario de contacto con una función textual primaria exhortativa. <a href="http://www.adea.es">www.adea.es</a> .	358
Figura 59. Análisis contrastivo de terminología en bloque contacto (B).	360
Figura 60. Análisis contrastivo de terminología en bloque empresa (C).	366
Figura 61. Análisis contrastivo de la terminología más frecuente en bloque empresa (C) según sinónimos.	367
Figura 62. Análisis contrastivo de terminología en sección localización (C1).	369
Figura 63. Análisis contrastivo de terminología en sección misión (C3).	372
Figura 64. Ejemplo de enlace a sección D5, RSS. De <a href="http://www.gruporh.com">www.gruporh.com</a> .	378
Figura 65. Análisis contrastivo de terminología en bloque «noticias» (D1).	380
Figura 66. Análisis contrastivo de terminología en sección productos (E1).	384
Figura 67. Análisis contrastivo de terminología en sección servicios (E2).	385
Figura 68 Análisis contrastivo de terminología en sección ofertas (E3).	386
Figura 69. Porcentaje de distribución de secciones en bloque «legal» en textos originales.	388
Figura 70. Análisis contrastivo de terminología en sección aviso legal (F1).	392
Figura 71. Análisis contrastivo de terminología en sección privacidad (F2).	393
Figura 72. Análisis contrastivo de terminología en sección condiciones generales, F3.	395
Figura 73. Análisis contrastivo de terminología en sección trabajo (G1).	402
Figura 74. Calco tipográfico en la sección trabajo (G1). Sitios web de EMC y Dell.	403
Figura 75. Análisis contrastivo de terminología en sección inversores (G8).	404
Figura 76. Errores de localización en la activación de la sección buscar (H2).	407

Figura 77. Análisis contrastivo de terminología en sección mapa web (H1). .....	410
Figura 78. Análisis contrastivo de terminología en sección buscar (H2). .....	411
Figura 79. Análisis contrastivo de terminología en sección preguntas-ayuda (H3). .....	413
Figura 80. Análisis contrastivo de terminología en sección enlaces (H4). .....	414
Figura 81. Análisis contrastivo de terminología en sección registro (H5). .....	415
Figura 82. Título de la página web de <a href="http://www.innova31.es">www.innova31.es</a> con indicación del título general del sitio web y de la página específica del mismo. ....	423
Figura 83. Menú de navegación con texto codificado en imágenes. <a href="http://www.pandasoftware.es">www.pandasoftware.es</a> . ....	431
Figura 85. Página protegida que indica un área reservada de <a href="http://www.diode.es">www.diode.es</a> . ....	436
Figura 86. Segmentos textuales en evento « <i>Select</i> ». ....	436
Figura 87. Ejemplo de texto a localizar en etiqueta <code>&lt;input&gt;</code> . .....	437
Figura 88. Relación entre sección, unidades significativas y subunidades significativas en bloque contacto (B). .....	440
Figura 89. Extracto de concordancias de «información» en subcorpus de contacto original. ....	443
Figura 90. Extracto de concordancias de la palabra «información» en subcorpus de contacto localizado. ....	444
Figura 91. Extracto de concordancias de «en contacto» en subcorpus localizado de contacto con predominio del modo imperativo. ....	447
Figura 92. Extracto de concordancias de «en contacto» en subcorpus original de contacto con predominio de formas de cortesía. ....	448
Figura 93. Formas lingüísticas convencionales asociadas con la colocación «contactar con». ....	449
Figura 94. Colocaciones de la palabra clave «formulario». ....	452
Figura 95. Verbos asociados a la palabra «formulario» en subcorpus comparable de contacto. ....	453
Figura 96. Subunidad significativa «tratamiento» (B-III-a-1) exclusiva de los textos localizados. ...	458
Figura 97. Extracto de colocaciones del verbo «leer» en subcorpus original legal que ilustra la presencia convencional de adverbios intensificadores. ....	482
Figura 98. Extracto de concordancias de la palabra «aceptación» en subcorpus legal original que ilustra su fraseología convencionalizada. ....	484
Figura 99. Extracto de concordancias de la palabra «aceptación» en subcorpus legal localizado. ...	484
Figura 100. Jurisdicción en sitios web localizados del corpus. ....	489
Figura 101. Extracto de concordancias de la palabra «aplicable*» en subcorpus legal original. ....	490
Figura 102. Extracto de concordancias de la palabra «aplicable*» en subcorpus legal localizado. ...	490
Figura 103. Extracto de concordancias del verbo «someter» en subcorpus legal original. ....	491
Figura 104. Extracto de concordancias del verbo «someter» en subcorpus legal localizado. ....	491
Figura 105. Extracto de concordancias del verbo «regir» en subcorpus legal original. ....	492

Figura 106. Extracto de concordancias del verbo «regir» en subcorpus legal localizado. ....	492
Figura 107. Extracto de concordancias de lema «renunc*» en subcorpus legal original.....	493
Figura 108. Extracto de concordancias de lema «renunc*» en subcorpus legal localizado.....	494
Figura 109. Extracto de concordancias del lema «prohib*» en subcorpus legal original. ....	497
Figura 110. Extracto de concordancias del lema «prohib*» en subcorpus legal localizado. ....	497
Figura 111. Extracto de concordancias del término «ley orgánica» en el subcorpus original legal.	503
Figura 112. Extracto de concordancias del término «ley orgánica» en el subcorpus legal localizado. .....	503
Figura 113. Extracto de concordancias de la palabra «virus» en subcorpus legal original. ....	512
Figura 114. Extracto de concordancias de la unidad fraseológica exclusiva de los textos localizados «tal cual» en subunidad significativa sobre errores o interrupciones en el sitio (F-3-I-b). ....	515
Figura 115. Falta de coherencia en los menús de navegación. Sitios webs de EMC y de Black&Decker. ....	544
Figura 116. Falta de coherencia en los menús de navegación. Sitio web de EMC. ....	544
Figura 117. Falta de coherencia en los menús de navegación. Sitio web de Mattel.....	545
Figura 118. Error en la sección H2, «buscar». Sitio web de ADM. ....	548
Figura 119. Error en la sección H2, «buscar». Sitios web de Otis y Alcoa. ....	548
Figura 120. Inconsistencia en la pluralización de siglas. Ejemplo de HP.....	552
Figura 121. Error sintáctico relacionado con la variedad dialectal del español no propia de España. Sitio web de EMC.....	554
Figura 122. Error en la falta de coherencia en el uso del tratamiento formal e informal en la misma página. Sitios web de Motorola y Sun.....	555
Figura 123. Ejemplos de errores ortográficos o cacografías, incluyendo la sustitución de la «ñ» por la letra «n».....	558
Figura 124. Inconsistencia en el uso inadecuado de mayúsculas en títulos en Cigna y Xerox. ....	558
Figura 125. Uso erróneo de mayúsculas en fechas en Electronic Arts. ....	558
Figura 126. Anglicismo ortográfico en la descripción de las décadas y años en Purina.es ....	558
Figura 127. Anglicismo ortotipográfico y error en cifras con el uso de punto para marcar tanto los millares como los decimales. ....	559
Figura 128. Error ortotipográfico en la falta de signo inicial de exclamación en sitios de Motorola y Mattel.....	559
Figura 129. Subunidad significativa «tratamiento» (B-III-a-1) exclusiva de los textos localizados, con errores de traducción. ....	563

## ÍNDICE DE TABLAS

Tabla 1. Ciclos de vida en productos de software y contenido web con mayores posibilidades de actualización en los textos web (Mata 2005: 194).....	62
Tabla 2. Metodología de análisis de subgénero web «sitio web de empresa de informática» de Bolaños <i>et al.</i> (2005).....	71
Tabla 3. Tipología de corpus adaptada de Sara Laviosa (2002: 34-38), con el objeto de estudio de la presente tesis. ....	235
Tabla 4. Composición final del corpus y análisis comparativo. Estadísticas elaboradas con Webbudget. ....	273
Tabla 5. Correlación de las distintas fases de análisis y el uso de corpus o elaboración de subcorpus comparables. ....	282
Tabla 6. Verbos más frecuentes en lista global de palabras en corpus original lematizada. ....	295
Tabla 7. Extracto de líneas de concordancia de «desea» con Wordsmith Tools. ....	296
Tabla 8. Verbos más frecuentes en lista global de palabras en corpus original no lematizada. ....	296
Tabla 9. Distribución de versiones localizadas de sitios web españoles. ....	304
Tabla 10. Frecuencia de aparición en el corpus original de las variables de análisis que responden a los bloques y secciones.....	340
Tabla 11. Análisis contrastivo de la superestructura de los textos originales y localizados. ....	346
Tabla 12. Análisis contrastivo de bloques y secciones recurrentes. ....	348
Tabla 13. Análisis contrastivo de secciones comunicativas en el bloque empresa (C). ....	364
Tabla 14. Análisis contrastivo de secciones comunicativas en el bloque noticias (D).....	376
Tabla 15. Análisis contrastivo de secciones comunicativas en el bloque productos-servicios (E)...	381
Tabla 16. Palabras de mayor frecuencia global en ambos corpus. Estadística realizada con Wordlist. ....	382
Tabla 17. Análisis contrastivo de secciones comunicativas en el bloque legal (F). ....	387
Tabla 18. Términos legales en lista de palabras clave del corpus localizado realizado con la herramienta <i>Keywords</i> . ....	389
Tabla 19. Análisis contrastivo de secciones comunicativas en el bloque zonas de clientes (G). ....	398
Tabla 20. Análisis contrastivo de secciones comunicativas en el bloque interactividad (H). ....	405
Tabla 21. Cuadro resumen contrastivo del grado de convencionalización de la terminología en menús de navegación. ....	418

---

Tabla 22. Tabla contrastiva relativa a frecuencia uso de unidades léxicas convencionalizadas en menús de navegación. ....	419
Tabla 23. Palabras de mayor frecuencia en subcorpus de segmentos textuales en <title>. Estadísticas obtenidas con Wordlist de WordSmith Tools. ....	422
Tabla 24. Listado de palabras en subcorpus de sección contacto (B). ....	442
Tabla 25 . Colocaciones de la palabra «información» en subcorpus de bloque contacto. ....	444
Tabla 26. Formas lingüísticas convencionales asociadas con la colocación «en contacto». ....	446
Tabla 27. Subunidades significativas identificadas en la unidad significativa «formulario de contacto», B-III, y análisis contrastivo de las unidades léxicas en las que se expresan. ....	456
Tabla 28. Tabla contrastiva de la media de palabras en secciones legales (F) por sitio web. ....	469
Tabla 29. Análisis contrastivo de unidades y subunidades significativas identificadas en bloque legal (F). ....	474
Tabla 30. Perfiles textuales característicos en bloque legal (F) en ambos subcorpus. ....	476
Tabla 31. Listado de frecuencia contrastivo por grupos de cuatro palabras en subcorpus legal. ....	477
Tabla 32. Palabras clave en el subcorpus legal localizado utilizando el subcorpus legal original como corpus de control. ....	479
Tabla 33. Formas verbales asociadas a la palabra «aceptación» en la unidad significativa (F-1-II). ....	483
Tabla 34. Tabla contrastiva de las colocaciones de la palabra «protegido» en subcorpus legal. ....	495
Tabla 35. Tabla contrastiva de unidades fraseológicas asociadas a la unidad significativa sobre el titular del sitio web (F-I-IV). ....	500
Tabla 36. Análisis contrastivo de la terminología empleada para el concepto «enlace». ....	506
Tabla 37. Análisis contrastivo de la formas lingüísticas asociadas al lema [responsab*] en subcorpus legal. ....	510
Tabla 38. Formas lingüísticas asociadas al lema [responsab*] exclusivas del subcorpus legal localizado. ....	511
Tabla 39. Tabla contrastiva de colocaciones de «se reserva el derecho» en la unidad (F-3-II). ....	519
Tabla 40. Tabla contrastiva de colocaciones de «se reserva el derecho» en la unidad (F-3-III). ....	521
Tabla 41. Cuadro resumen de formas lingüísticas asociadas a la superestructura del bloque legal (F). ....	528
Tabla 42. Unidades y subunidades con mayor diferencia de frecuencia entre ambos subcorpus legales. ....	530



# 1. INTRODUCCIÓN

Plain text is the foundation of most web information, especially on homepages, which are dominated by navigation and news, both of which usually require words. Other media types are usually best used in conjunction with words, but they do have their place on many homepages (Nielsen y Tahir 2002: 49).

## 1.1. Presentación

Durante los últimos años hemos sido testigos del desarrollo exponencial de la localización, un fenómeno que en la actualidad reclama su estudio pormenorizado desde una perspectiva traductológica (Dunne 2006; Foralón 2006; Bolaños *et al.* 2005; Pym 2004a; Quirion 2003; Gouadec 2003). Este proceso, que engloba a la traducción como eslabón fundamental (Esselink 2001), presenta unas características fruto de un ciclo de producción, distribución y recepción textual radicalmente diferente al de otros procesos de traducción, como la técnica, jurídica, literaria o audiovisual<sup>1</sup>.

La localización web, también llamada *e-localización* (Cronin 2003), se desarrolló adaptando y modelando ciertos procesos ya existentes en la localización de software (Dunne 2006; Yunker 2003). Este proceso ha experimentado un crecimiento exponencial a partir de los años noventa de forma paralela a la expansión de la *World Wide Web*, que a su vez extendió y popularizó el uso de

---

<sup>1</sup> La localización se ha considerado tradicionalmente como una rama de la traducción audiovisual, que incluye además el doblaje y subtitulación. No obstante, ciertas publicaciones (Hurtado Albir 2001; Pacte 2000, 2005, 2007) separan la traducción audiovisual de la localización ya que constituyen modalidades de traducción diferentes, y la denominan «traducción de productos informáticos». Nuestra visión está en consonancia con la propuesta de Hurtado Albir (2001) y la considera una modalidad diferenciada, que denominamos localización, y que engloba varios tipos de traducción como la técnica, la jurídica, la publicitaria o audiovisual.



una nueva forma de organización textual, los hipertextos. En conjunción con los nuevos modelos de comunicación (Janoschka 2003), su impacto en el campo de la traducción es enorme. En general, Internet como medio para distribuir información se considera la invención más significativa desde que Guttemberg creara la imprenta en el siglo XV (Lockwood y Scott 1999: 8; Posteguillo 2003). Inexorablemente, esta revolución se ha trasladado al mundo de la traducción con el surgimiento de una nueva modalidad, la localización web, que está teniendo un profundo impacto en su práctica, investigación y docencia.

## **1.2. Motivación**

El presente estudio se ha fraguado tras varios años de labor localizadora de sitios web<sup>2</sup>, motivado por la esmerada orientación de nuestra directora Dr. Maribel Tercedor, y matizado por los comentarios tras las ponencias basadas en nuestro estudio preliminar (Jiménez 2005a), las discusiones tras ponencias en congresos internacionales que surgieron de la presente tesis (Jiménez 2006b, 2007b, 2007c, 2007d; Tercedor y Jiménez 2007), así como los comentarios de revisión de los artículos ya publicados (Jiménez 2008a, 2008b). En su vertiente didáctica, los resultados de la presente tesis se han aplicado a la docencia de la traducción en el curso de localización web impartido en Wake Forest University, EEUU.

En la motivación de la presente tesis ocupa un lugar primordial que la mayoría de investigaciones en localización partan del sector empresarial (Quirion 2003; Bouffard y Cignon 2006), cuyo objetivo consiste en producir textos localizados que sean recibidos como producciones originales en el seno de la cultura meta: «with the look and feel of locally made products» (LISA 2003: 5). Desde la perspectiva traductológica, este hecho equivaldría a que se ajusten a las convenciones tipológicas, discursivas y de género propias de la cultura meta. La presente

---

<sup>2</sup> El autor está certificado por la American Translators' Association en la combinación En>Es.

investigación surge por la necesidad de dilucidar desde una perspectiva traductológica ciertas cuestiones básicas de interés: ¿qué características textuales, discursivas, pragmáticas o estructurales poseen los textos web producidos originalmente en España a los que se han de parecer los textos localizados? ¿Qué características son las que el sector de la localización aprecia como pertenecientes a la variedad regional o *locale* de España? ¿Cómo se han llegado a establecer? ¿Cómo influye la localización como proceso económico y tecnológico en este proceso de traducción? Finalmente, y puesto que constituye una de las principales aportaciones de la presente tesis, ¿se asemejan realmente los textos localizados a los originalmente desarrollados en España? ¿Se mantienen en éstos las convenciones establecidas en un género específico y representativo?

Tras nuestra revisión sobre investigaciones en localización hemos constatado que la dirección del proceso, y por ende sus análisis previos, parten generalmente desde el contexto sociocultural estadounidense hacia las culturas meta. De hecho, raramente se produce en la dirección contraria, en análisis que partan desde cada cultura y cuyos resultados se reflejen en los textos localizados (Schäler 2002). Este hecho ha motivado una de nuestras hipótesis de partida: un análisis descriptivo que parta del concepto de género ayudará a dilucidar qué elementos discursivos, estructurales, lingüísticos y culturales son prototípicos en los sitios web españoles y qué convenciones de utilidad para el traductor se han consolidado (Jiménez 2006a). Este aspecto se ha investigado en profundidad previamente con objetivos didácticos (Bolaños 2002; Bolaños *et al.* 2005; Rennau 2004), aunque nuestras bases teóricas, objetivos y análisis difieren en gran medida.

Un interés adicional, tras el análisis descriptivo con fines traductológicos y didácticos del género web seleccionado, surge por la escasez de investigaciones contrastivas en el ámbito hispánico (Gamero 2001). Realizaremos un análisis contrastivo empírico que en Traductología constituye:

Una herramienta para dilucidar las diferencias interculturales en el ámbito de las convenciones de género, lo cual tiene aplicación tanto en el ejercicio de la traducción profesional como de la enseñanza de la traducción (Gamero 1998: 422).

Durante los últimos años ha aumentado la popularidad de estudios contrastivos de géneros textuales basados en corpus (Baker 2004: 169). En éstos se han tratado los géneros de mayor importancia en la práctica de la traducción (e. g. Angost 1999; Borja 1999; Gamero 1998, 2001; Bolaños 2002; Vázquez del Árbol 2003; Fernández Sánchez 2004; Sánchez Nieto 2005, 2006)<sup>3</sup>. Desde un punto de vista contrastivo, estos estudios descriptivos investigan las divergencias en culturas diferentes entre un mismo género o géneros cuyas funciones y situaciones comunicativas son similares. Los objetivos de estas investigaciones empíricas suelen ser descriptivos y didácticos, ajustándose a la línea traductológica marcada por Reiss y Vermeer, (1984: 167), Hartman (1980), Hurtado Albir (2001) o Nord (1997b). Partiendo de un enfoque didáctico, esta última investigadora señala la validez e importancia de este tipo de investigaciones:

Teachers of translation mostly have to justify general stylistic conventions by referring to their experience or native-speaker competence. Large corpus-based comparative studies would provide valuable insights (Nord 1997b: 56).

No obstante, nuestro punto de partida queda marcado por el hecho de que hasta la fecha en España no se han realizado estudios empíricos amplios que traten el estudio contrastivo de textos originales y traducidos como base para investigar el proceso traductor. Esta constituiría la segunda línea de investigación propuesta por Hurtado Albir en este ámbito (2001: 505):

La formación de corpus formados por originales y traducciones es también de interés para describir la manera en que funcionan las traducciones y para analizar si al traducir se siguen o no las convenciones estructurales y lingüísticas propias de la lengua de llegada. En

---

<sup>3</sup> Esta lista no pretende ser exhaustiva, y se circunscribe al ámbito de España. Baker (2004) señala que este incremento de los estudios de corpus se observa en gran cantidad de países y combinaciones lingüísticas.

este quehacer, la metodología de corpus electrónicos abre grandes posibilidades (Hurtado Albir 2001: 505).<sup>4</sup>

En la presente tesis situamos estas dos posibles vías de investigación contrastiva en un mismo nivel. El estudio empírico contrastivo de textos traducidos-localizados y de textos originalmente producidos en una lengua y región determinada (*locale*) ayudará a dilucidar en qué aspectos los primeros se asemejan o distancian del prototipo textual compartido por los miembros de la comunidad de discurso meta. La metodología que empleamos constituye una sonda directa hacia el proceso de traducción (Olohan, 2002: 6), y sus resultados ahondarán en aspectos de éste imposibles de obtener con investigaciones contrastivas bilingües. Defendemos este doble nivel de análisis que argumentamos producirá resultados aplicables en dos ámbitos: (1) en la práctica de la localización, desde la descripción del género web más traducido y convencionalizado clarificando en qué aspectos divergen los textos localizados del prototipo establecido en España, y (2) en la didáctica, ya que se proporcionará además una descripción de las principales dificultades que comporta este tipo de traducción.

Dentro del mundo de la localización ya se ha anotado la pertinencia de realizar estudios contrastivos de sitios web, aunque no se especifican los objetivos, marco teórico o metodología apropiados (Foralón 2006; Austerlühl 2006). Además, hemos constatado que el sector de la localización no ha mostrado un interés claro por el uso práctico de corpus textuales. Tal como apunta el investigador Gregory Shreve:

Corpora, for the most part, are under-used and have not been effectively integrated into translation technology or into the internationalization strategies of most companies (Shreve 2006: 313).

En este sentido, el sector de la localización se ha decantado mayoritariamente por el uso de memorias de traducción o traducción automática, junto con una fase de

---

<sup>4</sup> Énfasis propio.

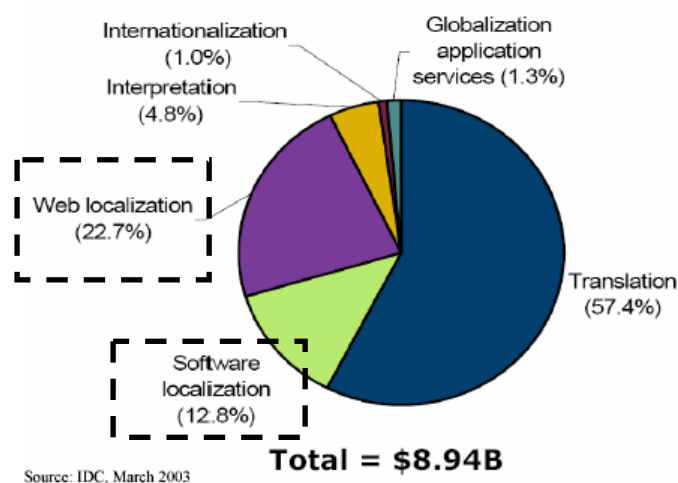
control de calidad de mayor importancia. A nuestro parecer, este proceso conduce en su conjunto a un procesamiento centrado en el microcontexto y que, normalmente, suele obviar aspectos textuales más globales como las diferencias entre géneros y tipologías textuales entre lenguas.

En lo que respecta a la Traductología, durante los últimos años ésta se ha enriquecido progresivamente con un marco de estudio interdisciplinario, que abarca la introducción de la psicología, la lingüística, la sociolingüística, la retórica o la lengua para fines específicos. Fruto de esta cooperación se han incorporado gradualmente conceptos como el *género*, que abordamos en nuestro trabajo ya que puede ayudar al traductor a conocer las limitaciones y opciones que impone cada cultura por cuanto a las convenciones tipológicas, textuales, discursivas y estructurales (Trosborg 1997; Nord 1997a). El conocimiento y la selección del género apropiado mejora la efectividad comunicativa de un texto en una situación social convencionalizada y puede presentar super, macro y microestructuras prototípicas que pueden diferir según la lengua y cultura. A su vez el género, que se define por características externas, impone unas restricciones en la organización y estructura internas de un texto, en la tipología textual. El traductor necesita ser consciente de las diferencias interlingüísticas tanto en un género determinado como en las tipologías textuales en las que se exprese (Gamero 2001). Estos conceptos cobran especial importancia en los textos web, en donde la mezcla de tipologías textuales suele ser la norma (Santini 2007a, 2005a; Tercedor 2005).

### **1.3. Localización y traducción de sitios web**

El desarrollo de la localización de sitios web se produjo adaptando y modelando procesos en localización de software a las particularidades de textos digitales que se distribuyen por Internet, con lo que se considera una rama de la localización con características propias (Yunker 2003: 30; Corte 2002). A pesar de su procedencia,

su importancia queda atestiguada al haber superado en volumen de negocio a la localización de software (LISA 2004; Schäler 2005)<sup>5</sup>. Dado su origen, algunos autores han acuñado los términos *e-localization* (Schäler 2001: 22-26; Cronin 2003: 14), *content localization* (Esselink 2006), o localización de contenidos web (Mata 2005). En la presente tesis emplearemos el término localización web<sup>6</sup>, que a nuestro parecer es suficientemente conciso y claro para nuestros fines.



**Figura 1. Predicción en el volumen de negocio en industrias de la lengua para el 2007. Fuente IDC Research (cf. Schäler 2005).**

Este proceso no se puede contextualizar sin su relación con Internet, cuya importancia como plataforma mundial de comunicación en el siglo XXI ha revolucionado el mundo de la localización y ha sido determinante en la instauración de nuevas prácticas y modelos empresariales (Lommel 2007: 5; Dunne 2006: 6; Esselink 2006). Internet se ha convertido en un modo esencial de comunicación en un mundo globalizado, de ahí el vertiginoso aumento en las cifras de sitios web localizados. Si analizamos las cifras de usuarios, queda patente

<sup>5</sup> Las predicción de [IDC Research](#) para el 2007 señala un 22,7% de volumen de negocio de localización web y tan solo un 12,8% de volumen global para la localización de software (cf. Schäler 2005).

<sup>6</sup> En la situación de la localización dentro de la Traductología, Hurtado Albir (2001) emplea el término «traducción de productos de software» que no incluiría por tanto la localización web. A pesar de las reticencias, el término localización se encuentra a nuestro parecer suficientemente extendido dentro realidad traductológica en España.

que su número sigue creciendo en todo el mundo, aproximándose a mil millones en 2005 (LISA 2004) y más de cien millones de sitios web en la red mundial (El Mundo 2006).

Los textos objeto de estudio se enmarcan en el contexto de la *World Wide Web*, que ha sido definida por su creador Tim Berners-Lee (1999) como «the universe of network-accessible information, an embodiment of human knowledge». De esta concepción como red de información surge el aspecto textual de mayor interés en este medio, la hipertextualidad (véase 3.3). En general los conceptos «web» o «red» e «Internet» se suelen usar de manera indistinta, aunque este paralelismo no se corresponde con la realidad: la WWW constituye una de las diversas situaciones comunicativas en las que se puede usar Internet (Berners-Lee 1999, cf. Crystal 2001). Otros intercambios comunicativos posibles son los chats, correos electrónicos, foros o mensajes SMS, y sus peculiaridades lingüísticas han sido estudiadas ampliamente desde la Lingüística<sup>7</sup> (Crystal 2001; Posteguillo 2003; López Quero 2003), e incluso la Traductología (O'Hagan y Ashworth 2003). No obstante, estas situaciones comunicativas tienen un menor impacto en el sector de la localización.

En lo que respecta a España, el número de personas conectadas a la WWW ha aumentado de manera exponencial, de 10 millones en 1999 según cifras de la Unión Española de Internautas<sup>8</sup>, a 15,3 millones en noviembre del 2007 según Nilsen Netratings<sup>9</sup>. Esta empresa, la más reconocida en estadísticas web, constató que en los últimos años España fue el país con el mayor crecimiento porcentual en número de usuarios del mundo (Nilsen-Netratings 2003). Este aumento en el número usuarios es paralelo al incremento de sitios web en ámbito geográfico

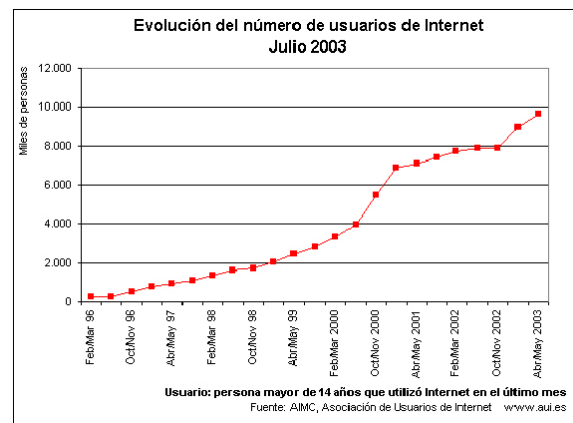
---

<sup>7</sup> Para una bibliografía de las diferentes publicaciones es posible consultar el sitio web de Joaquim Llisteri: [http://liceu.uab.es/~joaquim/general\\_linguistics/internet\\_ling/us\\_llengua\\_internet.html](http://liceu.uab.es/~joaquim/general_linguistics/internet_ling/us_llengua_internet.html).

<sup>8</sup> [www.uai.com](http://www.uai.com).

<sup>9</sup> [www.nielsen-netratings.com](http://www.nielsen-netratings.com).

español<sup>10</sup>. Si nos adentramos en la producción textual en español, sigue igualmente creciendo el número de páginas web activas en este idioma: aproximadamente 17,5 millones según cifras de Nielsen-Netratings<sup>11</sup>. El universo digital en español, aun cuando nuestro estudio está acotado al espacio de España, sería mucho mayor si se incluyeran el resto de variedades regionales del español<sup>12</sup>.



**Figura 2. Evolución del uso de Internet en España. Fuente: Asociación de Usuarios de Internet.**

Estas cifras muestran claramente que Internet se ha consolidado como medio social de comunicación, el objetivo del creador de la Web (Berners-Lee 1999)<sup>13</sup>. Este hecho ha dado lugar a un nuevo modelo de comunicación, que en nuestro caso específico de los sitios web denominaremos Modelo de Comunicación de Masas Interactivo (*Interactive Mass Communication*), un término que adoptamos de las publicaciones de Anna Janoschka (2003: 96 y ss.). En este contexto, el uso de sitios web para las empresas internacionales ya no es una opción, sino que se ha

<sup>10</sup> Como veremos posteriormente, el sector de la localización emplea el término «locale» en este caso. Según el lenguaje HTML, el código para la *locale* de España es lang="es-ES". Es decir, en las páginas dirigidas a España, aun cuando éste no se suele incluir, la World Wide Web Consortium recomienda la inclusión de la etiqueta <HTML lang="es-ES">. Se puede obtener información sobre los códigos de las diferentes lenguas en [www.W3c.org/O-HTML-tags.html](http://www.W3c.org/O-HTML-tags.html).

<sup>11</sup> En abril del 2006.

<sup>12</sup> Con 113 millones de usuarios, el español ocupa el tercer lugar tras el inglés y el chino en número de usuarios en el mundo (Internet World Statistics <http://www.internetworldstats.com/stats7.htm>, consultado el 22 de febrero del 2008).

<sup>13</sup> Según menciona el creador de la web «The web is more a social medium than a technical one» (Berners-Lee 1999: 113).



convertido en un requisito para competir en el mercado internacional (Lommel 2007; Yunker 2003).

Ahora bien, gran cantidad de publicaciones señalan que la mayoría del contenido web se encuentra en inglés. Según los cálculos de LISA (2003), del 62 al 87% de documentos activos en la WWW se encuentran en inglés, aun cuando el 90% de los usuarios prefieren usar una lengua diferente (Tercedor 2005). Diversas publicaciones indican la primacía del inglés en las comunicaciones por Internet (Lommel 2007; Yunker 2003), en parte por el origen estadounidense de Internet que influyó en su codificación (Quah 2005: 165) y que se ha superado con la adopción gradual de Unicode<sup>14</sup>. En el caso del español, por ejemplo, el problema de la codificación supuso que sólo a partir del 2001 se pudieran registrar dominios con signos diacríticos como la «ñ» (Posteguillo 2003: 26)<sup>15</sup>.

La primacía del inglés en las comunicaciones conlleva la adopción de terminología en esta lengua (Aguado de Cea 1994; Carobolante 2001; Belda 2003) y, en cierta medida, de tipos textuales y géneros digitales estadounidenses entre las diferentes comunidades del mundo (Bolaños 2002). En general, partimos de la premisa de que las tipologías textuales y géneros en la web poseen unas características diferenciadas en cada cultura, tal como observamos en nuestra investigación previa sobre la variación terminológica y tipológica entre un corpus comparable de páginas web localizadas al español y páginas producidas originalmente en España (véase 1.3.1). En nuestra opinión, se suele confundir la influencia que ejerce el gran universo web en inglés en registros, géneros, y tipologías textuales digitales con la necesidad de adaptar el texto o contenido web al nuevo medio y modelo de comunicación interactivo (Janoschka 2003: 96). Asimismo, el texto web se debe adaptar a las peculiaridades cognitivas de la lectura en pantalla. En la actualidad

---

<sup>14</sup> [www.unicode.org](http://www.unicode.org)

<sup>15</sup> En septiembre del 2007 se posibilitó la introducción de la totalidad de los signos diacríticos de las lenguas peninsulares en los dominios «.es» (El Mundo, 3-10-2007).

existen gran cantidad de investigaciones en esta área, entre las que ocupan un lugar preponderante las referentes a usabilidad<sup>16</sup> (Nielsen 2006, 1999; Nielsen y Tahir 2002; Adkisson 2002; Brinck *et al.* 2002; Krug 2005). Estos estudios han desarrollado, desde un punto de vista cognitivo, unas pautas básicas que se basan en la interactividad con los usuarios<sup>17</sup> y en ellas se definen los patrones básicos del procesamiento cognitivo de páginas web. Por ejemplo, la lectura en pantalla es 25% más lenta que en papel o la mayoría de usuarios no leen las páginas sino que las escanean (Nielsen 2000: 101). Estas peculiaridades conllevan su consiguiente impacto en el tipo de lengua que se emplea en este medio (véase 5.9).

Puesto que el germen de la presente investigación lo constituyó nuestro estudio piloto (Jiménez 2005a)<sup>18</sup> y su marco teórico, resultados, limitaciones y deficiencias influyeron en la decisión de iniciar la presente tesis, pasamos a describirlo brevemente.

### **1.3.1. Investigación previa sobre la localización web (Jiménez 2005a)**

El núcleo de esta investigación lo constituyó la recopilación de un corpus comparable (Baker 1995) que analizamos empíricamente con técnicas de Lingüística de corpus. La sección localizada del corpus se compuso de 74 páginas web localizadas al español y 84 páginas originales en español. La hipótesis de trabajo inicial consistía en que las páginas localizadas mostrarían estadísticamente unas frecuencias de unidades léxicas y colocaciones diferentes de las páginas españolas, debido a la influencia de la lengua y cultura de origen, el inglés en

---

<sup>16</sup> «Usability» en inglés. La denominación en español de este término no está exenta de discusión. La usabilidad ha sido definida por la norma ISO 9421 (1998) como «the effectiveness, efficiency and satisfaction with which specific users achieve specified goals in particular environments».

<sup>17</sup> De nuevo, usamos la palabra «usuario» para referirnos al receptor del TM.

<sup>18</sup> Trabajo de investigación realizado en el curso de doctorado «Herramientas informáticas y bases para el análisis textual y de procesos cognitivos de la traducción» dentro del programa de doctorado en Procesos de Traducción e Interpretación de la Universidad de Granada e impartido por la Dr. Clara Inés López Rodríguez y la Dr. Maribel Tercedor Sánchez, directora de la presente tesis.

nuestro caso. Esta investigación supuso la introducción en localización de la metodología de corpus comparable en el estudio de la lengua de la traducción (Baker 1992, 1993, 1995, 1998; Laviosa 1997; Kenny 1998, 2001). Esta corriente se basa en que la traducción «is a communicative event which is shaped by its own goals, pressures and context of production» (Baker 1996: 175), por lo que los textos fruto de este proceso pueden mostrar rasgos lingüísticos distintivos que observables con técnicas de análisis de corpus (Zanettin 2000).

El análisis empírico se realizó sobre la terminología empleada en los menús de navegación, un elemento básico y convencionalizado presente en todo sitio web. En particular, se estudiaron las unidades léxicas altamente convencionalizadas en las páginas web anglosajonas, *contact us*, *about us*, *privacy policy*, *home* y *terms of use* (Nielsen y Tahir 2002). En los resultados observamos diferencias significativas en las características discursivas y tipológicas entre ambas secciones del corpus, que atribuimos a los distintos contextos de producción de ambos conjuntos textuales. Las principales diferencias identificadas se centran en la frecuencia de uso, la variación terminológica, el tenor digital o la aparición de unidades léxicas localizadas inexistentes en el corpus original español (Tirkkonen-Condit 2004; Kenny 1999; Gellerstam 1996). Como ilustración de los resultados obtenidos presentamos una de sus gráficas, en la que podemos observar las frecuencias de las distintas unidades léxicas que se corresponden con el concepto representado por la unidad léxica convencionalizada en inglés *contact us*. Las barras azules corresponden al corpus de páginas originalmente desarrolladas en España, y las rojas al corpus localizado al español.

Los resultados muestran que las distintas unidades léxicas presentan diferencias de frecuencia entre ambas secciones del corpus, de ahí que los resultados fueran son congruentes con los de investigaciones traductológicas similares anteriores (Gellerstam 1996; Laviosa-Baithwate 1996). Podemos observar además que el

término de mayor frecuencia en la sección traducida, «contáctenos», refleja la influencia del término original *contact us*, altamente convencionalizado en el contexto estadounidense con un 89% de uso (Nielsen y Tahir 2002).

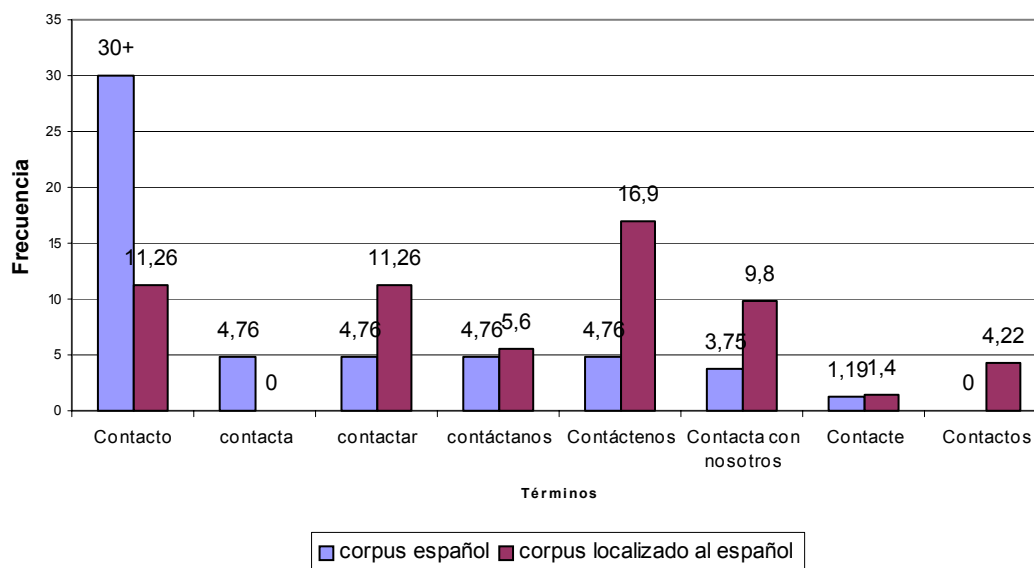


Figura 3. Frecuencia de uso en corpus original y localizado. De Jiménez (2005a).

En lo que respecta a la variación terminológica, el análisis contrastivo reveló cómo existían unidades léxicas como *privacy policy* con un bajo grado de fijación en español, y este hecho repercutía en una gran variación en sus formas localizadas. Las variación aparece tanto en el primer componente de la unidad léxica, *privacy*, con «privacidad» o «confidencialidad», como en *policy*, con «cláusula», «política», «declaración», o «compromiso». En este caso, atribuimos la baja fijación de los distintos términos que responden a este concepto por el diferente grado de convencionalización de este bloque comunicativo en un mismo género en culturas diferentes, la anglosajona y la española. El desarrollo natural de estas conclusiones supuso la introducción en la presente investigación del concepto de género en el estudio contrastivo de la textualidad web. Asimismo, su introducción supuso la concepción del sitio web como unidad textual de análisis en Traductología, superando así el uso de páginas web como unidad de análisis traductológico al no

constituir una unidad operativa en la que recaiga el proceso de traducción (véase 3.4).

### **1.3.2. La localización web y su importancia en Traductología**

El aumento de la importancia de los sitios web como plataforma de comunicación conlleva la necesidad de establecer criterios de calidad, un aspecto que se recoge en el conjunto de publicaciones sobre diseño y desarrollo web (Brinck *et al.* 2002). Esta necesaria calidad se traslada igualmente a sus subsecuentes localizaciones, ya que la calidad en la traducción de todo texto emitido<sup>19</sup> por la empresa contribuye al reforzamiento positivo de su imagen (Jeney 2007; Gamero 2001; Bolaños 2002a).

Como veremos en el siguiente capítulo, la localización es parte de un proceso económico más amplio que incluye la globalización e internacionalización (véase 2.2), y presenta diferencias dependiendo del «grado de localización» (véase, 2.7). En este concepto encontramos ciertos paralelismos con los conceptos funcionalistas de *escopo* (Reiss y Vermeer 1984) y el «encargo» de la traducción (Nord 1997), ya que la necesidad de calidad y acercamiento a las expectativas del usuario en este medio son fundamentales. De hecho, el icono más usado durante la navegación en este medio tan competitivo es «Atrás», puesto que los usuarios han de encontrar de manera efectiva la información que necesitan. En caso contrario abandonan el texto web y siguen buscando (Nielsen 2000; Nielsen y Loranger 2006). Sirva como ejemplo la media de tiempo por página web en España, 52 segundos<sup>20</sup>. Los manuales de usabilidad y guías de estilo web remarcan la importancia de que las páginas posean una estructura clara y concisa que permita su escaneado (Nielsen 2006; Nielsen y Tahir 2002), y que se ajusten a las convenciones discursivas, de registro, y tipológicas que espera el usuario (*ibid.*).

---

<sup>19</sup> Como son los textos técnicos o publicitarios.

<sup>20</sup> Según Nielsen-Netratings en diciembre del 2007. En todo el mundo la media es de 46 segundos. De [www.nielsennetratings.com](http://www.nielsennetratings.com).

Estos aspectos se relacionan con potencial éxito de todo sitio web localizado, entendido como la capacidad de atraer y mantener unos niveles de audiencia adecuados:

One key benchmark of Web success is stickiness-the ability to attract new and repeat visitors and keep them on a site (LISA 2004: 35).

Se respalda así la importancia de una localización de calidad, el fin último que el sector prima en este proceso junto con la necesidad de que los textos localizados se perciban como creaciones originales en cada región (LISA 2004: 13). Recogemos las palabras de la asociación internacional en la materia:

Localization needs to go beyond language questions to address issues of content and look and feel with the ultimate aim of releasing a product that looks like it has been developed in country<sup>21</sup> (LISA 2004: 11).

Esta posición no está libre de críticas (Schäler 2002; House 2001), puesto que aunque se localice de manera adecuada el sitio web con cambios de colores, iconos, etc., el contenido suele ser el mismo: materiales globalizados que ingenuamente se suponen culturalmente neutros (Schäler 2002). Juliane House asocia esta noción de textos culturalmente neutros con la invasión e infiltración de convenciones culturales y lingüísticas anglosajonas en todo mundo (2001: 253-254). En este proceso de «universalización», la distribución de textos web localizados por Internet desempeña un papel clave.

De igual manera, la calidad y conformidad con las expectativas del receptor son de gran importancia en otro concepto clave, la credibilidad (Posteguillo 2003: 154). La importancia de que el texto meta se ajuste a las convenciones de género que esperan los usuarios constituye un aspecto primordial, puesto que de no ser así, el texto corre el riesgo de que no se perciba como verídico o creíble y, por lo tanto, no sea funcional en la cultura meta. Este aspecto ha sido estudiado por los analistas

---

<sup>21</sup> Énfasis propio.

de captología, que incluso han creado unas «Guidelines for Web Credibility»<sup>22</sup> (Frogg 2002; cf. Posteguillo 2003:154).

No obstante, la corriente en Traductología que se centra en las peculiaridades de los textos traducidos postula mediante estudios empíricos de corpus que estos textos presentan características y frecuencias atípicas si se comparan con textos originales en cualquier idioma (Baker 1992, 1993, 1995, 1998, 1999; Kenny 1998, 2001; Laviosa 1998, 2002), aun cuando éstos se aprecien como originales por los receptores. Una de nuestras premisas de partida se basa, por consiguiente, en que este *desideratum* de producir textos semejantes a los producidos en la región lingüística de destino es iluso en cierta manera, puesto que los textos localizados son el resultado de un proceso de producción textual diferenciado (Baker 1998). En este sentido, hasta la fecha no existe investigación alguna sobre este aspecto en localización y esta será una de las aportaciones básicas de la presente tesis.

Nuestro estudio surge además de la necesidad de que la localización sea investigada en un proceso de abajo-arriba, desde el seno de cada cultura o región meta hacia el hegemónico sector anglosajón (Schäler 2002). Tradicionalmente, la localización ha supuesto un proceso de arriba-abajo, que parte desde empresas mayormente estadounidenses con textos producidos en inglés hacia cada cultura. Esta dirección repercute en los textos localizados, ya que el proceso global no se nutre necesariamente de las peculiaridades textuales propias de cada género en culturas diferentes. Sirva como ejemplo nuestra investigación previa (Jiménez 2005a), en la que identificamos términos localizados que reflejan directamente su origen anglosajón y que, por consiguiente, constituirían extranjerismos léxicos y sintácticos como «sobre nosotros», «cláusula de privacidad», «declaración de privacidad» o «términos y condiciones». De hecho, estas formas léxicas no se identificaron en el corpus de 79 páginas de inicio de sitios web corporativos

---

<sup>22</sup> Creadas en la University of Stanford. Accesible a través de [www.webercredibility.org/guidelines](http://www.webercredibility.org/guidelines).

españoles. Los resultados que obtendremos en la presente tesis ampliarán estos resultados previos.

El proceso traductor en localización se encuentra inmerso en un ciclo global más amplio (Dunne 2006), que por su posible influencia en la actividad traductora y su producto, hemos de explicitar. Por consiguiente, nos es ineludible revisar cómo se produce y qué peculiaridades posee el proceso de localización en profundidad, aspecto que desarrollaremos en el capítulo 2.

### **1.3.3. Los sitios web corporativos como objeto de estudio**

Los sitios web en general suponen una de las situaciones comunicativas que surgieron con el advenimiento del nuevo medio, Internet, y son uno de los posibles modos de comunicación en los que se han centrado investigaciones traductológicas<sup>23</sup> (Mata 2005; O'Hagan y Ashworth 2003; Días Fouçes 2003; Bolaños *et al.* 2007, 2005; Bolaños 2002).

Dentro entre los múltiples géneros digitales, y para centrar nuestro estudio, seleccionamos los *sitios web corporativos*<sup>24</sup> por constituir un género representativo. Su selección se realizó de acuerdo con dos principios básicos: (1) en primer lugar, de entre los distintos textos web que existen, los sitios web corporativos se han consolidado como género y, consecuentemente, se han convencionalizado gran cantidad de sus características (Askehave y Nielsen 2005; Adkisson 2002). Las investigaciones sobre géneros web han comprobado que los sitios corporativos presentan un estilo y forma más definidos que otros géneros

---

<sup>23</sup> Dentro de los posibles modos de comunicación definidos por el medio Internet que pueden tener impacto en la Traductología encontramos los chats, correo electrónico, mundos virtuales, listas de correo etc. (O'Hagan y Ashworth 2003).

<sup>24</sup> La Real Academia Española sugiere los términos «cibersede» o «sede electrónica» en este caso. Hemos seleccionado el término «sitio web corporativo» por ser más claro y conciso para nuestros objetivos.



web, como los sitios web personales<sup>25</sup> o los institucionales (Kennedy y Shepherd 2005) y, (2) la elección se basa en el alto volumen de traducción en esta área (Göpferich 1995). La localización de sitios web presenta el mayor crecimiento en la actividad de traducción (Lockwood 1999), y a su vez, dentro de las traducciones de sitios web, nos limitamos a los **corporativos** puesto que la presencia de las empresas en la WWW se ha convertido en un imperativo (Dunne 2006; Schmitt 2000), al haberse constituido en el principal medio de información sobre sus productos, servicios, y sobre la empresa en sí.

La presente tesis se centra así en este género web representativo altamente convencionalizado, los «sitios web corporativos» o *corporate homepages* en inglés, que en nuestra opinión constituyen un punto de partida óptimo para la investigación en localización web. Preliminarmente, este género se podría definir como<sup>26</sup>:

Corporate homepages were defined as web pages describing the interests and ambitions of companies whose purpose for existing is to make profit through selling some product or service (Shepherd *et al.* 2004: 237).

Según esta definición, el género de estudio estaría representado por sitios web que describen los intereses, actividades y objetivos de una empresa cuya existencia se dirige hacia la obtención de un beneficio económico mediante de la venta de productos o servicios. Estas empresas pueden tener presencia exclusivamente en Internet, como [www.amazon.com](http://www.amazon.com), o puede tratarse de una empresa consolidada con presencia física, [www.iberdrola.es](http://www.iberdrola.es) (Singh y Pereira 2005: 2). En los capítulos 4 y 5 de la presente tesis revisaremos en profundidad las clasificaciones textuales en

---

<sup>25</sup> Dillon y Gushrowski (2000) defienden que los sitios web personales serían los primeros géneros digitales en establecerse como tales.

<sup>26</sup> Las definiciones previas de «corporate homepage» que hemos identificado suelen diferir: Liu *et al.* (1997) la define como «a document on a web server that can be referenced under the corporate name and that is developed and maintained by the corporation and serves as an entry point for customers to access specified information utilizing an electronic medium» (Liu *et al.* 1997: 336).

Traductología y describiremos detalladamente el marco teórico adoptado a este respecto.

#### 1.4. Hipótesis de trabajo

El presente trabajo parte de cuatro hipótesis de trabajo:

- i. En primer lugar, que cada género impone unas restricciones a todos los niveles textuales que pueden divergir entre lenguas y culturas diferentes, con lo que este hecho repercutirá el proceso de localización.
- ii. En segundo lugar, el proceso de localización impone unas restricciones en la labor traductora sobre todo debidas al contexto tecnológico y al acceso al texto por parte del traductor, y este aspecto podrá observarse por medio de un análisis empírico de corpus.
- iii. En tercer lugar, el género web que constituyen los «sitios web corporativos», o *corporate homepage* (Kennedy y Shepherd 2005), se encuentra convencionalizado en España y los receptores llegan al mismo con unas expectativas en lo que se refiere a la superestructura, macroestructura, elementos intratextuales o funcionalidad<sup>27</sup>, que en principio el traductor debería respetar y mantener para producir un texto plenamente funcional en la cultura meta.
- iv. En cuarto lugar, y en consonancia con Baker (1995, 1996, 1998), las localizaciones son el resultado de un proceso de producción textual sujeto a una condiciones específicas y un contexto de producción diferenciado que resultará en textos con ciertos aspectos cuantificables empíricamente por

---

<sup>27</sup> Este concepto lo entendemos como la interacción entre la máquina y el usuario (véase 5.5.6).

medio de la metodología de la Lingüística de corpus. Entre éstos aspectos, los textos localizados mostrarán un fenómeno de asimilación a las convenciones de género establecidas en el contexto sociocultural de partida (Reiss y Vermeer 1984: 195).

## 1.5. Objetivos

Para validar estas hipótesis de estudio se presentan los siguientes objetivos que se investigarán empíricamente sobre la base de un corpus comparable (Baker 1995) representativo de sitios web corporativos españoles por un lado, junto con sitios web corporativos localizados desde un único contexto sociocultural y geográfico, EEUU, hacia otro conformado por España. Los objetivos entroncan con la necesidad existente en Traductología de estudios descriptivos (Toury 1980: 81) y se basan en poner a disposición de los investigadores, profesionales, y docentes de la traducción análisis contrastivos de ciertos géneros textuales que muestren las convenciones respectivas dentro de un par de lenguas específico. Más concretamente nuestros objetivos serían:

- a. Revisar las publicaciones sobre localización para observar la influencia, cambios y restricciones que este proceso impone en la tarea de la traducción.
- b. Delimitar qué constituye un «texto web» desde una perspectiva traductológica.
- c. Revisar la bibliografía sobre tipologías textuales en Traductología para proponer un marco teórico de estudio que dé cabida a hipertextos como los sitios web y, consecuentemente, garantizar la comparabilidad de los textos

recopilados y su estatus como unidad en la producción, recepción y traducción.

- d. Delimitar, desde un enfoque interdisciplinario, qué constituye el género *sitio web corporativo*.
- e. Desarrollar las bases metodológicas para la recopilación de un corpus comparable de textos web completos naturales tanto en español como localizados hacia España, introduciendo el concepto de *hipertexto* (Storrer 2002) o sitio web completo como unidad de análisis.
- f. Realizar un estudio descriptivo empírico de este género web en el que se obtengan las convenciones que se han asentado a diferentes niveles, para proporcionar un análisis que sea de interés a profesionales, investigadores y docentes en esta área.
- g. Realizar un estudio descriptivo comparativo de ambas secciones del corpus para investigar si al traducir este género representativo determinado se siguen o no las convenciones extra e intralingüísticas que se han identificado en la primera fase de análisis.

## **1.6. Estructura de la tesis**

La presente tesis se ha estructurado en consonancia con los tres niveles que ha de presentar toda investigación empírico-experimental en Traductología: el nivel conceptual, metodológico y analítico (Hurtado Albir 2001:190). Esta estructuración sigue por tanto el proceso heurístico característico de las investigaciones científicas y se presenta de modo cíclico, con la modificación de

hipótesis a través de las distintas etapas a lo largo del tiempo durante las investigaciones previas.

1. El primer nivel, o nivel conceptual, queda recogido en la primera parte de la presente tesis, que gira en torno al marco teórico desarrollado. Los capítulos correspondientes a esta fase tratan los cuatro conceptos básicos en nuestra investigación mediante una revisión crítica de la bibliografía pertinente sobre: (1) el proceso de localización (2), la noción de texto en este proceso (3), la introducción de tipologías textuales en Traductología (4) y la introducción del concepto de género como herramienta en la construcción y análisis contrastivo de nuestro corpus monolingüe comparable (Baker 1995). Como último punto en nuestro marco teórico revisaremos (5) la Lingüística de corpus y su introducción en investigaciones empíricas en Traductología.
2. El segundo nivel, o nivel metodológico, queda recogido en la segunda parte de la tesis, en la que expondremos la metodología de la investigación que hemos adoptado. Esta etapa queda recogida en el capítulo 7, en el que expondremos los criterios básicos de recopilación del corpus monolingüe comparable: la representatividad, comparabilidad, la estandarización y la tipología, y estableceremos los criterios para la identificación plena de los textos como originales o producto de un proceso de localización-traducción. Finalmente, se presentan las distintas etapas en el análisis y las herramientas de software empleadas para llevarlo a cabo.
3. En tercer nivel, o nivel analítico, se corresponde con la tercera parte en la que analizaremos y discutiremos los resultados obtenidos, para pasar a contrastar las hipótesis. Este nivel se corresponde con cuatro capítulos en la tesis. En el capítulo 8 se recoge el análisis extralingüístico contrastivo del género objeto de estudio, ya que se procederá desde el contexto extralingüístico más amplio

hacia los aspectos intratextuales. En este capítulo se revisan contrastivamente los factores pragmáticos, los factores comunicativos entre los que se incluyen las estadísticas de localizaciones hacia otras lenguas y el análisis contrastivo del tono o tenor digital en este género, el condicionamiento sociocultural en el que se incluye el marco normativo y jurídico, y finalmente el marco socioprofesional. Los capítulos 9 y 10 contienen el análisis contrastivo intralingüístico, que parte de la superestructura textual por representar el mayor rango jerárquico en este nivel, pasando a analizar los distintos bloques y secciones de este género que se corresponden con las distintas áreas identificadas en los sitios web corporativos. En el análisis de las formas lingüísticas convencionales nos centramos en la terminología empleada en los menús de navegación, ya que estos segmentos textuales presentan los mayores niveles de convencionalización en este género. El análisis superestructural se centra además en la superestructura interna de cada página web, analizando contrastivamente aquellos segmentos textuales determinados por los estándares de base (Wright 2006), tales como el título de la página o los segmentos textuales en los atributos <alt>. El capítulo 10 está dedicado al análisis contrastivo intratextual adicional de las formas lingüísticas convencionales en los dos bloques comunicativos que, tras el análisis de los resultados obtenidos, poseen un mayor interés: las páginas de contacto y las páginas de información legal. Por último, en el capítulo 11 se analizan los tipos de errores más recurrentes en los sitios web localizados, ya que las peculiaridades de la multitud de errores identificados les confiere el carácter de marcadores para los textos producto de un proceso de traducción. En especial, se analiza la relación entre el nivel de convencionalización de ciertos elementos en el género y la mayor o menor presencia de errores.

4. La cuarta parte queda reflejada en el capítulo 12, dedicado a las conclusiones, en el que se revisan las hipótesis iniciales con el fin de contrastarlas con los resultados y se exponen las futuras vías de investigación.
  
5. Finalmente en la quinta parte, que se corresponde con los capítulos 12 y 13, se presenta un listado de la bibliografía consultada en la preparación de la presente tesis, así como los distintos apéndices en los que se incluyen entre otros las composiciones finales de las dos secciones del corpus comparable o distintas listas de frecuencia léxica obtenidas.

**PARTE I:**

**MARCO TEÓRICO**





## 2. LA LOCALIZACIÓN COMO OBJETO DE ESTUDIO

En la presente investigación nos es esencial delimitar cómo se desarrolla el proceso de localización, qué aspectos lo diferencian de otros procesos de traducción, y su posible impacto en su producto, los textos localizados. Este proceso suele seguir un ciclo relativamente bien definido (LISA 2004:15), aunque debido a las diversas perspectivas de estudio no existe aún un consenso claro sobre qué constituye la localización en sí (Dunne 2006; Foralón 2006)<sup>28</sup>. Las diferentes definiciones de localización provienen en su mayor parte del sector, que no de la Traductología, y en las mismas se observa esta multiplicidad de perspectivas: (1) desarrolladores informáticos, (2) agentes económicos del sector, (3) traductores profesionales o (4) investigadores en Traductología<sup>29</sup>. Una primera aproximación a las definiciones muestra ciertos denominadores comunes del proceso:

- a) La separación entre un componente cultural y otro lingüístico (Pym 2004a; LISA 2003; Esselink 2000; O'Hagan y Ashworth 2003; Quirion 2003; Schäler, 2001).
- b) La denominación de producto a la unidad de localización (LISA 2003; Yunker 2003; Esselink 2001; Schäler 2001).
- c) La dirección hacia una *locale* o «región» meta (Pym 2004a; LISA 2003; Esselink 2000; O'Hagan y Ashworth 2003; Comité sectoriel de l'industrie

---

<sup>28</sup> Dunne (2006) señala de entre las perspectivas de estudio la de (1) los desarrolladores de software, (2) agentes económicos del sector, (3) traductores o investigadores en Traductología.

<sup>29</sup> Dunne se refiere a esta multiplicidad de perspectivas de la siguiente manera «there exists no consensus as to what precisely constitutes localization, due in large part to the ways in which it is perceived» (Dunne 2006: 1). En esta misma línea, y adhiriéndose a la misma postura, Debbie Foralón (2006) reseña que las definiciones de localización dependen de la perspectiva desde la que partan y reflejan la distinta adscripción del que las formula.

canadienne de la traduction 1999, cf. Quirion 2003; Dr. International 2003).

- d) La inclusión de aspectos técnicos (LISA 2004, 2007; Dunne 2006).
- e) La frecuente omisión de la palabra «traducción» (Lommel 2007; LISA 2003; O'Hagan y Ashworth 2003; Yunker 2003)<sup>30</sup>.

A nuestro parecer, este último aspecto reside en la importancia que el sector otorga a una compleja faceta técnica que no suele presentarse en otros tipos de traducción<sup>31</sup>, y surgió en los albores de la localización debido a la voluntad de destacar ciertas adaptaciones técnicas específicas. En este contexto, el sector comenzó a separar la localización de una concepción arcaica de la traducción (Uren *et al.* 1992) que, como veremos, quedaba reducida en su visión a una mera equivalencia entre palabras. Aún hoy en día, y a pesar del avance exponencial de los Estudios sobre Traducción, el sector de la localización sigue concibiendo la traducción en parte como una simple comunicación de significados y mensajes entre palabras. Esta concepción queda patente en la siguiente cita:

This process came to be known as *localization*, to distinguish it from *translation*, which focused on communicating the meanings and messages of words<sup>32</sup> (Lommel 2007: 5).

La Traductología contemporánea incluye a la localización como una modalidad de traducción diferenciada (Hurtado Albir 2001: 94), y se asume que conjuga unas características propias de ahí se acepte ampliamente la denominación de «localización» y no exclusivamente traducción (Pym 2004a: 26). A pesar de que en Traductología la engloba y queda circunscrita como traducción marcada por el modo, el sector de la localización, tal y como se desprende de la cita anterior, sigue haciendo hincapié en la necesidad de separarlas (Lommel 2007: 7).

---

<sup>30</sup> Excepto en la definición de Bert Esselink (2000), Dr. International (2003) y Chandler (2004).

<sup>31</sup> Varias definiciones provenientes del sector mencionan expresamente el término «traducción» (Chandler 2004; Dr. International 2003; Esselink 2001).

<sup>32</sup> Énfasis propio.

En general, nos parece más productivo considerar que la localización se clasifique por su modalidad, puesto que no se podría clasificar exclusivamente por un tipo de traducción, entendido como la traducción de textos que pertenecen a determinados ámbitos socioprofesionales que por tanto se asocia a tipos y géneros textuales específicos (Hurtado Albir 2001: 58). Por su parte, la localización web va más allá de la traducción de textos técnicos en la que a veces se sitúa, e incorpora la traducción de textos jurídicos, publicitarios, subtitulación de presentaciones o videos, etc.

## 2.1. El origen de la localización

La localización<sup>33</sup> comenzó en los años 80 centrada en los productos de software y hardware (Esselink 2006)<sup>34</sup>, y se fue extendiendo a otros campos como la localización de sitios web, interfaces de productos de electrónica, videojuegos, móviles, etc. (Esselink 2001: 5 y ss). Históricamente, los primeros pasos de la localización se produjeron en el seno de la traducción, aunque este proceso fue diferenciándose al constatarse ciertos problemas técnicos durante las primeras traducciones de programas informáticos<sup>35</sup>. Estos supusieron la creación de un entorno de trabajo colaborativo con programadores (Dunne 2006) y, consecuentemente, fue inevitable desarrollar un proceso cíclico más amplio que englobara tanto a traductores como a desarrolladores.

La localización se circunscribía en sus orígenes primordialmente al área geográfica de los EEUU, por lo que en sus albores el proceso se desarrollaba casi exclusivamente del inglés hacia otras lenguas<sup>36</sup> (Uren *et al.* 1992). Sin embargo,

---

<sup>33</sup> El sector ha acuñado la abreviación L10n, y se debe a que existen 10 letras entre la «l» y la «n» en la palabra localización.

<sup>34</sup> Para una revisión sobre el origen y desarrollo de la localización, consúltese el artículo «The evolution of Localization» (Esselink 2006).

<sup>35</sup> Tal como los formatos erróneos en las fechas o los problemas con cadenas concatenadas de texto que los traductores no podían resolver.

<sup>36</sup> Este aspecto supuso que la localización inicialmente se redujera a los traductores dentro de la combinación de lenguas del inglés hacia otras lenguas, y no entre francés y español, por ejemplo.

desde los años 90 surgió paulatinamente la corriente opuesta, con procesos de localización hacia el inglés centrados principalmente esencialmente en textos web (Lommel 2007). Nuestros resultados atestiguan la importancia de esta dirección inversa, con un 46% de sitios web españoles con una versión localizada hacia esta lengua mayoritaria en Internet (véase 8.1.2.1).

El crecimiento exponencial de este procesos durante los últimos años (LISA 2004:18)<sup>37</sup> ha supuesto que hoy en día el sector de la localización, o *Localization Industry* en inglés, se compone no sólo de *Localization Service Providers* y desarrolladores de software/hardware, sino que además se considera parte del sector a los centros de formación universitarios y al sector de las telecomunicaciones (Quah 2005).

## 2.2. G.I.L.T.

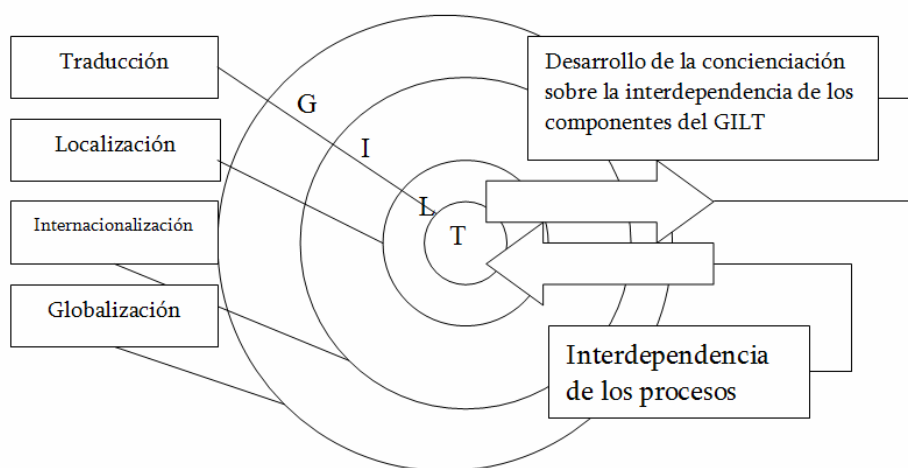
Para acotar el proceso de localización, en primer lugar tenemos que definir el conjunto de procesos interdependientes conocido con el acrónimo GILT: Globalización, Internacionalización, Localización y Traducción (Cadieux y Esselink 2002). La pertinencia de delimitar y definir estos procesos se debe a que su ausencia puede afectar directamente a la tarea de traducción durante la localización. Existe cierta confusión general entre los objetivos de estos procesos interdependientes, confusión que se extiende asimismo al plano terminológico. No obstante, la terminología empleada se está normalizando recientemente por los esfuerzos de ciertas asociaciones que han dotado al campo de la coherencia conceptual necesaria<sup>38</sup>. Sirva como ejemplo de dicha confusión que Microsoft denomina «globalización» al concepto que se considera «internacionalización» en el sector (Cadieux y Esselink 2002). La aparición de estos dos procesos surgió por la

---

<sup>37</sup> Según los cálculos de LISA el volumen anual de negocio es de 3.700 millones a 5.000 millones de dólares anuales.

<sup>38</sup> Como puede ser la Localization Industry Standard Association (LISA).

pertinencia de realizar cambios y adaptaciones tanto en la organización empresarial como en el desarrollo técnico de los productos de software. Estas dos vertientes dieron lugar a la globalización y la internacionalización respectivamente. La aplicación efectiva de estos dos procesos resultó en un ahorro considerable de tiempo y recursos para las empresas dedicadas a la localización de software (Brooks 2000).



**Figura 4. Interdependencia de los procesos GILT y desarrollo de la concienciación sobre las interdependencias de los GILT. Adaptado de Dunne (2006: 5).**

El proceso GILT se inicia con la globalización y culmina en la traducción. En virtud de esta dirección, Anthony Pym (2004: 30) señala que sólo existe un proceso general, la globalización, siendo la internacionalización y la localización partes integrantes del mismo. A pesar de esta dirección, coincidimos con el investigador Keiran Dunne (2006) en que, independientemente del desarrollo actual de los procesos GILT, desde una perspectiva traductológica los mismos evolucionaron en la dirección opuesta; el único proceso primigenio fue la traducción, y las exigencias impuestas por este proceso culminaron con la aparición de la globalización, el proceso más general del conjunto. La Figura 4 recoge las relaciones bidireccionales en el proceso global.

El proceso de traducción en localización se encuentra inmerso en estos otros procesos, aunque la internacionalización es la que facilita directamente el proceso de localización. En el caso de la localización web, su menor complejidad técnica (Esselink 2006) puede conllevar la inexistencia de estos procesos interrelacionados, aunque este hecho dificultaría indudablemente la labor de los traductores. Esta interdependencia de estos procesos GILT queda también patente en el siguiente gráfico publicado por LISA (2004:15). El mismo presenta un modelo práctico que recoge de forma cíclica la dirección del proceso durante su implementación, que no el desarrollo de concienciación de los mismos presentado en la figura previa:



Figure 1. The global product development cycle.

Figura 5. Esquema del ciclo de desarrollo de un producto global según LISA (2007:19).

Ésta no sería la única propuesta que se ha presentado, aunque sí que constituye la más aceptada. Si retornamos a la confusión terminológica y conceptual que aludimos previamente, en la propuesta presentada por Microsoft el proceso global de localización se denomina «internacionalización», y de ahí se separa la localización de la *world-readiness*, que en el esquema anterior comprendería la internacionalización y la globalización (Dr. Internacional 2003). A su vez, la *World-readiness* se compondría de la globalización y la «habilitación de la

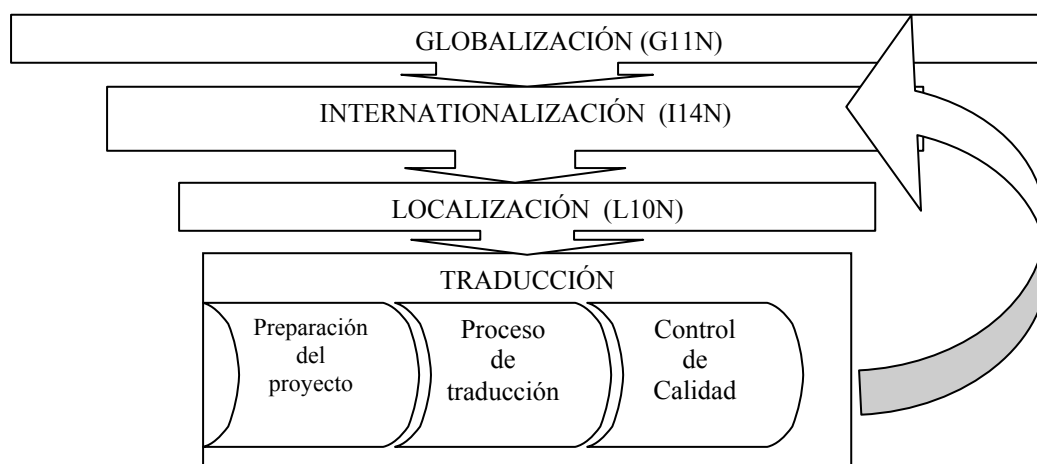
localización» o *Localization enablement*, que se corresponde con la fase técnica de internacionalización mencionada anteriormente (Esselink y Cadieux 2002).



**Figura 6. Proceso global de globalización según Microsoft (Dr. International 2003).**

A pesar de las aparentes diferencias entre los modelos presentados, a nuestro parecer es incuestionable la existencia de una fase técnica de preparación de la localización, que denominaremos «internacionalización» (véase 2.2.2), una base global que trata aspectos organizativos empresariales denominada «globalización» (véase 2.2.1), y finalmente, la «localización» (véase 2.3), en la que quedaría inmerso el proceso de traducción. Ésta a su vez se subdivide a grandes rasgos en la preparación del proyecto, el proceso de traducción y el control de calidad (Quah 2005:114).

Pasamos a tratar en mayor detalle cada una de estas fases interdependientes que contextualizan el proceso de localización.



**Figura 7. Propuesta de interdependencia de los procesos GILT en los que se enmarca la traducción.**



### 2.2.1. Globalización

La globalización<sup>39</sup> constituye el proceso más amplio y se centra primordialmente en la organización empresarial, definiéndose por sus objetivos. Según LISA (2007):

Globalization [...] refers to all of the business decisions and activities required to make an organization truly international in scope and outlook. Globalization is the transformation of business and processes to support customers around the world, in whatever language, country, or culture they require (LISA 2007:1).

La globalización se centra en organizar la estructura empresarial y los recursos existentes para que sea posible ofrecer productos-servicios o atender a cualquier cliente independientemente de su lengua, región o cultura. Esta definición ha evolucionado desde las primeras publicaciones de LISA (2003), observándose un esfuerzo generalista que incorpora todo tipo de interacción entre empresa y cliente en cualquier lengua. En este sentido, la definición de globalización se aproxima al concepto de globalización económica<sup>40</sup>. Esta evolución queda patente si se contrasta con la definición previa de esta misma asociación, de afán más universalista, puesto que el objetivo anterior de la globalización consistía en la facilitación exclusiva del proceso de localización:

...making all the necessary technical, financial, managerial, personnel, marketing, and other enterprise decisions necessary to facilitate localization (LISA 2004: 14).

En la actualidad, la globalización constituye un proceso más amplio que engloba a la propia localización, y en el plano temporal se produce tanto antes de la localización, que facilita, como después de la misma, en la distribución de productos o asistencia multilingüe al cliente (Lommel 2007).

---

<sup>39</sup> El sector emplea el sigloide, *glln*, que al igual que los anteriores se debe a la existencia de once letras entre la «g» y la «n».

<sup>40</sup> Desde una perspectiva económica el controvertido concepto de la globalización se entiende como: «The ever-increasing integration of national economies into the global economy through trade and investment rules and privatization, aided by technological advances» (Hinges 2000: 5).

En el ámbito de la localización web, hemos constatado que el término «globalización» se ha identificado en ciertos estudios con la aspiración de un sitio web determinado de alcanzar la mayor audiencia posible, algunos autores llegan incluso a denominar «globalización web» a la localización de sitios web (Yunker 2003).

### **2.2.2. Internacionalización**

La internacionalización<sup>41</sup> se centra en la parte técnica del proceso. En general, consiste en el conjunto de procedimientos que garantizan que (1) un documento o aplicación no esté limitado culturalmente, (2) que sea independiente de la lengua de desarrollo, normalmente el inglés (Lommel 2007: 28)<sup>42</sup>, y (3) que sea fácil de localizar sin que se requieran modificaciones técnicas a gran escala (Tercedor 2005). Este proceso se realiza durante la fase de desarrollo técnico, y aunque el sector no incluye la participación expresa del traductor en esta etapa, no podemos obviar el hecho de que su intervención como mediador intercultural es indispensable a la hora de establecer qué adaptaciones culturales o técnicas son necesarias (Chandler 2004). La definición más aceptada de internacionalización proviene de LISA que la define como: «The process of enabling a product at a technical level for localization» (Lommel 2007: 17).

En la publicación anterior, LISA (2004) incluyó en su definición la abstracción de la funcionalidad como aspecto fundamental del proceso de Internacionalización. Este aspecto reviste una importancia en los estudios de localización en Traductología:

---

<sup>41</sup> El sector ha acuñado la abreviación I14n para denominar a la internacionalización. Se debe a que existen 14 letras entre la «i» y la «n».

<sup>42</sup> En nuestro estudio centrado en los sitios web desarrollados en España, se observa en gran medida la dirección opuesta, del español al inglés, en un 46% de los sitios web originalmente producidos en España.

...primarily consists of abstracting the functionality of a product away from any particular language so that language support can be added back in simply, without worry that language-specific features will pose a problem when the product is localized (LISA 2004: 14).

El objetivo de la internacionalización se centraría por tanto en el desarrollo de un producto que no dependa de una lengua determinada, de ahí el intento de abandonar paulatinamente el concepto de TO y la aparición del «texto internacionalizado» como origen de las localizaciones a todos los posibles idiomas (Pym 2004a).

Este aspecto, el objetivo de abstraer la funcionalidad comunicativa de un producto software o web e independizarla de cualquier lengua, posee un cierto paralelismo con los paradigmas funcionales y pragmáticos traductológicos (Nord 1991, 1996, 1997a, 1997b; Hatim y Mason 1990, 1997; Reiss y Vermeer 1984; Holz-Mänttari 1984). En estas investigaciones se otorga gran importancia a la necesidad de mantener la función comunicativa de un TO en el TM cuando el escopo o encargo de traducción lo requiera (véase 4.2.1.4). En el caso de la localización, el concepto de «funcionalidad» equivaldría a la función comunicativa<sup>43</sup> de cada componente o segmento textual. Ahora bien, hemos de aclarar que en el contexto interactivo de Internet el concepto de «funcionalidad» se solapa con la concepción de interacción entre máquina y usuario o entre usuario y emisor del texto (véase 5.5.6), que será el sentido con el que lo usaremos en la presente investigación.

Uno de los objetivos de la internacionalización sería por tanto definir y acotar claramente la función pragmática de cada parte o segmento textual de un texto a localizar y abstraerla, con el fin de que se pueda expresar en cualquier lengua y cultura. En este sentido, esta abstracción de la funcionalidad se ha relacionado con

---

<sup>43</sup> En el contexto de la localización defendemos la concepción de «función comunicativa» se ajusta mejor que «los propósito retóricos» a los que aluden Hatim y Mason (1990, 1997), ya que la figura del emisor desaparece gradualmente en la localización (Pym 2004).

el tan criticado concepto de la equivalencia en traducción, que en el caso de la internacionalización se puede tratar en palabras de Pym (2004) de una «internationalization-based equivalence», diferente del tan criticado y debatido concepto traductológico de equivalencia entre dos lenguas (Nord 1997a). En este caso en particular, nos parece productivo aducir que los elementos de la situación comunicativa son mucho más limitados al tratarse de un entorno digital, con una interacción claramente definida por los desarrolladores y limitada por las capacidades del medio, Internet.

Finalmente, en el caso específico de la localización web, un entorno más flexible que la localización de software, la «internacionalización web» se ha identificado además con cambios en la parte técnica del sitio web o el *back-end*, mientras la localización se asocia a la estructura de presentación o el *front-end*, lo que el usuario ve (Singh y Pereira 2005: 7). Debido a la posible confusión con el concepto de estructura superficial de la gramática generativista de Chomsky, en la presente tesis emplearemos el término «estructura de presentación» en lugar de «estructura de superficie»<sup>44</sup>. Estos dos conceptos revierten una importancia considerable en localización puesto que indican la distinción entre el texto sobre el que trabajan los desarrolladores y traductores, y el texto con el que interactúa el usuario final de un sitio web o un programa de software.

### 2.3. Definición de localización y *locale*

La localización se encuentra inmersa en un proceso global que incorpora la globalización e internacionalización. Como hemos visto, sus objetivos respectivos

---

<sup>44</sup> En ciertos estudios se denomina «estructura visual o virtual» (Mata 2005), *surface structure-deep structure* (Kersten *et al.* 2001) o *front-end – back-end* (Cronin 2003), conceptos estos más usados. Puesto que el lenguaje HTML es un lenguaje de presentación y organización, emplearemos los términos estructura profunda y estructura de presentación. Ciertos autores como Mata (2005) contraponen la estructura de presentación o «visual» con la «estructura subyacente», que en un sentido distinto al nuestro se describe como «un conjunto organizado de carpetas, subcarpetas y archivos almacenados en el servidor en el que está hospedada la Web» (Mata 2005: 200).

son preparar la organización empresarial o la vertiente técnica del producto para que sea posible su localización hacia cualquier región o *locale*<sup>45</sup>. Este concepto, del que surgió el término localización, se define por un conjunto de información específica que combina la lengua, región y codificación informática (Esselink 2000), y desempeña un papel fundamental en la recopilación de nuestro corpus (véase 7.2.1). De entre las definiciones de *locale* encontradas, a nuestro parecer la más clarificadora sería la presentada por Microsoft (Dr. International 2003), según la cual constituye:

The collection of features of the user's environment that is dependent on language, country/region, and cultural conventions. The locale determines conventions such as sort order; keyboard layout; and date, time, number, and currency formats. In Windows, locales usually provide more information about cultural conventions than about languages (Dr. International 2003: 7).

A pesar de que esta definición parte de la localización de software, por su concreción nos parece suficientemente clarificadora. Para los objetivos de este trabajo, coincidimos con esta definición en la importancia de las convenciones culturales como la base fundamental de cada *locale*. No obstante, las convenciones no se limitan a las culturales tal como recoge esta cita, sino que se extienden a distintos niveles, incluyendo las convenciones lingüísticas, tipológicas y de género.

En lo que respecta a la definición de localización, la más extendida en investigaciones en este campo es la propuesta por la Localization Industry Standards Association (LISA) en el 2003:

Localization involves taking a product and making it linguistically and culturally appropriate to the target locale (country/region and language) where it will be used and sold. [La localización supone tomar un producto y adecuarlo culturalmente y lingüísticamente a la población de destino (país/región e idioma) donde se usará y se venderá<sup>46</sup>] (LISA 2003: 13).

---

<sup>45</sup> En el presente trabajo se emplea el término *locale*, puesto que no se ha acuñado un término específico o aceptado en español este término. En Herman y Sachse (2005) se propone para nuestra lengua «parámetros específicos de una región».

<sup>46</sup> Traducción propia.

Como mencionamos anteriormente, en esta definición de localización se encuentran los aspectos básicos comunes al resto de definiciones: en primer lugar, el objeto de la transformación son productos<sup>47</sup> y no textos, postura esta que defendemos en este trabajo de corte traductológico. En segundo lugar, el proceso incluye una vertiente lingüística y otra cultural, aun cuando las lenguas se encuentran evidentemente inmersas y condicionadas por la cultura en las que se encuentran (Malinowski 1923) y son inseparables (Bassnet y Lefevere 1991)<sup>48</sup>. Y en tercer lugar, nos encontramos con el concepto de *locale*, un término que emana de la parte técnica del proceso y que, como expusimos anteriormente, se identifica con un conjunto determinado de información asociado a un lugar y-o cultura (Esselink 2000).

Consciente de la insuficiencia de la definición, la propia LISA propuso en el *Localization Primer* del 2004 una nueva definición ampliada, cuyo objetivo se dirigió a ampliar el campo de aplicación de este concepto más allá de programas de software o sitios web:

Localization is the process of modifying products or services to account for differences in distinct markets [...] (LISA 2007: 11).

La organización LISA completa esta definición con la disgregación de cuatro componentes esenciales en la localización: (1) el lingüístico, (2) el contenido/cultural, (3) el físico<sup>49</sup> y (4) el técnico. Esta nueva propuesta incide en la importancia de la globalización y la internacionalización, puesto que incluye no sólo los productos, sino también los servicios, e ignora el concepto de *locale*,

---

<sup>47</sup>«Productos» en este caso se refiere al concepto de bienes desde el campo de la economía. La Traductología se centra en el estudio del producto del proceso de la traducción, un concepto diferente.

<sup>48</sup> Esta es una concepción errónea que se repite en las publicaciones desde el sector de la localización. Según Microsoft: «when translating a book or a movie, there is no need to account for cultural variance» (Dr. International 2003).

<sup>49</sup> Este aspecto no es relevante en la presente investigación puesto que se refiere a cambios físicos como el cambio del volante a la derecha en un coche que ha de ser vendido en el Reino Unido (Lommel 2007: 12).

pasando a denominarlos «mercados». Al fin y al cabo, el concepto de *locale*, que irregularmente corresponde con el concepto de cultura meta, surge del sector y está claramente influenciado por el tamaño e importancia comercial del mercado en cuestión (Pym 2004a: 40).

Aunque la definición presentada anteriormente es la más aceptada, en nuestra investigación bibliográfica hemos identificado veinticuatro definiciones de localización. De ellas, diez surgen de la Traductología (Dunne 2006; Quah 2005; Sandrini 2005; Schäler 2004; Pym 2004a; O'Hagan y Ashworth 2003; Corte 2002; Quirion 2003; Schäler 2001: 22; Comité sectoriel de la Traduction 1999), y quince desde el sector de la localización o sectores afines (Lommel 2007, 2003; Müller 2005: 16; Chandler 2004; *Multilingual Computing and Technology* 2004; Lingo Systems 2004; Dr. International 2003; Yunker 2003; Symmonds 2003; EEEL Online 2002, cf. Foralón 2006; Esselink 2000; Uren *et al.* 1992; Microsoft 1994; ÖNORM D 1200<sup>50</sup>). Tras un análisis detallado de las diferentes propuestas, consideramos que la definición de mayor importancia por ser la más completa es la presentada por Keiran Dunne (2006), en parte porque a su vez aúna ambas perspectivas:

The processes by which digital content and products developed in one locale (defined in terms of geographical area, language and culture) are adapted for sale and use in another locale. Localization involves: (a) translation of textual content into the language and textual conventions of the target language, (b), adaptation of non-textual content (from colors, icons and bitmaps, to packaging, form factors, etc.) as well as input, output and delivery mechanisms to take into account the cultural, technical and regulatory requirements of that locale. In sum, localization is not so much about specific tasks as much as it is about the processes by which products are adapted.

Moreover, localization is but one of a number of interdependent processes and cannot be fully (or correctly understood without being contextualized in reference to them. These processes are referred to collectively by the acronym GILT (Dunne 2006: 4).

---

<sup>50</sup> ÖNORM D 1200 Austrian standard: «Localization: Linguistic and cultural adaptation and translation of software applications, technical descriptions, user manuals, etc., for local markets» (Cf. Pym *et al.* 2006: 61).

Esta definición está en consonancia con la que propondremos, ya que separa la traducción de la adaptación técnica necesaria en este proceso. Ahora bien, en la misma se encuentran deficiencias ya que no se adopta una visión precisa y englobadora sobre qué constituye un «texto»<sup>51</sup>. El investigador separa el aspecto gráfico y visual del texto en sí, que denomina aspectos no textuales, si bien en nuestra opinión, y en consonancia con las investigaciones en traducción técnica, el componente gráfico es parte integrante de todo texto y necesario en su interpretación (Göpferich 1995; Gamero 1998, 2001). Hemos de abogar por la adopción de una concepción amplia del texto en este ámbito, ya que consideramos que esta separación recurrente entre aspectos textuales y visuales en las publicaciones del sector es contrario a principios básicos en la Traductología contemporánea, tales como la concepción de «texto» en esta disciplina (véase capítulo 3). De hecho, es común relegar la traducción a un simple componente lingüístico y textual supuestamente de menor complejidad, mientras que se reserva un componente cultural y técnico para los expertos en localización cuyo perfil se supone más especializado.

Los Estudios de Traducción están gradualmente reclamando la importancia del componente traductológico del proceso (Quirion 2003), ya que se le suele restar importancia aun cuando la traducción de texto constituye la mayor parte del trabajo de localización (Lommel 2007), alrededor del 50% de inversión en este campo (LISA 2004). Sólo en tres de las definiciones que parten desde el sector encontramos una mención expresa a la traducción, en Esselink (2001: 2), Microsoft (Dr. International 2003)<sup>52</sup> y Chandler (2004)<sup>53</sup> :

---

<sup>51</sup> El investigador reduce el texto exclusivamente a las palabras.

<sup>52</sup> «The process of adapting a program for a local market, which includes **translating** the user interface, resizing dialog boxes, customizing features (if necessary), and testing results to ensure that the program still works» (Dr. International 2003).

<sup>53</sup> «Localization is the process of translating a game into different languages» (Chandler 2004: 12) [Énfasis propio].



Localization is the translation and adaptation of a software or web product, which includes the software application itself and all related product documentation<sup>54</sup> (Esselink 2001).

La definición de Esselink incluye expresamente los productos web, el objetivo de nuestro estudio, de características ligeramente diferentes a los productos de software. No obstante, se sigue separando la traducción de un proceso de «adaptación», que en este caso circunscribimos a la adaptación técnica, ya que la adaptación cultural es parte integrante del proceso de traducción global (Vinay y Dalbarnet 1958; Bastin 1996). Este aspecto surge en parte por el desconocimiento del sector de la localización de ciertos fundamentos bien establecidos en la Traductología contemporánea (Pym 2003, 2005). Según las diferentes teorías y paradigmas de la traducción, la adaptación textual, cultural, o tipológica es parte integrante del un proceso lingüístico, comunicativo y cognitivo global que enmarca todo proceso de traducción (Gouadec 2003; Baker y Malkmjaer 2001). En este caso, esta labor simplemente se refuerza de forma especial en el proceso de localización:

...on ne saurait nier qu'elles [adaptación cultural] aient toujours appartenu au travail du traducteur, du terminologue et du rédacteur; elles trouvent leur prolongement dans celui du localisateur (Quirion 2003: 547).

Como ilustración de esta integración en la labor traductora de la adaptación, en la traducción técnica no se alude al concepto de «localización» al incluir un número de teléfono diferente o una sección específica por las peculiaridades de la red eléctrica en el Reino Unido al traducir un manual de instrucciones español (Gamero 2001). De igual manera, no se considera localización la substitución de la palabra *biscuit* por la de *cookie* en la edición estadounidense de Harry Potter, o el título *Harry Potter and the Phylosopher's Stone* por *Harry Potter and the Sorcerer's Stone* (Yunker 2003: 432). Todos estos son casos de adaptación cultural inter o intralingüística, parte integrante del proceso de traducción en sí<sup>55</sup>. A

---

<sup>54</sup> Énfasis propio.

<sup>55</sup> Para un tratamiento más detallado de la adaptación en Traductología consúltese Bastin (1998).

nuestro parecer, la localización no supone más que una proyección hacia el futuro de la traducción (Quirion 2003), siendo el componente técnico o adaptación técnica necesaria durante el proceso de localización la única gran diferencia que puede separar una «traducción tradicional» (Gouadec 2003), opuesta a la «localización que es más sofisticada que la traducción» (Pym 2004a: 25). Esta reciente nueva modalidad sólo requiere una consiguiente actualización en las competencias técnicas del traductor (Quirion 2003), así como el conocimiento de las convenciones textuales propias de estos nuevos textos (véase 8.1.4.2.1).

Como pone de manifiesto el especialista en localización Reineke (2005), la controversia entre qué constituye localización o si ésta no constituye más que una traducción especializada se seguirá produciendo hasta que no se realicen mayores esfuerzos por un estudio conjunto en esta disciplina. En la presente tesis nos adherimos a esta postura, puesto que pensamos que la compartimentalización de las diferentes perspectivas dificulta el consenso sobre varios puntos básicos, como por ejemplo cuál es el objeto de estudio específico de la localización. La posición que adoptamos en la presente tesis, dentro de la propuesta de Hurtado Albir (2001: 58), enmarca la localización como una modalidad de la traducción, que es a su vez multimodal por la posible inclusión de traducciones audiovisuales, y que conjuga varios tipos de traducción como la técnica, la publicitaria o la jurídica. Dependiendo de la perspectiva, la localización web también puede caracterizarse como un tipo de traducción multimedia (Gambier y Gottlieb 2001). La innegable existencia de la localización como proceso diferenciado ha supuesto, a pesar de la diversidad de perspectivas sobre la localización y su posible dependencia o no de la Traductología, que la profesión del localizador haya surgido durante los últimos años y que sea una realidad en el mundo laboral (Dunne 2006: 1).

Si nos centramos en otras propuestas de definición en diferentes publicaciones, la localización se suele disgregar entre una adaptación cultural y lingüística del

producto. Por ejemplo, la definición de localización del Comité Sectorial de la Traducción del Canadá (1999, cf. Quirion 2003: 547) sería «l'adaptation à une langue et une culture étrangères de logiciels et de documents techniques accompagnateurs».

En la definición de O'Hagan y Ashworth (2003), desde la perspectiva de la Comunicación Mediada por Ordenador (*Computer Mediated Communication*), también se separa el contenido lingüístico del cultural:

...a process to facilitate globalization by addressing linguistic and cultural barriers to the Receiver who does not share the same linguistic and cultural background as the Sender (O'Hagan y Ashworth 2003: 66).

En este estudio se separa además la traducción dentro de la localización mediante la división entre el contenido y la presentación, (O'Hagan y Ashworth 2003), indicando que la traducción sólo se centra en el *content* y la localización trata de la manipulación tanto del contenido como de la presentación<sup>56</sup>. Esta separación se asemeja a la separación anterior entre *front end* y *back end* en un producto de software.

En general, en la mayoría de definiciones hemos constatado un intento del sector por reducir la traducción a un mero intercambio lingüístico en el proceso global, como si no fuera más que un proceso mecánico, obviando la vertiente comunicativa, cognitiva y textual del proceso traductor y separándola del componente cultural y técnico (Pym 2004a). El sector tardó en reconocer la importancia y el esfuerzo necesario en este proceso para una localización de calidad (Dunne 2006; Brooks 2000), y aún se puede observar el anhelo de automatizar el proceso, de manera que la internacionalización facilite la

---

<sup>56</sup> Los autores se basan en la separación que realiza Gile (1995) en el contexto de la interpretación.

localización hasta el punto de que sea mecánica e instantánea por medio de la traducción automática (Pym 2004a)<sup>57</sup>.

Esta separación queda patente además en las prácticas empresariales, puesto que según los datos de LISA (2004), el 80% de las labores de traducción se realizan por empresas subcontratadas y no directamente por las empresas de localización o *Localization Providers* (ibid).

No podemos concluir este epígrafe sin una mención expresa a aquellas definiciones propuestas exclusivamente para la localización web. De forma concisa y desde el sector Yunker la define como «a process of modifying a website for a specific locale» (Yunker 2003: 17). El investigador Peter Sandrini, con una perspectiva funcionalista sobre el proceso de localización web, entiende este fenómeno como «a process of modifying a website for a specific locale according to the goals outlined by the client» (Sandrini 2005: 3). El investigador sitúa en primer lugar al iniciador de la traducción, aunque en nuestra opinión las expectativas del receptor se han de situar en un mismo plano. Tras esta revisión de las distintas definiciones de localización, en un intento de observar este fenómeno desde su multiplicidad de perspectivas, creemos pertinente presentar nuestra definición desde una perspectiva pragmático-cognitiva y traductológica:

*La localización constituye una modalidad de traducción que comprende un complejo proceso textual, comunicativo, cognitivo y tecnológico por el que un texto en formato digital y en un entorno interactivo se modifica para su uso en una lengua y contexto sociocultural de recepción distintos a los originales, siempre según las expectativas de la audiencia a la que se dirige y las especificaciones o grado de localización que encargue el iniciador.*

---

<sup>57</sup> Según Schäler (2003) director del Centro de Localización de Limerick, el uso a gran escala de las memorias de traducción es la «golden rule» de la industria de localización.

## 2.4. Agentes en el proceso de localización

Desde la óptica del sector de la localización, el objetivo del proceso es que el producto o texto sea recibido como si se hubiera desarrollado en la región de destino (LISA 2004: 11). Este hecho se canaliza hacia una de las características principales del texto localizado: éste no necesita representar al texto origen, sino que necesita representar la función textual del mismo en otra cultura. Según Pym<sup>58</sup>:

The localized text is not called on to represent any previous texts, it is instead part of one and the same process of constant material distribution, which starts in one culture and may continue in many others (Pym 2004a: 6).

Desde un punto de vista traductológico, es de interés que en las definiciones mencionadas no se menciona a los agentes del proceso, ni al autor del TO, ni al traductor, ni al receptor del TM. Sólo en la definición de O'Hagan y Ashworth (2003), al encontrarse dentro de un modelo de Comunicación Mediada por Ordenador<sup>59</sup> (*Translation Mediated Communication*), se menciona al emisor y al receptor. En general, esto se debe a que el software necesita ser anónimo, «software is anonymous technology speaking to anonymous markets» (Pym 2004a: 6). En el caso de los sitios web corporativos, el proceso comunicativo que se entabla es doble: entre el sitio web como ente autónomo y el emisor a través de la interactividad que proporciona el medio, y por otro lado entre la empresa como entidad y el receptor en un nuevo modelo de comunicación (Janoschka 2003: 98).

Si comparamos el proceso de localización con otros tipos o modos de traducción, el primero requiere una mayor participación de agentes para completar una localización de calidad, de ahí que el trabajo conjunto de desarrolladores,

---

<sup>58</sup> El concepto de distribución textual es esencial en el trabajo de Pym sobre la localización (2004: 4), en sus palabras: «The principle of dynamic distribution [of text] is essential to our view of localization and translation».

<sup>59</sup> Se considera como un caso de «Comunicación Mediada por Ordenador» que incluye la interpretación de chats, traducción de correos, etc.

traductores - localizadores, y *QA testers* lingüísticos y técnicos sea, en nuestra opinión, el germen de la separación entre la traducción y la localización. Varios investigadores hacen hincapié en que esta necesaria cooperación entre los diversos agentes fue la razón por la que se comenzó a separar la traducción de la localización (Quah 2005: 11). Al hilo de esta afirmación, nuestra visión propia de la localización se articulará en torno a este necesario entorno cooperativo en el que el traductor no es totalmente responsable del producto final. Los agentes se corresponden con las distintas fases del proceso, el análisis de mercados, la preparación del proyecto o internacionalización, la traducción y el control de calidad o QA (ibid). Como ejemplo, Chandler (2004: 61) señala que el proceso de localización de videojuegos precisa la intervención de veinte agentes diferentes, con dos únicas menciones a traductores; en el proceso de traducción y en proceso de control de calidad.

En los estudios traductológicos se suele mencionar como agentes básicos al iniciador, emisor del TO, traductor o intérprete, y el receptor del TM. A nuestro parecer, el distanciamiento entre el sector y la Traductología se ha producido en parte por la dificultad para imbricar en este proceso al equipo multidisciplinar de localización en el seno de los diferentes paradigmas y teorías en Traductología. Como mencionamos previamente, las diferencias entre ambas perspectivas han dificultado el intercambio de conocimientos entre ambas (Pym 2003, 2005), y en cierta manera, han llevado a que la localización se haya establecido y se practique sin una aplicación plena de los conocimientos de la Traductología (Pym 2003, 2005; O'Hagan y Ashworth 2003). En este sentido, y a pesar de que el trabajo colaborativo es inherente a la labor traductora, abogamos por un modelo de proceso traductor-localizador específico en el que se incorpore la cooperación los desarrolladores responsables de la vertiente técnica o tecnológica. Esta interacción localizador-desarrollador ha de producirse necesariamente en ambas direcciones, mediante la identificación de problemas de traducción-localización previos, fase

esta correspondiente a la internacionalización, o la identificación de problemas de traducción durante el proceso que requieran de la intervención o asistencia del desarrollador.

De entre los agentes en el proceso, el receptor del TM es de especial interés en la localización web. Éste es parte fundamental en un enfoque textual puesto que su conocimiento global (Beaudegrande y Dressler 1981), sus expectativas y su conocimiento de las convenciones tipológicas afectan al proceso de producción y traducción (Nord 1997a). En localización se suele denominar «usuario» (Lommel 2007). Este matiz es esencial puesto que señala que los textos digitales no necesitan ser leídos, sino «usados», en un contexto de interactividad que marca claramente la tipología de los mismos. Sin embargo, el «usuario» en general no necesita leer la totalidad del texto, sino que escanea la información y lee o usa sólo la que necesita (Nielsen 2000). En la publicación de Pym (2004: 34) se critica la premisa de que todos los receptores van a «usar» el texto, y de hecho, no todos los receptores van a usarlo, como en la lectura de una ventana emergente de publicidad web que en general se cierra sin más<sup>60</sup>. Debido a esta peculiaridad, en la presente tesis se empleará el concepto de «usuario» y no «lector», «destinatario» o «receptor»<sup>61</sup>, el concepto más usado en Traductología, con el fin de ajustarlo a las peculiaridades de la situación comunicativa que contextualiza este tipo de textos.

## **2.5. La adaptación técnica**

Como avanzamos en el epígrafe anterior, la adaptación técnica durante la fase de internacionalización constituye uno de los aspectos que diferencia la localización de otras modalidades de traducción. No obstante, la labor del traductor primordial

---

<sup>60</sup> El usuario no usaría el mismo pero sí que «interactúa» al cerrar esta ventana emergente.

<sup>61</sup> En la Traductología contemporánea un número creciente de investigadores emplean el término «audiencia» (Muñoz Martín 1995; Mason 2000; Nobs 2005; e. o.). Dado que nuestra tesis se fundamenta en el intercambio comunicativo en un medio específico, Internet, con una interactividad determinada, emplearemos el término «usuarios».

en esta fase, dado que sus conocimientos sobre la lengua y cultura meta son esenciales a la hora de establecer qué adaptaciones son necesarias (Chandler 2004). Asimismo, se requieren sus competencias lingüísticas y culturales durante la fase de Control de Calidad lingüística (*linguistic QA*). No obstante, gran parte de las adaptaciones a llevar a cabo, tales como los problemas de codificación, son competencia exclusiva de los desarrolladores. A continuación presentamos una relación de las diferentes adaptaciones técnicas necesarias que resumimos de las publicaciones de mayor prestigio sobre la internacionalización (Herrmann y Sashse 2005; Dr. International 2003; Symmons 2003; Esselink 2000; Uren *et al.* 1992):

- Separación del texto a traducir del código fuente
- Formatos de archivos propietarios
- Codificación: conjuntos de caracteres y fuentes
- Dirección y ordenación de los textos
- Moneda y formato de los números
- Adaptación de sistemas de medida y posiciones geográficas
- Calendarios y zonas horarias
- Direcciones y códigos postales
- Representación de números de teléfono
- Organización de listas alfabetizadas
- Formato mayúsculas y minúsculas
- Límite de palabras
- Límite de los caracteres: de importancia para idiomas como árabe, farsi o urdu.
- Silabificación y palabras unidas por guiones
- Expansión del texto. En general, 25% más de espacio.
- Teclas rápidas
- Operaciones aritméticas
- Tamaño estándar de papel



- Colores
- Diseño de iconos
- Disposiciones legales

Durante las dos últimas décadas el sector de la localización ha consensuado las adaptaciones técnicas más comunes, sobre todo en productos de software-hardware. La asignación de ciertas áreas a una fase determinada puede resultar complicada, como sería el caso de la adaptación al marco legal en la localización hacia otro país. A pesar de que las adaptaciones técnicas se desarrollan durante la fase de internacionalización, un cierto número ha de investigarse y establecerse durante la fase de globalización, que Microsoft denomina «internacionalización» (Dr. International 2003). Para ilustrar este aspecto, podemos recurrir al caso de la localización de videojuegos para Alemania, cuya legislación vigente no permite el uso excesivo de sangre, ya que este hecho podría suponer la prohibición de su venta. Consecuentemente, durante la fase de internacionalización se suele incorporar la posibilidad de reducir o cambiar a verde el color de la sangre en las versiones localizadas hacia este país (Chandler 2004). Como se verá en los resultados obtenidos en la presente tesis, parte de los errores identificados son atribuibles a deficiencias en la fase de internacionalización (véase capítulo 11).

## 2.6. Componente cultural

La adaptación cultural de un producto de software o texto web supone uno de los pilares de la localización (Lommel 2007: 14). Esta adaptación cultural se asemeja a la necesaria en otros tipos de traducción<sup>62</sup>, combinada con otros aspectos específicos determinados por el medio y el nuevo modelo de comunicación. Dada

---

<sup>62</sup> La adaptación del espacio disponible se lleva realizando desde hace muchos años en subtitulación o en la traducción de impresos en diferentes campos como el judicial, el publicitario etc. La traducción de iconos puede ser comparable a los incluidos en la traducción de manuales de instrucciones o en publicidad, y las adaptaciones a la forma cultural en las tiendas web virtuales son comparables a las que se realizan en la traducción de catálogos de compra por correo.

la recomendación de que los textos localizados se reciban como producciones originales de la región de destino, hemos de argumentar que la consecución de este objetivo pasa por incluir aquellos parámetros que sean determinados por la cultura, como puede ser la estructura textual (Neubert y Shreve 1992) o las convenciones de cada género o tipo textual (Reiss y Vermeer 1984; Hurtado Albir 2001). En este sentido, ciertos investigadores en localización web apuntan que este objetivo no se corresponde con «engañar» al receptor para que éste crea que la empresa es local, sino que perciba que la empresa respeta y aprecia la cultura hacia la que se dirige (MacDonough 2006; Yunker 2003: 18). No obstante, no podemos obviar el hecho de que el receptor no identifica necesariamente a una empresa determinada con una lengua en particular o una cultura de origen determinada.

Desde los comienzos de la localización, el sector ha sido consciente de la primordial importancia de la adaptación cultural (Esselink 2001). Su línea de actuación se ha encaminado a desarrollar productos con el mayor nivel de neutralidad cultural posible (Cronin 2003: 18), un discutible objetivo que se observa desde los años 80. Este componente cultural ha atraído la atención de un número creciente de investigadores durante los últimos años (MacDonough 2006; Tercedor 2005; Singh y Pereira 2005; Singh *et al.* 2004; Schäler 2002), y ha sido el objeto de varios estudios empíricos que lo relacionan con una mejora en la navegación, interacción y valoración de los sitios web, puesto que esta adaptación conlleva un menor esfuerzo cognitivo y el entorno de interacción es más claro y eficiente (Singh *et al.* 2004; Luna *et al.* 2002).

Si nos centramos en qué elementos forman parte de esta adaptación cultural, desde un enfoque traductológico el estudio de Tercedor (2005: 153) señala cinco elementos clave desde una doble perspectiva pragmático-textual y cognitiva:

- Los iconos

- Los aspectos lingüístico-textuales como la intertextualidad, el registro o el macrotexto
- Los aspectos gráficos
- Los aspectos técnicos
- Los aspectos cognitivos, tales como la navegación, las metáforas, los modelos mentales y la interacción

Desde una disciplina diferente, el marketing, y más centrado en la localización web, los investigadores Singh y Pereira (2005) han realizado un estudio pormenorizado de las adaptaciones culturales en sitios web. La percepción, el simbolismo y el comportamiento se sitúan en su visión como elementos clave que definen la cultura, que a su vez establecen los valores compartidos y los patrones estructurados de pensamiento. La contribución esencial de los autores consiste en elaborar un marco de estudio de valores culturales en el que basan su investigación sobre las adaptaciones culturales en la localización web. Los valores identificados se centran en el elemento del comportamiento y se fundamentan en las investigaciones de Hofstede (1991)<sup>63</sup>, de gran impacto en el diseño de software y sitios web:

- Individualismo frente a colectivismo (*Individualism-Colectivism*)
- Distancia al poder (*Power distance*)
- Tolerancia hacia la incertidumbre (*Uncertainty Avoidance*)
- Masculinidad-Feminidad (*Masculinity-Femininity*)
- Contexto bajo o alto <sup>64</sup>(*Low-High Context*)

Los investigadores identifican ciertos elementos en los sitios web como variables cuantificables que permiten la comparación entre culturas diferentes. Las

---

<sup>63</sup> Para mayor información sobre estas dimensiones se puede visitar el sitio web del autor. <http://www.geert-hofstede.com/>.

<sup>64</sup> Este valor se basa en los estudios de Hall (1976).

implicaciones de este estudio en nuestra tesis se relacionan con las diferencias significativas entre España y los países anglosajones en todos los valores o variables, un hecho que apoyaría la pertinencia de un estudio como el nuestro. Estas diferencias implican posibles divergencias tanto en la tipología textual como en la estructura de los géneros. La importancia del individualismo en la cultura anglosajona, a diferencia de la española, puede conllevar la convencionalización de la política de privacidad en la totalidad de los sitios web, mientras que en España, una cultura más colectivista, no se ha convencionalizado este aspecto en el género (véase 9.1.2.6). Por otro lado, España presenta un alto índice en el valor «tolerancia hacia la incertidumbre», y esto se cristaliza en el uso de una terminología más local o la alta recurrencia de los datos de registro mercantil y código de identificación fiscal (CIF) de la empresa, una garantía de seguridad para el usuario (véase 10.3.5.3).

The image shows a screenshot of the IKEA website in Hebrew. The layout is adapted for right-to-left reading. On the left side, there are three green arrows pointing to the left, indicating navigation options. On the right side, there are three green arrows pointing to the right, indicating navigation options. The main content area is centered. At the bottom, there is a table with contact information for different regions. The footer includes the text 'Inter IKEA Systems B.V. 2003 - 2005' and a 'go to IKEA global' button.

מוקד שרות טלפוני:	פקס:	כתובת למשלוח דואר:
09-8928888	09-8928880	ת.ד. 8607
שלוחה מס' 3 - בירור מלאי		רחוב גיבורי ישראל 1
שלוחה מס' 4 - קבלת שירות לאחר רכישה		אזור תעשייה פולג
שלוחה מס' 5 - קבלת נציג שירות בנושאים אחרים		נתניה דרום
שעות פעילות המוקד		מקוד 42504
א- ה 8:00 - 14:00 שישי וערבי חג 9:00 - 20:00		

Figura 8. Adaptación cultural a la percepción según Singh y Pereira (2005). Las flechas apuntan a la dirección de lectura sea a la derecha en hebreo o la izquierda en inglés. www.ikea.com.

En resumen, la adaptación cultural supone un elemento esencial en el proceso de localización, avalada por estudios empíricos que relacionan la mejor adaptación cultural de sitios web con el aumento de su valoración en España (Singh *et al.*

2004). No obstante, en localización la posibilidad de realizar adaptaciones culturales más allá de las textuales depende del objetivo del iniciador y su encargo (Nord 1997b)<sup>65</sup>, de donde surge el concepto de «grado de localización» que trataremos a continuación.

## 2.7. Diferentes grados de localización

La localización como proceso económico entraña la introducción de diferentes grados de localización dependiendo de la importancia económica del mercado o la *locale* (Brooks 2000). Este aspecto nos es de interés al encontrarse el proceso de localización inevitablemente influenciado por el escopo o encargo (Reiss y Vermeer 1984; Nord 1997b), por lo que no podríamos delimitar establecer unas características homogéneas del proceso al que se enfrentan los traductores en su labor localizadora, al depender éstas del grado de localización. Según Microsoft, el concepto «grado de localización» se define como:

The amount of translation and customization necessary to create different language editions. The levels, which are determined by balancing risk and return, range from translating nothing to shipping a completely translated product with customized features (Dr. International 2003: 15).

Si tomamos este ejemplo de Microsoft (Brooks 2000: 49-50), un modelo repetido en el sector y de gran impacto en la localización web, observamos que se divide en tres grados:

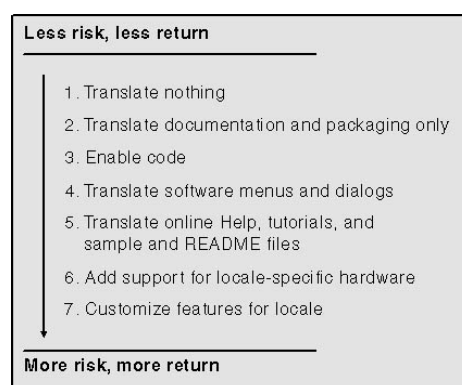
- Primer grado o «enabled». Los usuarios pueden componer textos en su propio idioma, pero el software y los archivos de ayuda están en la lengua de desarrollo, el inglés en este caso. Este grado no tendría importancia en

---

<sup>65</sup> Ciertas adaptaciones culturales, como la posibilidad de lectura de derecha a izquierda en el árabe o hebreo, requieren de una compleja etapa técnica ajena a la labor traductora en sí.

una investigación sobre la textualidad web, puesto que en este caso la adaptación sería sólo técnica.

- Segundo grado o «localizado». La interfaz de usuario y la documentación se traducen, pero las herramientas específicas de cada lengua (como los correctores ortográficos) y el contenido permanecen en inglés, o en otra lengua, como el español por defecto en el caso de las localizaciones al catalán (Pym 2004a).
- Tercer grado o «adaptado». Se revisan o se recrean las herramientas lingüísticas, el contenido y las funciones para el mercado de destino.



**Figura 9. Diferentes grados de localización en los productos de software (Dr. International 2003)**

Este modelo podría aplicarse al proceso de localización web, aunque sólo existirían el segundo y tercer grado. Por ejemplo, el sitio Web de Ikea<sup>66</sup>, señalado como el mejor sitio web localizado (Singh y Pereira 2005), presenta gran cantidad de localizaciones, como el sitio turco o el saudita, en los que el contenido, la funcionalidad y el aspecto gráfico son similares, mientras que el francés o el español presentan diferencias notables fruto de su localización a un tercer grado.

---

<sup>66</sup> [www.ikea.com](http://www.ikea.com).

Con respecto a los distintos grados, nos es preciso apuntar que los modelos de globalización web pueden ser *centralizados* o *descentralizados*, (Yunker 2003; O'Hagan y Ashworth 2003: 74), por lo que la localización web a un tercer grado se podría confundir con una producción local. En el primero caso, todos los sitios web localizados se controlarían desde el país de origen por medio de programas de *Global Content Management* (véase 2.8), como Goolge.com, y el en el segundo, los sitios locales para cada región se desarrollarían y mantendrían en cada país con contenidos y promociones específicas, pero manteniendo la imagen global corporativa del sitio.

Esta identidad gráfica y visual es de gran importancia para su identificación (Hunt 1995), y uno de los aspectos a mantener en el proceso de localización (Dr. International 2003). Un ejemplo en nuestro corpus sería [www.adeco.es](http://www.adeco.es) puesto que, aun siendo una empresa multinacional y manteniendo el mismo aspecto visual, el sitio web español responde a una producción local<sup>67</sup>.

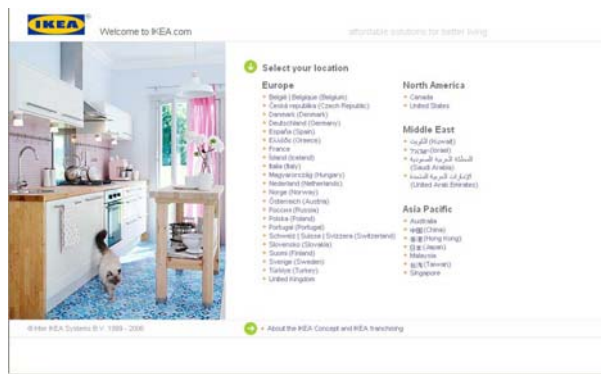
Son varios los modelos de grados de localización web que se han propuesto, y en éstos presentan ciertas diferencias dependiendo de sus objetivos (Singh y Pereira 2005; Yunker 2003: 128-130)<sup>68</sup>. En nuestra opinión, la propuesta de clasificación más concisa y práctica sería la presentada por los investigadores Singh y Pereira (2005: 10-15), que se basan en la importancia de la cultura meta en la localización:

---

<sup>67</sup> Determinamos que es una producción local puesto que los comentarios de programación y las variables en los script java se encuentran en español (véase 7.2.3).

<sup>68</sup> Yunker diferencia entre: (1) Localización incremental: localización parcial sin inversiones en el apartado técnico para observar la importancia del mercado. (2) Localización exhaustiva: la localización del sitio web a partir de un sitio original en el que se debe crear una experiencia del usuario equivalente. (3) Localización adaptada: es la recreación total del sitio web para que los usuarios sientan que el sitio web se ha creado en la región geográfica y no es un sitio globalizado que dificulta su usabilidad.

## Capítulo 2. La localización como objeto de estudio



Portal internacional de Ikea.com. La selección de sitios localizados se realiza en el idioma original, y en paréntesis aparece su traducción al inglés.  
<http://www.ikea.com/>



Página web de Ikea España. La localización ha añadido un mapa local para la selección de tiendas.  
[http://www.ikea.com/ms/es\\_ES/](http://www.ikea.com/ms/es_ES/)



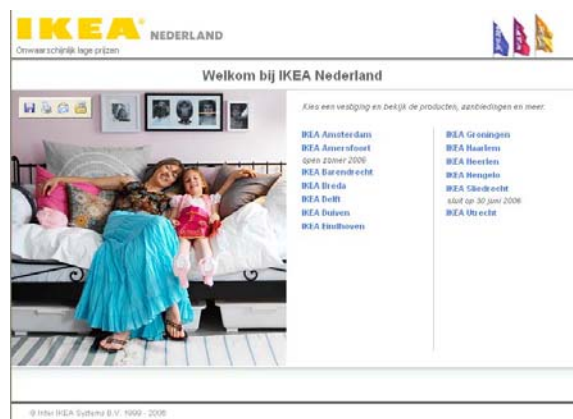
Sitio web de Turquía. El contenido y forma es similar a la web de Arabia Saudita, pero con orientación de izquierda a la derecha.  
<http://www.ikea.com.tr/content/index.asp>



Sitio web de Arabia Saudita. Se puede observar la barra de navegación a la izquierda para acomodar la escritura de la derecha a la izquierda.  
<http://www.ikea.com.sa/content/index.asp>



Sitio web de Francia. Contenido similar al de Holanda o Rep. Checa.  
[http://www.ikea.com/ms/fr\\_FR/](http://www.ikea.com/ms/fr_FR/)



Sitio web de Holanda con contenidos similares al francés.  
[http://www.ikea.com/ms/nl\\_NL/](http://www.ikea.com/ms/nl_NL/)

**Figura 10. Diferentes grados de localización en Ikea.com. El sitio de Holanda es similar al de Francia. Igualmente el sitio turco y saudita presenta contenidos similares aunque adaptados a la percepción con su lectura de izquierda a derecha o al contrario.**



- Sitios web **únicos**<sup>69</sup>: los sitios de empresas globales cuya web es similar para todos los países, en una sola lengua sin traducción alguna, por ejemplo [www.tyco.com](http://www.tyco.com).
- Sitios web **semilocalizados**: sitios web en los que se ofrece la información de contacto de la filial local, pero no se ofrece contenido específico para países diferentes. [www.gap.com](http://www.gap.com).
- Sitios web **localizados**: en los que encontramos ciertas páginas web localizadas específicas para cada país, pero no se adapta el *back-end* o estructura profunda de las páginas.
- Sitios web **localizados extensivamente**: en lo que nos encontramos una localización global y el contenido se adapta totalmente a la *locale* de destino, como puede ser [www.amazon.com](http://www.amazon.com).
- Sitios web **culturalmente adaptados**: en los que se refleja una inmersión total en la cultura de destino, objetivo de su estudio. Diferencian tres niveles globales de adaptaciones culturales: la percepción, el simbolismo y el comportamiento. A pesar de que indican que ningún sitio web llega a este nivel de localización, señalan que [www.ikea.com](http://www.ikea.com) sería el más cercano.

A pesar de que los investigadores no describen con la necesaria concreción este último grado, en la presente investigación adoptamos esta división operativa, ya que la propuesta de Microsoft se refiere expresamente a los programas de software. Durante la fase de recopilación del corpus localizado se registraron estos niveles, y se incluyeron como localizaciones los sitios hacia España en un grado *semilocalizado*, aun cuando mantuvieran una sola página en español, como <http://www.allergan.es/>.

---

<sup>69</sup> Los investigadores denominan este nivel en inglés *standardized*.

Como resumen, si analizamos los diferentes grados que el sector aplica a sus procesos de localización, llegamos a la conclusión de que este proceso y su producto, el texto localizado, no poseerá necesariamente unas características homogéneas sino que el encargo (Nord 1997b) o el grado de localización repercutirá en el texto de llegada. De hecho, éste podrá presentar una estructura profunda o *back-end*, superestructura, estructura discursiva y tipología textual con distinto grado de dependencia del TO.

## 2.8. La localización de software y la localización web

Como hemos apuntado, la localización web se desarrolló a partir de la localización de productos de software (Yunker 2003), de ahí que se considere como una de sus ramas. Desde un punto de vista traductológico, son varias las diferencias fundamentales que encontramos entre sendos procesos de traducción, si bien existen un gran número de características similares.

Desde el punto de vista **comunicativo**, Internet es un medio electrónico, global, e interactivo y cada una de estas particularidades afectan en gran medida al tipo de texto y lengua que se encuentran en la misma (Crystal 2001; Vengadasamy *et al.* 2004). En lo que respecta al proceso de traducción, la «interacción» e «interactividad» son esenciales<sup>70</sup>, y las separaremos siguiendo el trabajo de Anna Janoschka (2003: 115). La autora define la interacción como la comunicación entre el emisor del mensaje y el receptor, que se puede realizar a través de correos electrónicos, chats o formularios. La interactividad, por su parte, se refiere a la comunicación del usuario con el medio (Janoschka 2003; Jucker 2003). Ésta se fundamenta en el uso de funciones de búsqueda, la lectura selectiva de la página, la

---

<sup>70</sup> La interactividad se ha apuntado como la principal diferencia entre el texto escrito y en línea (Mill 2005: 16).

activación de hipervínculos, la activación de menús de navegación o cualquier otra forma de participación activa por parte del usuario (Janoschka 2003: 116).

En la situación comunicativa que se establece durante el uso de un programa de software, la funcionalidad de los segmentos del TO se basa en la interactividad entre el programa de software como ente autónomo y el receptor del TM. Por el contrario, en los sitios web nos encontramos tanto la interactividad entre el sitio web y el receptor (como en los menús de navegación o ventanas emergentes en *javascript*), como interacción entre el receptor del TM y el ente responsable del sitio web (en formularios, contacto por correos electrónicos, compras electrónicas etc.). Es decir, que los textos digitales distribuidos por Internet no se ciñen a facilitar la interacción humana entre máquina y usuario, sino que por un lado su función es facilitar la comunicación entre los usuarios y el texto web por medio de la interacción (con la opción de búsqueda o la posibilidad de enlazar el texto con cualquier otro hipertexto<sup>71</sup>), y por otro lado con el «emisor» del sitio, la empresa en nuestro caso.

La traducción es asimismo un proceso **cognitivo** (Bell 1991; Kiraly 1995; Gile 1995; Lörcher 1991; etc.). En este sentido, el medio digital influye en el proceso de traducción, puesto que los textos están inmersos en el código fuente de diverso tipo. En el caso del software es normal que esté separado de las cadenas textuales que aparecen en pantalla<sup>72</sup>, aunque pueden incluir variables de diverso tipo (Esselink 2001: 65; Herrmann y Sachse 2005). En el caso óptimo de localización web, suelen ser los programas de traducción asistida los que separan el texto del código informático durante el proceso de traducción<sup>73</sup>. Varios autores dividen la localización entre (1) la localización de software, con un énfasis en avanzadas competencias técnicas de los traductores, y (2) la localización web o localización

---

<sup>71</sup> El fin principal del creador de la Web, Tim Beerner-Lee (1999) «The vision that I have for the Web is about anything potentially connected with anything».

<sup>72</sup> En el caso de que se haya producido una óptima internacionalización del producto de software.

<sup>73</sup> Como puede ser el caso del programa de memoria de traducción Trados S-Tagger.

de contenidos, más simple que la anterior y con una mayor importancia de las competencias lingüísticas y culturales (Esselink 2006). De igual manera, y como apuntamos anteriormente, el grado de localización también ejerce una influencia determinante en el proceso de traducción.

De igual modo, el medio digital supone la visualización del texto en pantalla con unas restricciones de espacio (Nielsen 2000). En la localización de software, este punto suele presentar mayores problemas que se resuelven con programas de localización WYSIWYG<sup>74</sup> como *Catalyst* o *Passolo*. Sin embargo, la limitación de espacio es un problema de menor importancia en la web, puesto que el texto se puede visualizar de manera «líquida»<sup>75</sup>, es decir, adaptándolo al navegador, la configuración de pantalla y tamaño de letra que haya elegido el receptor (Nielsen 2000: 318). De hecho, suelen ser las preferencias del receptor las que establecen cómo se visualiza la página, por lo que varias publicaciones mencionan que el concepto de WYSIWYG posee una menor importancia en localización web (ibid).



**Figura 11. Disposición líquida en pantalla con posibilidad de variar el tamaño de la letra. Icono de [www.ono.es](http://www.ono.es).**

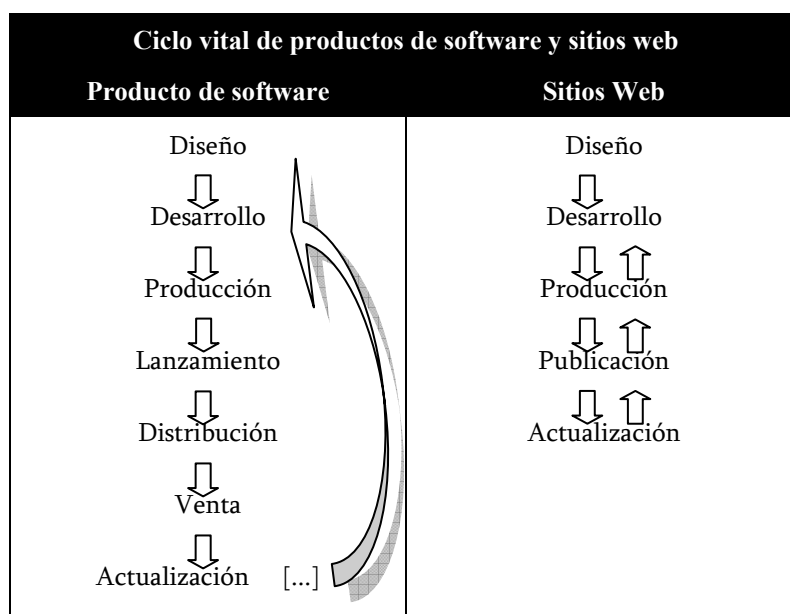
Por último, en cuanto a la traducción como operación **textual** (e. g. Neubert y Shreve 1992; Hatim y Mason 1990 1997; Baker 1992, 1998), una de las mayores diferencias entre la textualidad de ambos casos sería la estabilidad del texto y su posibilidad de actualización (Mata 2005). El texto en el seno de un programa de software es más estable, puesto que las actualizaciones y cambios se producen de

---

<sup>74</sup> WYSIWYG. - *What You See Is What You Get*. En estos programas de software el traductor trabaja sobre la misma estructura de presentación que ve el usuario final y no sobre segmentos textuales separados de su presentación.

<sup>75</sup> Esta es una recomendación básica de los manuales de usabilidad. No obstante, en el estudio de Nielsen y Tahir (2002) sobre sitios web corporativos sólo el 18% de los sitios lo incluían.

versión en versión; un periodo de tiempo que suele ser de uno o varios años. En el caso de los textos web, la actualización del contenido y su consiguiente localización puede producirse con gran rapidez, ya que es posible añadir subtextos o elementos textuales en cualquier momento. De hecho, ésta es una de las características intrínsecas de los hipertextos (véase 3.3.2). Hoy en día, la actualización de sitios web se considera primordial y se asocia a la actualidad de la información. De no ser así, ésta puede recibirse como anticuada en el dinámico mundo de Internet (McAlpine 2001: 122)<sup>76</sup>. Para ilustrar las diferencias en los ciclos vitales, hemos recurrido al siguiente esquema presentado por Mata (2005). En el mismo se recoge cómo el proceso de actualización del texto web se puede iniciar durante todo su ciclo vital, mientras que en la localización de software la actualización presenta mayores restricciones:

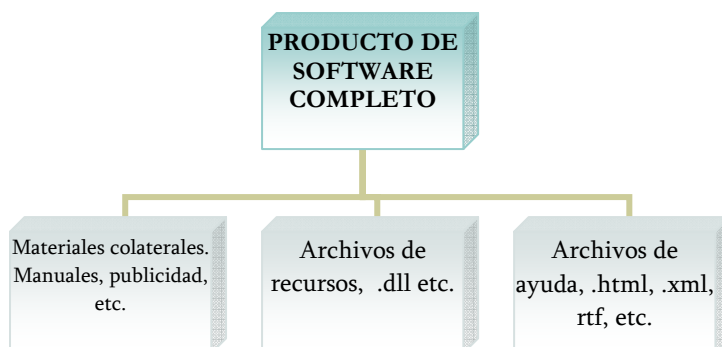


**Tabla 1. Ciclos de vida en productos de software y contenido web con mayores posibilidades de actualización en los textos web (Mata 2005: 194).**

<sup>76</sup> La autora señala que en Internet existe una cantidad ingente de sitio web antiguos no actualizados, y este aspecto se relaciona con la falta de credibilidad esencial en todo sitio web (Standford Credibility Guidelines 2001).

Otro concepto igualmente importante en los textos web es la «dinamicidad textual». Un texto web «dinámico» es aquel que no posee una versión estable y definitiva en un momento temporal determinado, por lo que dos usuarios distintos pueden recuperar textos distintos al acceder al mismo sitio. Estos textos dinámicos se producen a través de un servidor activo por medio de un *Content Management System* o programa específico en el servidor. Los textos dinámicos se contraponen a los textos web «estáticos», aquellos textos web estables cuya presentación es similar entre usuarios o en el tiempo.

Por último, la localización web surgió de la localización de software, y por tanto, también se encuentran subtextos con similares procesos de traducción, tales como la localización de presentaciones en formato Macromedia Flash en sitios web (Macromedia 2004), cuya localización es similar a la de los archivos de recursos en software. Otro caso en el que encontramos semejanzas son los archivos de ayuda en la localización de software<sup>77</sup>. Estos textos suelen encontrarse en formato html o Xml y están interconectados por hipervínculos, por lo que su traducción es similar a la de contenido web. Los archivos de ayuda serían uno de los tres grandes bloques en los que se separa tradicionalmente la localización de software, junto con los archivos de recursos y los materiales colaterales (Esselink 2001):



**Figura 12. Producto de software completo. Elaboración propia.**

<sup>77</sup> De hecho, la mayoría de los archivos de ayuda en la actualidad se producen en html o xml.

## 2.9. Revisión de estudios sobre localización

En este último apartado resumimos los resultados de nuestra búsqueda bibliográfica sobre los procesos de localización. En general, la mayoría de los estudios monográficos sobre la localización parten del sector y aún son escasos los estudios de impacto que surgen desde la perspectiva de los Estudios de Traducción (Quirion 2003). De hecho, y como manifestamos anteriormente, podemos constatar que el sector de la localización se ha desarrollado instaurando sus propios modelos, obviando hasta cierto punto el cuerpo de conocimiento de la Traductología (Pym 2003; O'Hagan y Ashworth 2003: 130). Este hecho se repite en el desarrollo de otros tipos de traducción como la traducción automática (Quah 2005: 176), y refleja una división entre el mundo de la traducción como actividad económica y la Teoría de la Traducción (Pym 2004b).

La mayoría de las publicaciones giran en torno a los aspectos técnicos del proceso de localización (Quirion 2003), ya que la implementación de la fase técnica revierte una gran dificultad. No podemos obviar el hecho de que en estas publicaciones se intenta separar el aspecto textual o parte lingüística como algo independiente y de menor complejidad (Pym 2004a: 25). Como mencionamos anteriormente, la traducción se considera simplemente un problema lingüístico, «just a language problem» (Brooks 2000)<sup>78</sup>, o «translation of words» (Esselink 2006: 28). A pesar de los recientes esfuerzos desde la Traductología, la traducción se sigue percibiendo como un simple componente relacionado con el significado de palabras aisladas, o «focused on communicating the meanings and messages of words.» (Lommel 2007: 7). Esta desconexión ha supuesto que durante los últimos años se hayan alzado voces desde la Traductología a favor de la investigación especializada en este campo. Como comenta Quirion:

---

<sup>78</sup> Este autor concibe a la traducción como «Translation is the core skill of converting text from one language to another, whether on hard copy or electronically» (Sprung 2000: 10).

L'établissement de la discipline sera appuyé par des recherches sur les aspects linguistiques de la localisation; les publications scientifiques sur la localisation bâtiront, pierre par pierre, la reconnaissance de ce nouveau champ de recherche (Quirion 2003: 557).

Coincidimos plenamente con el investigador, puesto que gradualmente el cuerpo de conocimiento de esta nueva disciplina se irá formando a partir de estudios empíricos desde una óptica traductológica, que el autor denomina aspectos «lingüísticos» al dejarse llevar por las definiciones de la localización desde el sector o una perspectiva lingüística en Traductología<sup>79</sup>. En este sentido, esperamos que nuestra tesis suponga una contribución de importancia. El investigador Anthony Pym ha publicado un gran número de artículos que recogen esta división entre el sector de la localización y la Traductología en un esfuerzo claro por tender puentes entre ambos (Pym 2003, 2004b, 2005). En nuestra opinión, su mayor aportación se centra en fomentar la relación bidireccional entre el sector de la localización y la Traductología ya que, como nos comenta, esta desconexión ha llevado al sector la localización a ignorar parcialmente treinta años de intensa investigación en Traductología (Pym 2003).

El número global de publicaciones sobre localización no es amplio debido a la juventud de esta incipiente disciplina. A continuación se repasan las publicaciones de mayor impacto separándolas entre las dos perspectivas mayoritarias de estudio; (1) la Traductología y (2) el sector, ya que su adscripción a ésta modela su visión global del fenómeno (Foralón 2006).

### **2.9.1. Estudios desde la perspectiva traductológica**

Desde la perspectiva de la Traductología, *The moving text* de Anthony Pym (2004) es el primer monográfico que estudia exclusivamente la localización desde una

---

<sup>79</sup> Como quedó patente anteriormente, el sector suele separar el componente lingüístico del cultural en el proceso global. En nuestro estudio defendemos que la definición aportada por Esselink (2001) sería más correcta, ya que acertadamente separa la «traducción» de la «adaptación técnica».



perspectiva teórica. Su libro es una recreación aplicada a la localización de su libro anterior, *Translation as Text Transfer*, (Pym 1992). El autor basa su teoría de la localización en su principio de la distribución textual como elemento esencial y determinante del proceso de localización. Según su teoría, la resistencia a esta distribución es el determinante que promueve o no la localización del texto. A nuestro parecer, una de las mayores contribuciones de su estudio es la progresiva desaparición del TO en el proceso de localización, ya que se basa en una internacionalización de la funcionalidad del producto en la que a partir de un producto internacionalizado se desarrollarían sus diferentes versiones, que a su vez serían TO y no TM. El autor ha publicado además gran cantidad de artículos sobre la división entre la Traductología y la Localización, o sobre cómo la teoría de la traducción puede ayudar a esta última (Pym 2003, 2004b, 2005).

En 2006 se publicó *Perspectives on Localization*, un volumen coordinado por Keiran Dunne y patrocinado por la Asociación Americana de Traductores (ATA). En el mismo se unen investigadores desde la Traductología y desde el sector y se aboga por un estudio conjunto de esta disciplina. El editor indica que las aproximaciones a la localización han sido históricamente descriptivas, puesto que los diferentes problemas que ha supuesto la localización se fueron encontrando gradualmente, y el autor indica que la práctica de la localización requiere una aproximación prescriptiva para su práctica efectiva (Dunne 2006: 6). A través de las diferentes contribuciones se exploran en el volumen los asuntos de mayor actualidad en la disciplina: el modelo empresarial, la calidad, terminología, ROI<sup>80</sup>, enseñanza de la localización, la localización de videojuegos, los estándares del sector y el futuro de esta disciplina.

El único libro en español sería el detallado volumen coordinado por Detlef Reineke (2005), *Traducción y Localización: mercado, gestión y tecnologías*, con un

---

<sup>80</sup> ROI.- *Return on Investment* o Rendimiento sobre la Inversión.

fin eminentemente práctico y descriptivo del proceso de localización desde una perspectiva mayoritariamente técnica. El coordinador conjuga eficazmente en este monográfico contribuciones tanto de investigadores en traducción como de informáticos o gestores de proyectos. En los diferentes capítulos se revisan las diferentes vertientes de la localización como el desarrollo de productos de software, la gestión de proyectos, el uso de memorias de traducción, la gestión de terminología, el proceso de localización web, de software o juegos, aunque se centra en aspectos técnicos del proceso que, como mencionamos anteriormente, suelen considerarse de mayor complejidad que la traducción en sí y definen la separación entre ésta y la localización. Desde nuestra perspectiva, en este volumen se echa de menos una mayor atención a aspectos textuales más allá de la gestión terminológica.

Otra publicación centrada en la localización es el trabajo de O'Hagan y Ashworth (2003). Constituye un estudio general sobre la Traducción Mediada por Ordenador (*Translation Mediated Communication*), que parte de la Comunicación Mediada por Ordenador (*Computer Mediated Communication*). Estudia los retos y estudia casos de traducción con Internet como medio de comunicación que va a ser el motor de cambio del el trabajo diario de los traductores. La localización de sitios web la denominan «Teletraducción» (*teletranslation*), que definen en dos vertientes; como «Traducción a través de Internet» (*translation operated via de Internet*) incluyendo traducción de chats, de correos electrónicos etc., y en segundo lugar como «traducción de contenidos para Internet» (*translation of Internet related content*).

El estudio de Cronin (2003), *Translation and Globalization*, supone una exploración de los cambios radicales en la profesión del traductor que conlleva la globalización de la economía mundial, junto con los avances técnicos como

Internet y los programas de memoria de traducción. Se centra, además, en el impacto de la labor del traductor en nuestras sociedades y economías globalizadas.

El proceso de localización se está empezando a incluir como objetivo de tesis doctorales, no obstante, una búsqueda en la base de datos TESEO sólo recuperó dos tesis doctoral presentadas en España sobre localización (Reineke 2002; Bolaños 2002)<sup>81</sup>. Esta última investigadora coincide en el género objeto de estudio, el sitio web corporativo, y ha constituido un grupo de investigación centrado en la investigación descriptiva de este género, DIGALTT (Bolaños *et al.* 2007, 2005). No obstante, los objetivos, el análisis que realiza, y la metodología de estudio difieren de los nuestros, con lo que la presente tesis pretende enriquecer el cuerpo de conocimiento sobre este género en particular. Por centrarse en el mismo objeto de estudio, revisaremos a continuación las diferencias entre ambas investigaciones (véase 2.9.1.1).

Si nos adentramos en la docencia de la localización, han aumentando considerablemente los centros en los que la localización pasa de ser una simple asignatura para convertirse en una especialización propia en traducción. Como ejemplo, proponemos el Centro de Localización de Limerick<sup>82</sup> en Irlanda dirigido por Reinhard Schäler que publica regularmente *The Localization Focus-The International Journal of Localization*<sup>83</sup>, en el que se acaba de iniciar un programa de doctorado en Localización, el master en Tecnologías de la Traducción y Localización de la Universidad Jaume I, el master en Localización del Monterey School of International Studies, o el certificado en Localización de Wake Forest University en el que ha contribuido el autor de la presente tesis.

---

<sup>81</sup> La tesis doctoral de Alicia Bolaños denomina «traducción digital» a la localización web.

<sup>82</sup> <http://www.localisation.ie/>

<sup>83</sup> Es una publicación trimestral que se encuentra en su sexto volumen.

El desarrollo del estudio de la localización nos lleva a apuntar la necesidad de la descomposición del proceso global y la concentración en diferentes aspectos de la misma. Estamos de acuerdo con Keiran Dunne<sup>84</sup> (2005), que aboga por la necesidad de descomponer en proceso en aspectos específicos ya que hasta la fecha se ha intentado realizar estudios que engloban la totalidad del proceso. Esta tendencia se observa en la aparición de investigaciones desde la Traductología sobre los aspectos culturales (MacDonough 2006; Tercedor 2005; Schäler 2002), educativos (Jiménez 2008; Foralón 2006; Altanero 2006; Austerlühl 2006; Reinke 2005; Reinke y Sánchez León 2005; Bermúdez 2005; Mata 2004; Días Fouçes 2004; Archibald 2004; Quirion 2003; Gouadec 2003; *inter alia*), ideológicos (Cronin 2003), tipológicos (Bolaños 2002, Jiménez 2005a), gestión terminológica (Karsch 2006; Schmitz 2005), variación terminológica en localización (Quirion 2006; Bouffard y Caignon 2006), políticos y estratégicos (Cronin 2003; Sandrini 2007), calidad en la localización web (Pierini 2007; Bass 2006; Pailhé 2005), sus diferentes áreas como la localización web (Jiménez 2008a; Sandrini 2008; Mata 2005; Sandrini 2005a; e. o.) o la localización de videojuegos (Muñoz Sánchez 2008; Dietz 2006; Bernal 2006; Vela 2005; *inter alia*) o el uso de herramientas de traducción en localización (Biau y Pym 2006; Sandrini 2005b; Reinke 2004; O'Brien 1998).

### **2.9.1.1. Las investigaciones de Alicia Bolaños (2002, 2003; Bolaños *et al.* 2005, 2007)**

Las investigaciones de Alicia Bolaños han sido pioneras en nuestro país en lo referente a la localización de sitios web. Sus trabajos se centran en un subgénero de nuestro género objeto de estudio, los sitios web de empresas informáticas, y surgen con fines descriptivos y sobre todo didácticos. A pesar de compartir el objeto de estudio, nuestra perspectiva centrada en el proceso de traducción distancia ambos estudios, con lo que pretendemos complementar sus investigaciones y enriquecer

---

<sup>84</sup> Profesor de Localización en Kent State University, EEUU. Comunicación personal durante la Conferencia Anual de la American Translators' Association en Seattle, 10 de noviembre del 2005.

aún más el cuerpo de conocimiento sobre este género. De entre las diferencias principales entre ambos estudios, la metodología de análisis del género de la investigadora se realiza a través de un corpus bilingüe comparable, mientras que en la presente tesis se empleará un corpus monolingüe comparable que constituye una sonda hacia el proceso traductor en sí (véase 7.2). No obstante, coincidimos con la investigadora en incluir los sitios web españoles y estadounidenses, la combinación de mayor importancia económica, aunque nuestro análisis descriptivo se dirige hacia el proceso y producto de la traducción y, consecuentemente, los sitios web estadounidenses se analizan en su localización hacia el español de España. La investigadora reconoce la importancia del análisis de sitios web localizados desde una perspectiva traductológica, y los excluye implícitamente en estas investigaciones para tratarlos en proyectos futuros (Bolaños 2002b: 44).

La segunda gran diferencia metodológica es la inclusión de páginas web como unidad de análisis; el análisis de la investigadora se centra exclusivamente en páginas de inicio, que como se establecerá en nuestro marco teórico, no constituyen un texto sino un subtexto o bloque comunicativo dentro de un género textual (véase 3.4). La investigadora señala en sus primeras investigaciones que la hipertextualidad imposibilita el análisis por textos completos (ibid: 43), en parte debido a las limitaciones técnicas existentes ya superadas en estos últimos cuatro años<sup>85</sup>. A nuestro parecer, las peculiaridades de las páginas de inicio en las que se centra la investigación no pueden ser representativas del género *sitio web corporativo*, puesto que éstas constituyen un índice o puerta de entrada hacia el sitio web global en la que se resumen brevemente los contenidos del sitio web de forma atractiva y, por tanto, su tipología textual se encuentra mayoritariamente marcada por una función primaria apelativa-persuasiva.

---

<sup>85</sup> En una investigación reciente (Bolaños *et al.* 2005) se incluyen sitios web aunque se limitan a tres enlaces de profundidad y el análisis intratextual se centra exclusivamente en las páginas de inicio.

Otra diferencia entre ambos estudios será la introducción del concepto de género y sus repercusiones en nuestra metodología. La metodología de análisis en las investigaciones de la autora surge de la caracterización del género de Shepherd y Watters (1998, 1999), recogida en los cuatro criterios de caracterización de Schmitt-Isler (2000): función, forma, contenido y funcionalidad (véase 5.5.6). Este modelo incluye por tanto tres criterios extratextuales, la función, la forma y la funcionalidad, y uno intratextual, el análisis del contenido. La investigadora prima no obstante la función sobre el resto de categorías (Bolaños *et al.* 2005), que por tanto se aproximaría al modelo de análisis de Göpferich (1995) en el que la función se constituye en el criterio de mayor importancia en el análisis extratextual, constituyendo la base del posterior análisis intratextual. El análisis combina consecuentemente métodos cualitativos, por cuanto a la función y forma, y con cuantitativos por cuanto al contenido y a la funcionalidad. Por nuestra parte, el modelo de caracterización de género que se adoptará en la presente tesis será el presentado por Silvia Gamero con ciertas modificaciones (véase 5.5), ya que por un lado surge desde una perspectiva traductológica, y por otro lado, tal como expondremos más adelante, es más concreto, amplio y específico que el anterior (Schmitt-Isler 2000).

	Imágenes	Marcadores Lingüísticos	Funcionalidad
Elementos de análisis	<ol style="list-style-type: none"> <li>1. Tipo: logo, banner, fotografía, icono u otro.</li> <li>2. Contenido: producto, persona, propio logo, el logo de otra empresa o otro.</li> <li>3. Posición en la página.</li> <li>4. Si es animada o no.</li> </ol>	<ol style="list-style-type: none"> <li>1. Coloquialismos: si su naturaleza es léxica, sintáctica, ortográfica o icónica.</li> <li>2. Verbos: persona, número, tiempo, voz y contenido semántico</li> <li>3. Pronombres personales: persona, número y función (sujeto u objeto)</li> <li>4. Adjetivos: posición (premodificador o postmodificador), contenido semántico.</li> </ol>	<ol style="list-style-type: none"> <li>1. Hipertexto: enlaces internos y externos (tipo, posición y estructura sintáctica)</li> <li>2. Funcionalidad.</li> </ol>

**Tabla 2. Metodología de análisis de subgénero web «sitio web de empresa de informática» de Bolaños *et al.* (2005).**

La primacía de la función como criterio clasificatorio se observa en su estudio posterior en el subgénero *sitios web estadounidenses de empresas informáticas* (Bolaños *et al.* 2005), en el que la analiza por medio del estudio de las imágenes, los marcadores lingüísticos y la funcionalidad. La Tabla 2 resume el conjunto de variables de análisis en el estudio de la investigadora.

Estos criterios seleccionados son objeto de análisis cuantitativo, y nuestra propia metodología incorpora varios de los marcadores lingüísticos y de funcionalidad presentes en este marco metodológico, con lo que será posible una comparación posterior una vez realizado nuestro estudio. No obstante, es necesario indicar que los resultados que se han publicado forman parte de un proyecto de investigación más amplio, DIGALTT (Bolaños *et al.* 2007), cuyos resultados finales aún no se han publicado.

### **2.9.2. Estudios desde el sector de la localización**

Desde el sector de la localización encontramos un gran número de monográficos de carácter eminentemente práctico, aunque estos poseen un mayor énfasis en la parte técnica del proceso, tales como *A practical guide to Localization* (2000) de Bert Esselink, en el que se estudian detalladamente todas las etapas de las que se compone el proceso global de la localización. La importancia de la localización supuso la aparición a principios de los 90 con las primeras monografía en este campo tales como el libro *Software Localization and Internationalization* (Uren *et al.* 1992). En el caso específico de los sitios web, *Beyond Borders: Web globalization strategies* de John Yunker (2003) se centra en la localización de sitios web corporativos desde un punto de vista económico. En ambos se menciona la traducción y los diferentes aspectos asociados con la misma, si bien no dejan de ser superficiales. Por la importancia del lenguaje Xml en la web, hemos de mencionar el libro de Yves Savourel (2001), *XML Internationalization and Localization*, que

sobresale como uno de los trabajos de mayor impacto en los estudios posteriores en localización. Igualmente, el uso de la plataforma Microsoft .NET en la localización constituye el núcleo del libro *Internationalization and Localization* de Nick Symmonds (2003). La labor localizadora de Microsoft se refleja además en el libro *Developing International Software* (Dr. International 2003).

La existencia de organizaciones como la Localization Industry Standards Association (LISA), es capital al realizar una gran función a hora de unificar y normalizar el sector. LISA se fundó en 1990 y desde entonces realiza una gran labor con la gran cantidad de publicaciones sobre localización de software, web, productos móviles, la internacionalización, la recopilación de estadísticas sobre la industria, y sobre todo, con la adopción de estándares tanto de memorias de traducción intercambiables (TMX) o OSCAR. Recientemente, la docencia de la localización se ha convertido en uno de sus objetivos, con su reflejo en la reciente conferencia «Teaching Localization for Global Business Readiness» en diciembre del 2007.

Son escasos los intentos de unir la industria de la localización y los estudios de Traductología, y en ese aspecto, desde el punto de vista de la industria esta el libro *Translating into Success* (Sprung 2000) un intento de la ATA y la industria de hacer puentes en los que intervienen responsables de localización de las empresas más punteras como Microsoft, Canon o Lockheed Martin.

La desconexión entre los interesados en la localización como proceso económico y la Traductología lleva a Schäler (2002, 2005) a remarcar cómo la localización puede empobrecerse sin estudios amplios desde una óptica traductológica. Como se mencionó anteriormente, Anthony Pym señala cómo el sector ha obviado el cuerpo de conocimiento de los últimos 20 años de la Traductología, a pesar de que



ésta podría ser de gran ayuda en ciertas áreas controvertidas como la equivalencia, adaptación o mediación (Pym 2004b).

## 2.10. Recapitulación

La localización surgió a partir de los años 80 centrada en la localización de software y gradualmente se extendió a la totalidad de productos que incorporan textos en formato digital. El sector de la localización instauró sus propios modelos empresariales dejando a lado en parte el cuerpo de conocimiento de la Traductología y, desde entonces, se han alzado las voces por un estudio conjunto de esta disciplina (Pym 2003, 2005; Dunne 2006; Schäler 2002). Hoy en día no existe un consenso claro sobre qué constituye la localización, cómo definirla, o qué aspectos constituyen esta disciplina emergente (Foralón 2006). A través de nuestra revisión crítica sobre este proceso y el análisis de las diferentes propuestas dependientes de su perspectiva de estudio, podemos concluir que la existencia de una adaptación técnica, tanto en la fase previa de Internacionalización como en el control de calidad por un lado, y su carácter dependiente del modo de traducción que la sitúa fuera de la traducción audiovisual por otro (Hurtado Albir 2001), serían dos puntos básicos que la justificarían la existencia de la localización como modalidad diferenciada y objeto de estudio independiente. Ciertos conceptos que han surgido desde la aparición de la localización son de interés a la Traductología en general, tales como el concepto de internacionalización textual, el concepto de *locale* de uso corriente y diario en localización toda vez que en Traductología se debate sobre qué constituye una lengua o dialecto, así como la controvertida equivalencia funcional, que en el entorno delimitado de la localización, pudiera tener una aplicación práctica (Pym 2004a).

De entre los diferentes aspectos de importancia en la localización, hemos desarrollado su pertenencia a un proceso global GILT (Globalización,

Internacionalización, Localización y Traducción) y hemos descrito las interdependencias entre estos distintos procesos. Además, hemos revisado la relativa falta de homogeneidad terminológica en los distintos estudios e investigaciones: por ejemplo, Microsoft utiliza los términos «internacionalización» y «globalización» en el sentido opuesto al del resto del sector (Esselink Y Cadieux 2002), o ciertas publicaciones sobre localización web prefieren denominar globalización web a este proceso (Yunker 2003). Adicionalmente, hemos repasado el uso por parte del sector del concepto «adaptación», tanto técnica como cultural y lingüística en un intento de explicar las diferencias entre la localización y la traducción, la existencia de diferentes grados de localización dependiendo del encargo de la localización, y la escasez de publicaciones en el área debido a la juventud de esta incipiente disciplina.



### 3. EL TEXTO EN LOCALIZACIÓN

La noción contemporánea de «traducción» se desarrolla desde una triple naturaleza: como acto de comunicación, como proceso mental y como operación textual (Hurtado Albir 2001: 40). Como actuación que se produce entre textos, la misma ha de participar de características textuales (ibid: 409). De hecho, es tal la importancia de la noción «texto» en el desarrollo de la Traductología que se ha llegado a identificar como su pilar fundamental y el punto de partida en ciertos paradigmas. Éste sería el caso del modelo textual que representa A. Neubert y cuya visión quedaría recogida en la siguiente cita:

Text is the central defining issue in translation. Texts and their situations define the translation process (Neubert y Shreve 1992: 5).

Por su parte, los estudios sobre localización se han preocupado en mayor medida por aspectos comunicativos, como el mantenimiento de la función de determinados segmentos textuales, ignorando que el proceso traductor-localizador incorpora estas vertientes cognitivas y textuales que se proyectan en el producto final, el texto localizado. Partiendo de esta concepción en el sector sobre la actividad traductora, que ha sido tildada de simplista<sup>86</sup>, no es de extrañar que la mayoría de estudios hayan obviado hasta la fecha uno de los fundamentos de nuestra investigación empírica; qué constituye un «texto» en localización y cómo caracterizarlo desde el punto de vista del proceso traductor. Los estudios surgen mayoritariamente desde la perspectiva técnica del proceso (Quirion 2003), por lo que el concepto «texto» pasa a denominarse:

- a) Material (Esselink 2001)

---

<sup>86</sup> Biau y Pym (2006: 11) señalan que esta concepción del sector de la localización nos traslada a los paradigmas de equivalencia de los años 60.

- b) Parte lingüística (LISA 2004, 2007),
- c) Contenido (Mata 2005; O'Hagan y Ashworth 2003; DePalma 2003; Dunne 2006; LISA 2006)
- d) Documento (Shreve 2006: 323)
- e) *Information elements* (Lockwood 2000)

La presencia del «texto» se ha esquivado en parte por falta de límites que acoten un texto web, la falta de aplicación de un modelo de hipertexto, la autoría múltiple del mismo (Pym 2004a), las peculiaridades de acceso al texto por parte del traductor, o la progresiva desaparición del texto unitario debido a usos tecnológicos<sup>87</sup> en la labor localizadora (Mossop 2006; Biau y Pym 2006). Existen gran cantidad de estudios de fondo en España sobre las particularidades de los textos jurídicos (Borja Albí 1998 2000), textos técnicos (Gamero 1998, 2001, Jiménez Serrano 2002), audiovisuales (Chaume 2004; Díaz Cintas 2003; Angost 1996), científicos (e. g. Gutiérrez Rodilla 2005; Alcina y Gamero 2002; López Rodríguez 2001; Tercedor 1999; Montalt i Resurrecció 2005) o periodísticos (Bustos *et al.* 2005)<sup>88</sup>, no obstante, en nuestra revisión bibliográfica no hemos identificado ninguno que defina claramente que constituye un «texto web» con fines traductológicos. Nuestra tesis incorpora un componente comparativo y, por consiguiente, debemos establecer qué constituye un texto, cómo definirlo y delimitarlo, y cómo afectan las peculiaridades de la textualidad web al proceso de traducción. El presente capítulo está dedicado a la noción de «texto» en Lingüística y en Traductología, para así aproximarnos a la caracterización del texto objeto de estudio, el *sitio web corporativo*.

---

<sup>87</sup> Como pueden ser las memorias de traducción o los *Global Content Management Systems*.

<sup>88</sup> Nos remitimos a los estudios generados en España y se pretende incluir un listado extensivo de los diferentes estudios traductológicos en diferentes campos.

### 3.1. La noción de texto

La noción lingüística de «texto» es relativamente reciente. El desarrollo de las definiciones de texto y sus características varían en consonancia con la perspectiva de estudio, de ahí que estas diferencias dificulten en sobremanera su definición (Janoschka 2003: 159). En el desarrollo de la Lingüística se ha analizado qué constituye un texto desde múltiples perspectivas: en los años 60 los estructuralistas se centraban en las características estructurales de los textos basándose en investigadores como Saussure (1916), Bloomfield (1935) o Chomsky (1965). Las críticas dirigidas a la falta de atención hacia las relaciones semánticas o al contexto supusieron el surgimiento de otra corriente más centrada en aspectos comunicativos, en la que se describe el texto por aspectos funcionales en su situación comunicativa, incorporando al emisor y al receptor como elementos cruciales.

Si nos adentramos en definir este concepto, en primer lugar hemos de reseñar cómo esta importancia cardinal del carácter comunicativo del texto aparece en la definición del investigador alemán Isenberg (1982: 80):

El texto es la unidad primaria de organización en la que se manifiesta el lenguaje humano [...] un texto es, en consecuencia, una unidad comunicativa, o sea, una unidad en la que se organiza la comunicación lingüística (Isenberg 1980, cf. Bernárdez 1982: 50).

Desde la Lingüística textual, Bernárdez (1982) se apoya igualmente en estos aspectos pragmáticos y comunicativos, a los que añade las determinantes coherencia y cohesión aportadas por Halliday y Hasan (1976: 23), incluyendo rasgos procesales de estructuración:

La unidad lingüística comunicativa fundamental, producto de la actividad verbal humana que posee siempre carácter social, está caracterizada por su cierre semántico y comunicativo, así como por su coherencia profunda y superficial, debida a la intención del

hablante de crear un texto íntegro, y a su estructuración mediante dos conjuntos de reglas: las propias del nivel textual y la del sistema de la lengua (Bernárdez 1882: 84).

Además de la importancia de los aspectos comunicativos, desde la Lingüística textual Beaugrande y Dressler (1981) defienden la relevancia de la «textualidad», la existencia de una naturaleza inherente a todo texto que se establezca como tal, determinada por ciertos criterios que ha de presentar:

A COMMUNICATIVE OCCURRENCE which meets seven standards of TEXTUALITY. If any of these standards is not considered to have been satisfied, the text will be non-communicative. Hence, non-communicative texts are treated as non-texts. (Beaugrande y Dressler 1981: 3).

Los siete pilares de la textualidad a los que se refieren son la cohesión, la coherencia, intencionalidad, aceptabilidad, informatividad, situacionalidad, e intertextualidad (ibid). Un texto que no cumpla estos parámetros no es considerado por los autores como tal. Como se verá en el punto 3.3.3, en el caso de los hipertextos son de esencial importancia la coherencia y la cohesión debido a ciertas particularidades en su procesamiento (Janoschka 2003: 165-167), tales como la falta de una secuencia textual fija, el procesamiento textual discontinuo, o la falta de límites hipertextuales reconocibles (Storrer 2002).

En Traductología, el análisis textual se incorpora con objetivos primordialmente didácticos, crítico-evaluativos o investigadores (Hurtado Albir 2001: 409). En esta disciplina se han realizado grandes esfuerzos teóricos encaminados a adoptar una unidad operativa como parámetro sobre el que se desarrolle la actividad traductora.

Dentro de este ámbito, Hatim y Mason (1990) acentúan en su definición la importancia de los aspectos comunicativos ya mencionados, que en estos investigadores se cristalizan en el propósito retórico, justo con la necesidad de que todo texto constituya una unidad coherente y cohesiva:

Text is a coherent and cohesive unit, realized by one or more than one sequence of mutually relevant elements, and serving some overall rhetorical purpose (Hatim y Mason 1990: 178).

Los autores defienden en su obra la importancia de reconocer la «textura», que se define como el conjunto global de aspectos organizativos y cohesivos que contribuyen a que un texto se reciba como una unidad coherente. Remarcan asimismo su relevancia durante el proceso de traducción, haciendo hincapié en la vital importancia de identificar límites a todo texto: «it is of vital importance for translators to identify text boundaries» (1990: 178). Indican además que en la traducción no existe una unidad relevante que se pueda identificar más allá del texto, «beyond the level of the text, it is difficult to perceive any regularly occurring patterns which would enable us to identify a unit of discourse» (ibid). Sitúan por tanto al texto como la unidad máxima operativa sobre la que se asienta la actividad traductora.

Entroncando con el carácter comunicativo de las definiciones presentadas anteriormente, Nord (1991) hereda la interrelación entre los rasgos internos y externos en la definición del texto y lo presenta como: «A text is the totality of communicative signals used in a communicative interaction» (Nord 1991: 14). El texto para la investigadora no se reduce a los elementos lingüísticos, sino que la acción comunicativa se puede realizar por medio de la combinación de medios verbales y no verbales, «A text is a communicative action which can be realized by a combination of verbal a non-verbal means» (Nord 1991: 15). Esta apreciación es esencial en la concepción de «texto» actual en Traductología, y coincidimos con esta afirmación sobre la inevitable importancia de los elementos visuales, interactivos, etc. en la concepción de texto en localización.

En esta misma línea, y dado que nuestro objeto de estudio son textos multimedia, nos es de interés la definición presentada por Susanne Göpferich (1995: 57), que



dentro del enfoque funcionalista en el que se encuadra, y siguiendo la definición de Nord, señala implícitamente la importancia de los elementos gráficos:

Un texto es un todo coherente compuesto por elementos lingüísticos, o gráficos y lingüísticos que presenta una orientación temática y-o funcional. Creado con una intención concreta, la intención comunicativa, cumple una función comunicativa identificable [...] y supone una unidad con un cierre temático y funcional (Göpferich 1995: 57, cf. Fernández Sánchez 2004: 28)<sup>89</sup>.

Si nos adentramos en el ámbito de la Localización, hemos de señalar que las definiciones de «texto» son muy escasas. Anthony Pym parte del contexto de la localización de software y lo define simplemente como: «A text is quite simply whatever unit is distributed as a unit» (Pym 2004a: 17). Para este autor, una de las características más importantes de este tipo de texto localizado es su dependencia de la distribución y la desaparición de la figura del texto origen (ibid: 6):

The localized text is not called on to represent any previous texts, it is instead part of one and the same process of constant material distribution, which starts in one culture and may continue in many others (Pym 2004a: 6).

Sobre esta afirmación quisiéramos aclarar que los hipertextos difieren en cierta manera del texto en programas de software, puesto que éstos no puede definirse como unidad de distribución, quizás más válida en programas de software, sino más bien como unidad de producción o de información, una **unidad que se desarrolle o presente de modo unitario**. En esta particularidad, y dentro del contexto de producción de sitios web en constante actualización (O'Hagan y Ashworth, 2003: 9; Biau y Pym, 2006), encontramos un paralelismo de interés con la definición en el ámbito de la *Information Management*, por la que la información se distribuye en forma de objetos de información digital o OI (Hofmann y Mehnert, 2000). Esta concepción se acerca a nuestra aproximación al texto web (véase 3.4), puesto que éstos constituyen una colección de información que se identifica como una unidad, y se define por su fin comunicativo, el usuario

---

<sup>89</sup> Énfasis propio. Traducción del alemán de Fernández Sánchez (2005: 28).

específico al que se dirige, el negocio producto o empresa a la que representa (una línea de productos o una función corporativa), la información que proporciona (en un formato específico y para una audiencia específica), así como unas restricciones publicitarias (ibid).

A modo de recapitulación, hemos intentado aproximarnos a la definición de texto observando el resultado de varios estudios sobre el mismo, procedimiento que sugiere Bernárdez (1982: 76) a la hora de definir qué constituye un texto.

Ofrecemos a continuación una síntesis de los puntos de interés para nuestra investigación que hemos constatado tanto en definiciones previas de texto como en varios estudios traductológicos. Esta síntesis será la base para nuestra caracterización del texto web que presentaremos con posterioridad (véase 3.4):

- a) Las diferentes perspectivas consideran el texto como una «**unidad**». Desde la lingüística sistémico funcional se indica que una propiedad esencial en todos los textos es el que forman un «todo unitario» o *unified whole* (Halliday y Hasan 1976: 1). En el repaso realizado queda patente que el texto es «la unidad fundamental del lenguaje entendido comunicativamente» (Bernárdez 1982: 40), ya sea a «communicative occurrence» (Beaudegrande y Dressler 1981), a «coherent and cohesive unit» (Hatim y Mason 1990) un «todo coherente» (Göpferich 1995), o simplemente, «a unit» (Pym 2004a).
  
- b) El carácter unitario de todo texto es una naturaleza inherente que se puede denominar **textualidad** (Beaudegrande y Dressler 1981: 3). Ésta aglutina tanto rasgos lingüísticos y semánticos (próximos a teorías estructuralistas y generativas), como pragmáticos (dentro de la perspectiva que le otorga gran importancia a la dimensión extralingüística). La textualidad como una esencia global se relaciona con la *textura*, que se puede definir como «la

propiedad por la cual un texto tiene consistencia lingüística y conceptual, es decir, tiene una continuidad en cuanto al sentido (es coherente), y a los elementos de la superficie (está cohesionado), y tiene una articulación de la evolución de la información (progresión temática)» (Hurtado Albir 2001: 415).

- c) Un texto es **situado**, es decir, sólo se convierte en tal a través de la actualización que realiza un receptor determinado en una situación de recepción concreta (Reiss y Vermeer 1984: 90). En este sentido, este aspecto implica la posibilidad de que pueda haber tantos textos como receptores del mismo (Nord 1997b: 31).
- d) El texto se encuentra por tanto en una situación y debe cumplir ciertos **criterios de textualidad** para que se reconozca como tal (Beaudegrande y Dressler 1981: 3). En los hipertextos, la coherencia se ha señalado como el de mayor importancia (Foltz 1996, Fritz 1999, Storrer 2002).
- e) En un sentido amplio, el texto incluye no solo **elementos verbales** sino que incluye igualmente **elementos no verbales** como gráficos, visuales, tipográficos etc. (Göpferich 1995). Un texto en su conjunto constituye una entidad cuyo valor es superior a la mera suma de sus elementos integrantes verbales y no verbales (Snell-Hornby 1996: 55).
- f) El texto tiene una **configuración determinada**, sea estructura u organización (López Rodríguez 2001: 80). Esta configuración lingüística de un texto se ve determinada por su **función comunicativa**, sin la que no puede existir un texto como tal (Göpferich 1995: 40).

Tras la exposición de esta síntesis, queda patente que las características fundamentales en las que se apoya la noción de texto se proyectan en su concepción en los diferentes modelos, teorías y estudios empíricos en Traductología. Si aplicamos estos aspectos a nuestros objetivos, vemos que en esta disciplina se asume la existencia de un texto origen **unitario** y **estable**, cuyo proceso de lectura y comprensión por parte del traductor se asemeja en cierta manera al de cualquier otro receptor<sup>90</sup>. Asimismo, se suelen usar textos completos, coherentes y cohesivos en los experimentos empíricos que intentan explorar ciertos aspectos del proceso de traducción<sup>91</sup> (véase Orozco 2002).

Por otro lado, la dependencia de una situación de recepción en la realización de un texto se manifiesta en que cada lectura constituye un acto único y dependiente de la situación de recepción: «a process subject to the particular contextual constraints of the occasion, just as much as the production of the texts is» (Beaudegrande 1978: 30). Este aspecto se ha denominado desde la perspectiva textual en Traductología «relatividad translativa» (Neubert 1996), por la que la traducción es el producto de un proceso de lectura y comprensión específico. Los textos web son dependientes de un medio interactivo, Internet, que conlleva la lectura en un espacio reducido, la pantalla. Durante el proceso de traducción el acceso al texto está mediatizado por la tecnología empleada y por **la organización textual interna del texto global, que se encuentra organizado (o no) según criterios informáticos y no comunicativos**. Aparte de cuestiones estructurales, el entorno tecnológico y los avances en producción y distribución textual inherentes a Internet suponen que la concepción de texto global unitario tiende a desaparecer de forma gradual (Mossop 2006; Biau y Pym 2006; Bowker 2006). En este sentido la adopción de *Content Management Systems* (véase 3.3.2), responsables de gran cantidad de sitios web de

---

<sup>90</sup> Estudios empíricos han demostrado no obstante que la lectura es más lenta y se activan mecanismos diferentes (Macizo y Bajo 2005).

<sup>91</sup> Ver Orozco (2002) para una relación de investigaciones empíricas en traducción escrita.

nuestro corpus, conlleva la desaparición del texto definitivo y unitario hacia un texto provisional y dinámico. En palabras de Biau y Pym (2006):

There is no final text, but a constant flow of updated, rearranged, re-sized and user-adapted provisional texts based on a large database of content in constant change (Biau y Pym 2006: 11).

A pesar de todo, defendemos que el traductor ha de trabajar necesariamente con unidades textuales coherentes y cohesivas en los que se reconozca una función o conjunto de funciones<sup>92</sup> (Göpferich 1995). A pesar de la descomposición del texto global en los procesos de localización, la necesidad de los traductores de identificar límites textuales es vital (Hatim y Mason 1990: 178). En este sentido, siguiendo la corriente funcionalista en el estudio de géneros textuales (Göpferich 1995; Fernández Sánchez 2004) y los estudios sobre la textualidad en los hipertextos (Fritz 1999; Storrer 2002; Janoschka 2003: 167), nos centraremos a continuación en el criterio de mayor importancia en el establecimiento de un texto en la web: la coherencia.

### **3.2. Elementos de textualidad: la coherencia en localización**

Como hemos señalado, son varias las condiciones que deben cumplirse para que un texto sea definido como tal (Beaudegrande y Dressler 1981: 3)<sup>93</sup>. Entre otras, en la definición de texto propuesta por Hatim y Mason (1990) se prima una concepción como unidad coherente y cohesiva (Hatim y Mason 1990: 178), que como tal tiene una función o un conjunto de funciones comunicativas (ibid). Independientemente del medio en el que se encuentre, la coherencia es una característica común a todo texto, por lo que ésta ha de recaer igualmente en los

---

<sup>92</sup> En localización defendemos que esta posible deficiencia puede y debe quedar compensada por el conocimiento experto del traductor en esta modalidad de la traducción.

<sup>93</sup> Para los autores todo texto debe cumplir los siete criterios de textualidad (Beaudegrande y Dressler 1981: 3). No obstante, dentro de la corriente funcionalista en Traductología se señala que los únicos rasgos obligatorios son dos, la coherencia y la función textual (Göpferich 1995; Fernández Sánchez 2005).

hipertextos (Storrer 2002: 157). Este particular ha supuesto que gran cantidad de publicaciones sobre textualidad web apunten a la coherencia como el criterio textual de mayor importancia para establecer un texto como tal en este medio (Fritz 1999; Janoschka 2003: 167), además del principal criterio en su producción y comprensión (Fritz 1999: 221).

El concepto de coherencia se encuentra estrechamente relacionado con el concepto de cohesión en la Lingüística textual<sup>94</sup>. La cohesión, según Beaugrande y Dressler (1981), se define en su lingüística textual estructural como:

The way in which the components of the SURFACE TEXT, i.e. the actual words we hear or see, are mutually connected within a sequence. The surface components depend upon each other according to grammatical forms and conventions, such as that cohesion rests upon GRAMMATICAL DEPENDENCIES (Beaugrande y Dressler 1981: 3).

Así, la cohesión se lleva a cabo a través de las relaciones gramaticales y sintácticas en la consecución de unidades textuales. La cohesión es además independiente de la coherencia. La coherencia va más allá de nivel sintáctico y se centra en el significado del texto, conectándola directamente con la comprensión del mismo. Según la definición de Beaugrande y Dressler (1981):

Concerns the ways in which the components of the TEXTUAL WORLD, i.e. the configuration of CONCEPTS and RELATIONS which underline the surface text, are mutually accessible and relevant (Beaugrande y Dressler 1981: 4).

La coherencia se apoya en las relaciones formales y sintácticas que ofrece la cohesión, pero ambas son independientes. Por ejemplo, existen textos como los poemas que no son cohesivos, pero que sí son coherentes, ya que la coherencia se puede producir por medio de una continuación del significado o del sentido, según Beaugrande y Dressler (ibid) un texto «tiene sentido» por la «continuity of

---

<sup>94</sup> Los conceptos de coherencia y cohesión de Beaugrande y Dressler (1981) difieren de la concepción dentro de la Lingüística sistémico–funcional de Halliday y Hasan (1967), que no emplean el término coherencia, sino cohesión léxica, aunque ambos conceptos no son equivalentes.

senses», que serían la base de la coherencia, entre los conocimientos que se activan por las expresiones del texto.

A texts "makes sense" because there is a CONTINUITY OF SENSES among the knowledge activated by the expressions of the text [...] We would define this continuity of senses as the foundation of COHERENCE, being the mutual access and relevance within a configuration of CONCEPTS and RELATIONS (Beaugrande y Dressler 1981: 84).

En esta concepción de coherencia subyace que es esencial la participación del receptor en este proceso. En esta importancia del receptor en su construcción se fundamenta una de sus perspectivas de estudio, la coherencia como parámetro estático en el producto, junto a la perspectiva que la estudia como un parámetro dinámico o como producto (Hurtado Albir 2001: 416).

### **3.2.1. La coherencia como parámetro estático en el producto**

Esta perspectiva, la coherencia como producto, se refiere al reconocimiento de sus características regulables y clasificables en un texto global (Hurtado Albir 2001: 416). Esta apunta a las interrelaciones de tipo lógico semántico (extratextuales) que se dan dentro los elementos que integran un texto (Göpferich 1995a: 50). Estas relaciones se refieren a elementos léxico-semánticos recurrentes, las relaciones lógicas o la progresión temática. Así, al considerar un texto como producto, su estudio se centra en analizar aquellas características objetivas y observables que proporcionen la coherencia a un texto. El análisis de la coherencia como un parámetro estático conlleva su relación directa con el análisis de la cohesión en el texto (Castellà 1992: 158; cf. Hurtado Albir 2001).

Desde la perspectiva del hipertexto y su localización, la perspectiva dinámica que entiende la concepción de la coherencia como proceso es de mayor importancia (Engbresten 2000; Storrer 2002). En parte se debe a la dificultad de estudiar una propiedad estática en textos dinámicos como los hipertextos, en los que el usuario

elige su propio camino, o un *self-selected path* (Fritz 1999). En este caso la aportación del usuario/lector a la construcción de la coherencia es fundamental, por lo que la perspectiva de la coherencia como proceso sería más pertinente a nuestro estudio.

### **3.2.2. La coherencia como parámetro dinámico en el proceso**

La concepción de coherencia como un proceso entraña una perspectiva dinámica en la que se reivindica la importancia del receptor en su proceso de construcción. En ésta se enfatiza, por un lado, la acción del emisor en su construcción y, por otro, la del receptor en su interpretación, analizando además las marcas de los participantes (Hurtado Albir 2001). Dentro de esta perspectiva dinámica, Van Dijk (1988) remarca la importancia de la coherencia como acción mental en lugar de una propiedad textual:

Empirically speaking, discourse does not have coherence, but is assigned coherence by language users (Van Dijk 1988: 62).

Dentro de la Traductología, Baker define la coherencia como la red de relaciones semánticas que organizan y crean un texto estableciendo una continuidad en el sentido (1992: 284), o también como «the result of the interaction between knowledge presented in the text and the reader's own knowledge and experience of the world» (Baker 1992: 219). La investigadora considera además que la aportación del receptor es esencial, por lo que la coherencia no se concibe como una característica intrínseca del texto, sino que el propio lector la infiere en el mismo (ibid: 222). Según la autora, las expectativas del receptor, sus conocimientos y expectativas lingüísticas y no lingüísticas desempeñan un papel esencial en este proceso. De la intervención del receptor depende la coherencia, o la continuidad del sentido pretendida por el emisor que señalaban Beaugrande y Dressler (1981). Así, puesto que la cohesión sería una característica objetiva del texto, ésta aparece de manera explícita en el mismo. Por el contrario, la coherencia es



subjetiva al ser la unidad de significado que percibe el receptor de un mensaje al relacionar el texto con la realidad extralingüística (López Rodríguez 2001: 101).

Para Hatim y Mason (1990: 239) la dependencia de la coherencia del receptor es igualmente importante, ya que la coherencia sería: «the requirement that texts hang together conceptually». Asimismo, y puesto que los textos son específicos de cada cultura, señalan que la coherencia no es un aspecto universal debido a esta dependencia de la intervención del receptor. En los mismos términos se expresa Baker (1992: 222), al señalar que la coherencia como proceso no sólo supone un fenómeno subjetivo dependiente del receptor, sino en cierta medida dependiente de la cultura, y por tanto, de relevancia en el proceso traductor.

La coherencia como parámetro dinámico se redefine en cierta manera en los textos web, por lo que revisaremos en primer lugar el concepto de hipertexto para definirlo y delimitarlo, pasando a revisar la importancia de la coherencia en su procesamiento, para concluir con nuestra definición operativa de texto web para la presente investigación.

### **3.3. El hipertexto**

El acceso a la información en la WWW se realiza normalmente mediante hipertextos enlazados por vínculos<sup>95</sup>. Estos textos poseen ciertas características que los separan de los textos impresos y les confieren un lugar especial en las investigaciones en Traductología.

Desde la aparición del concepto de hipertexto en los años sesenta, se constata que este concepto es heterogéneo, ya que no se clarifica si el «hipertexto» se refiere a

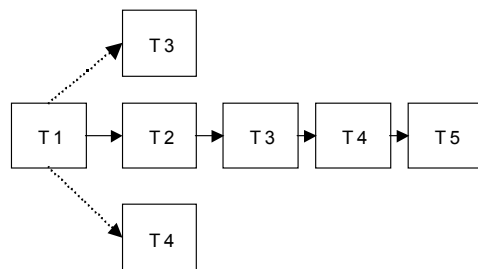
---

<sup>95</sup> Hoy en día existen otros procedimientos para acceder a la información en la WWW, como las subscripciones a servicios de información RSS (véase 9.1.2.4).

páginas web, a los diferentes componentes de una página web como los anuncios web, o a los sitios web completos (Janoschka 2003). Ted Nelson, reconocido por investigadores en hipertexto como el creador de este concepto<sup>96</sup> (Landow 1992: 4), lo define a mediados de los sesenta como:

Non sequential writing-text that branches and allows choices to the reader, best read at an interactive screen (Nelson 1993: 2).

La primacía de la concepción lingüística de texto como «unidad» desaparece en las definiciones de hipertextos. En su lugar, la falta de linealidad, la interactividad y la no impresión en papel se aportan como las primeras características claras del mismo. Para Landow (1992 1997), el hipertexto representa una textualidad diferenciada no lineal y organizada como una red abierta, sin un centro claro, en contraposición con el carácter lineal de todo texto impreso<sup>97</sup>.



**Figura 13. Modelo de texto lineal. Adaptado de Fritz (1997: 222).**

La linealidad y sus implicaciones en la producción y recepción textual es una de las características básicas que separan los hipertextos de los textos impresos (Fritz 1999: 222). En los hipertextos, la aparente falta de linealidad es en realidad una «multilinearidad» (ibid), puesto que el emisor configura previamente la estructura posible del sitio web. En este sentido, la labor del desarrollador de un sitio web se asemejaría más a la de un documentalista que a la de un escritor (Storrer 2002), ya

<sup>96</sup> Dentro de su investigación en a mediados de los 60, *Xanadu*, que no se llegó a implementar en ordenadores.

<sup>97</sup> A excepción de los hipertextos impresos como enciclopedias, etc. (Jucker 2003).

que se basa en la organización de una estructura textual global que permita su fácil uso, actualización y adición de material<sup>98</sup>.

El hipertexto, a diferencia del texto lineal, es una red de elementos textuales que suele tener una estructura, que en la teoría del hipertexto se describe como *axial* (o jerárquica), o en *red*<sup>99</sup> (Landow 1992). Estos dos tipos de estructuras se combinan en la mayoría de los hipertextos, aunque una de ellas suele predominar (Engbresten 2000: 212). En nuestro caso, los sitios web corporativos supondrían un caso de estructura esencialmente jerárquica, en los que subyace además un segundo nivel estructural en *red* debido a la inclusión de menús de navegación en todas ellas<sup>100</sup>. En general, suele ser difícil o casi imposible visualizar la estructura abierta de los hipertextos individuales. Existen, no obstante, ciertas características comunes que se pueden representar en un esquema, como la existencia de hipervínculos o cómo desaparece el concepto de estructura textual en su concepción lineal. Con el objetivo de representar visualmente la estructura hipertextual, Anna Janoschka (2003: 172) revisa las investigaciones previas y presenta una propuesta de representación hipertextual en la que observamos la estructura de navegación de un sitio en el contexto de la WWW, y se señalan como características básicas de los mismos:

- a) La importancia de la multilinealidad desde la perspectiva del usuario.
- b) Los hipertextos son abiertos<sup>101</sup>, es decir, existe la posibilidad de agregar más nodos o añadir enlaces externos.
- c) La inclusión de ciertos elementos textuales<sup>102</sup> independientes.

---

<sup>98</sup> Engbresten (2000) se expresa en los mismos términos en su artículo sobre la coherencia en el hipertexto: «The hypertext designer will have to select the text elements to be interconnected by electronic links. Normally, this selection will also involve choosing a global structure for the system of nodes and links» (Engbresten 2000: 211).

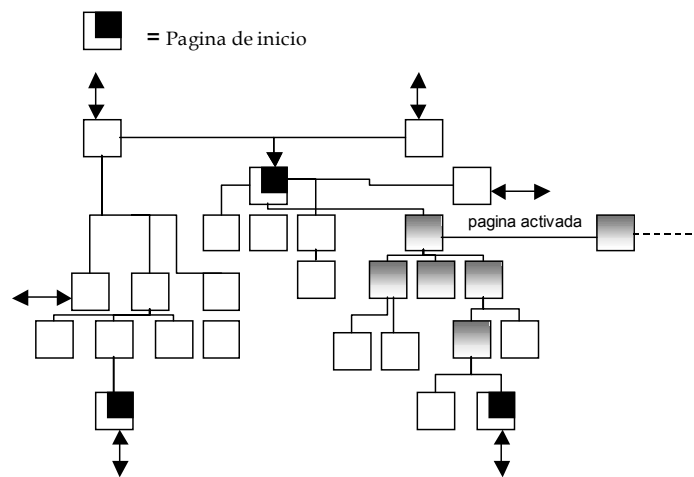
<sup>99</sup> El término empleado en inglés es *networked structure* (Landow 1992). La estructura en red se asemeja a una estructura neuronal sin un orden jerárquico.

<sup>100</sup> Como es el caso de los enlaces de contacto, o términos legales en la totalidad de las páginas.

<sup>101</sup> El término en inglés empleado por la autora es *open-endedness*.

<sup>102</sup> La publicidad web en el caso de Janoschka.

- d) La disminución de la importancia de las páginas de inicio al ser posible acceder a un sitio web por medio del *deep linking* (Nielsen 2000: 179), es decir, llegando directamente a una página interna del sitio web. Así, aun cuando conceptualmente las páginas de inicio son la entrada al sitio web, los hipertextos son descentralizados y no se pueden conceptualizar de manera lineal fija.



**Figura 14. Modelo de la estructura de un hipertexto en el contexto de la WWW. Adaptado de Janoschka (2003: 173).**

La conexión entre las diferentes páginas web de hipertexto se realiza a través de los enlaces o hipervínculos, concepto clave en el estudio de los hipertextos. Dentro del esquema de la autora encontramos enlaces «internos»<sup>103</sup>, es decir, hacia una página web o nodo perteneciente al mismo sitio web, o «externos», hacia una página web de otro hipertexto (Janoschka 2003: 179). Existen además los enlaces «intranodales», es decir, aquellos que apuntan al interior de una misma página cuya longitud sea superior al largo de pantalla, así como los enlaces «implícitos»

<sup>103</sup> Los hipervínculos se componen de un *trigger* o iniciador, de un enlace y de un *target* u objetivo (Janoschka 2003: 179). El *trigger* sería el elemento a activar, el enlace sería la conexión entre estos dos elementos y el *target* sería el nodo al que se accede.

(Engelbrecht 2000), aquellos que se presentan en un mapa o índice del sitio web y que contribuyen a su coherencia global.

### 3.3.1. Clases de hipertextos

Nuestra investigación se basa en la recopilación de un corpus de textos completos y unitarios, por lo que se hace necesario acotar esta aparente falta de límites de los textos web o hipertextos. En este sentido, Storrer (1999, 2002)<sup>104</sup> presenta una clasificación de gran utilidad para limitar qué constituye un hipertexto en una investigación empírica en Traductología. La autora indica la necesidad de diferenciar entre tres tipos de documentos digitales para facilitar los estudios empíricos en hipertextualidad<sup>105</sup>:

- En primer lugar presenta los **e-textos**<sup>106</sup>, textos organizados de forma secuencial disponibles en la WWW. Estos textos suelen ser copias de textos impresos y, por consiguiente, con mecanismos de cohesión y coherencia semejantes como tesis doctorales, artículos de periódico o copias de contratos. Normalmente se publican en formato .pdf o .html por ser una forma económica y eficiente de distribuir documentación. Estos textos suelen formar parte de un **hipertexto** global. Desde el punto de vista del proceso de traducción, estos documentos no representarían un caso especial, ya que su proceso de traducción se asemeja al proceso de traducción de estos mismos textos impresos<sup>107</sup>. Estos textos constituirían géneros textuales «existentes» o *extant* (Shepherd y Watters 1998) que presentaremos en el siguiente capítulo (véase 5.5.6.1).

---

<sup>104</sup> Desde la perspectiva de la Documentación.

<sup>105</sup> En este caso sobre la coherencia en hipertextos (Storrer 2002).

<sup>106</sup> Janoschka (2003: 171) indica que otros autores (Fritz 1999) denominan los e-textos como *textual elements*.

<sup>107</sup> Aun cuando argumentamos que la presencia de código de cualquier tipo en el documento influye de manera determinante en el proceso cognitivo de traducción.

- Los **hipertextos** son textos publicados electrónicamente con una estructura no lineal, con una función textual reconocible y que representan una unidad temática, como puede ser un sitio web corporativo. Son además abiertos ya que los emisores pueden añadir más «nodos»<sup>108</sup>. Estos hipertextos se diferencian de los anteriores en que su contenido se encuentra estructurado en nodos de hipertexto modulares que se encuentran conectados por medio de hipervínculos. El acceso a los diferentes nodos se lleva a cabo por medio de la navegación o la función de búsqueda del sistema hipertextual. Es precisamente su medio, Internet, el hecho diferencial entre estos hipertextos digitales y los textos no secuenciales denominados hipertextos impresos, tales como diccionarios, enciclopedias o periódicos (Jucker 2003).
- Por último, la **hiperweb** interrelaciona a los hipertextos y e-textos por medios de enlaces. La totalidad de la WWW se puede denominar a gran escala una hiperweb. La autora indica que a menor escala una hiperweb se podría caracterizar por un tema, o institución, como el caso de la hiperweb de la Universidad de Granada, [www.ugr.es](http://www.ugr.es), conformada por gran cantidad de hipertextos unitarios enlazados entre sí, los sitios web de diferentes departamentos, facultades o asociaciones de estudiantes. En este sentido, una hiperweb podría equivaler al género web «portal».

Los sitios web corporativos serían por lo tanto *hipertextos*, asociados con una sola compañía y con un solo dominio como puede ser [www.iberdrola.es](http://www.iberdrola.es). En nuestra opinión, estos sitios web corporativos representarían uno de los mejores ejemplos de hipertextos, puesto que casi todos los hipervínculos son «internos»<sup>109</sup>, hacia nodos o páginas dentro del mismo hipertexto. Los hipervínculos «externos», los que se dirigen hacia otros hipertextos de la hiperweb global, son casi inexistentes

---

<sup>108</sup> Los investigadores en hipertexto denominan «nodos» a los diferentes subtextos enlazados por medio de hipervínculos.

<sup>109</sup> Véase 3.3.

puesto que éstos conllevan el abandono del hipertexto por los usuarios, algo contrario a su objetivo (Nielsen y Tahir 2002)<sup>110</sup>.

Nuestra crítica a la propuesta de la investigadora se centra en que todo texto puede ser multifuncional, como sería el caso indiscutible de los textos web, y no exclusivamente responder a una función. Por otro lado, hemos de argumentar que pueden existir *e-textos* que no sean necesariamente copias de textos impresos, tal como las noticias en formato RSS producidas exclusivamente para la web (véase 9.1.2.4). A pesar de las posibles deficiencias de la propuesta de la autora, argumentamos que la separación que realiza la autora es válida y necesaria desde un punto de vista traductológico. Desde nuestra perspectiva es primordial separar los textos que se pueden investigar junto con los impresos, los **e-textos**, de aquellos que representan una textualidad diferenciada y dependiente del medio, los **hipertextos** y la **hiperweb**.

Si tendemos un puente entre esta clasificación textual y la Traductología, en la misma se observa un paralelismo de interés con la propuesta de tipos de textos simples, complejos y complementarios de Reiss y Vermeer (1984). Los e-textos se podrían asociar con los textos simples; los textos complejos son aquellos que pueden contener en su interior ejemplares de otros géneros textuales, por ejemplo un sitio web o hipertexto puede contener un contrato o una receta que constituyen e-textos o textos simples, pero un artículo de prensa en línea no puede contener un sitio web. Los hipertextos y la hiperweb constituirían por tanto textos complejos. Los textos complementarios son dependientes de un texto primario para su realización, como una reseña bibliográfica, o el resumen de una obra de teatro. En la WWW los resultados de un motor de búsqueda como Google constituirían un texto complementario ya que depende de textos primarios, los sitios web en sí, para su realización.

---

<sup>110</sup> Este no obstante puede ser el objetivo de los «motores de búsqueda» como Google, cuya función es dirigir al usuario hacia un hipertexto externo.

Pasamos a continuación a repasar las características específicas de los hipertextos

### **3.3.2. Características específicas de los hipertextos**

Son varias las características generales que definen a todo hipertexto. A continuación las esbozamos siguiendo la propuesta de O'Hagan y Ashworth (2003: 13), si bien complementada con aportaciones propias y de otras investigaciones:

- a) El texto se lee en la pantalla y no en papel, por lo que se dificulta la lectura (Pym 2004a), en general un 25% más lenta (Nielsen y Tahir 2002).
- b) El texto suele tener componentes multimedia, como audio, gráficos, iconos que tienen que ser apropiados para las normas de la lengua y cultura meta. Todos estos elementos son parte integrante del texto (Göpferich 1995) y su adaptación cultural es parte integrante del proceso de localización (Tercedor 2005).
- c) El texto se puede «adaptar» a los receptores de la lengua de destino incluyendo cambios de contenido. La adaptación de un hipertexto durante el proceso de localización estaría sujeta al grado de localización expuesto con anterioridad (véase 2.7).
- d) Los posibles receptores no son homogéneos sino una audiencia de una gran variedad. En Internet, cualquier receptor puede potencialmente llegar un sitio web. Desde los estudios de usabilidad de sitios corporativos se señala que no se puede identificar a un «receptor único medio» (Krug 2005: 135). No obstante, desde la perspectiva prescriptiva de los manuales de estilo y



desarrollo web se apunta al reconocimiento de las necesidades del usuario medio como paso previo a su producción (Price y Price 2002).

- e) El hipertexto<sup>111</sup> se puede leer en un orden diferente y, por lo tanto, en contextos diferentes dependiendo de los hipervínculos que se sigan. De igual manera, durante el proceso de traducción el traductor accede al texto de manera no lineal ya que como apuntamos, el texto al que accede el traductor está organizado (o no) por criterios informáticos y no necesariamente comunicativos.

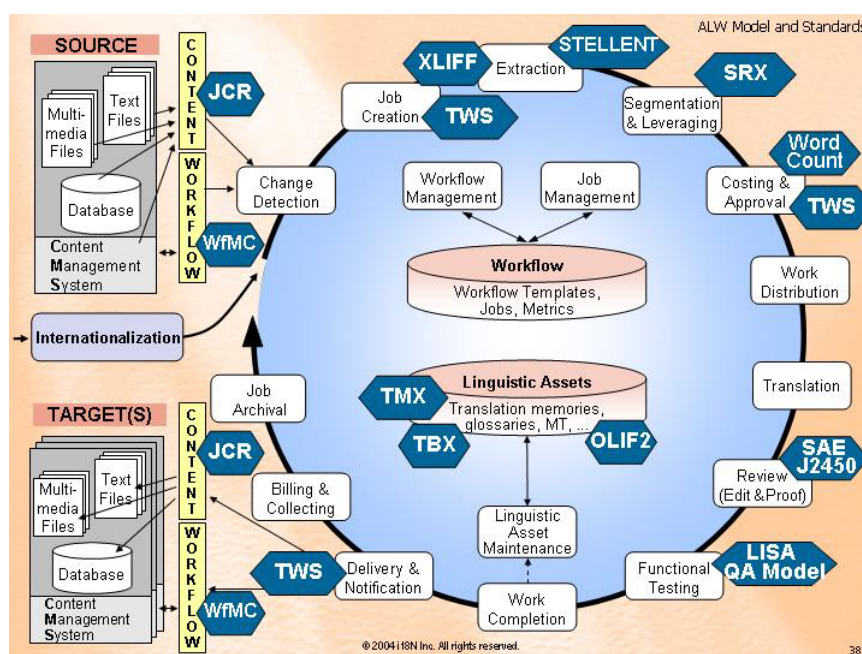


Figura 15. Esquema de *Global Content Management System*<sup>112</sup>.

- f) El texto se suele actualizar y cambiar con más frecuencia que el texto impreso, y a menudo el tiempo es un factor a tener en cuenta (LISA 2006: 4). Un ejemplo sería una promoción que caduca a los tres días. El sector de la localización, respondiendo a los retos que supuso la organización de

<sup>111</sup> El estudio de O'Hagan y Ashworth (2003) no define que constituye un texto o su contexto en localización.

<sup>112</sup> <http://www.i18n.ca/publications/GlobalizingContentModelStandards.jpg>

contenido web y su rápida traducción hacia diferentes lenguas, comenzó a desarrollar a finales de los años 80 los *Global Management Systems* o GMS<sup>113</sup> (LISA 2006; Yunker 2003: 355; Brooks 2000). Según LISA (2006), estos sistemas se definen como:

Systems and processes used to create, manage, publish and archive information as text, images, prices statistics, measurements, articles, documents, descriptions, sound bytes, video clips etc, in multiple languages and for use in multiple countries, regions and/or markets. A truly global content management solution provides the necessary infrastructure to enable an organization to ensure that its content is ready for publishing when and where it is needed, in the right format language and form. In addition, it should do so efficiently and cost-effectively, and be well integrated with the organization's other systems and processes (LISA 2006: 4).

Estos sistemas influyen en el proceso de traducción, puesto que actualizan solamente los segmentos textuales nuevos en un sitio web, los comparan con las memorias de traducción existentes, y separan aquellos segmentos sin traducción que hace llegar al traductor. En estos casos, también se agudiza la falta de contexto y co-texto aun cuando el traductor tenga acceso a las páginas web (Cheng 2000). También se ha señalado que aunque en localización se reutilice el material traducido lo más posible (Schäler 2002), es necesario editar incluso las equivalencias del 100%, puesto que el contexto de los nuevos segmentos puede ser diferente (LISA 2004: 25). Estos sistemas se asemejan a los *Content Management Systems* en que poseen varias funcionalidades dirigidas a la localización, tales como almacenar el contenido o el texto en segmentos en una o varias lenguas y producir páginas web diferentes según las preferencias del usuario (Yunker 2003: 354). No obstante, no están diseñados específicamente para el proceso de traducción (LISA 2006: 2). Estos CMS separan además el contenido de la forma, es decir, el texto del código y lo insertan de forma dinámica,

---

<sup>113</sup>Entre otros [www.glides.com](http://www.glides.com), [www.sdlint.com](http://www.sdlint.com), [www.globalsight.com](http://www.globalsight.com).

facilitando así la localización web al hacer más eficiente los procesos y, por tanto, reducir costes (Lynch 2006).

### 3.3.3. La coherencia en hipertextos

El estudio de la coherencia en hipertextos requiere una perspectiva dinámica y se circunscribe principalmente al estudio de la misma parámetro dinámico (Janoschka 2003; véase 3.2.2). La coherencia textual, como estructura lógica subyacente que guía al lector a través del texto (Neubert y Shreve 1992: 95), agrega a las palabras y construcciones una mayor significado que no tendrían por sí mismas, y se redefine en cierta manera en los hipertextos ya que no existe un «camino» predefinido que el usuario necesita seguir durante la fase de comprensión del texto. La progresión temática no desempeña por tanto la misma función que en textos lineales. Fritz (1999) señala la importancia de este *self-selected path* en la construcción de la coherencia en hipertextos:

Users make sense of a path or a segment of a path by seeing sequences of textual elements as realizations of sequencing patterns and by drawing inferences on the basis of their local and general knowledge (Fritz 1999: 223).

Esta posibilidad de elección surge de la estructuración hipertextual, ya que éstos han de dividirse en subtextos para ajustarlos al medio que los presenta en una pantalla de espacio limitado (O'Hagan y Ashworth 2003: 13). Cada subtexto constituye una unidad de contenido, información y almacenamiento en el nuevo medio (Nielsen y Loranger 2006). Según la perspectiva de estudio estos subtextos se suelen denominar «nodos» (Landow 1992), «lexias»<sup>114</sup> (Barthes 1977), «elementos textuales» (Fritz 1999), «secciones», (Codina 2003: 146), «hiperdocumentos» (Crownston y Williams 1999), «topics» (Horton 1990).

---

<sup>114</sup> El término «lexia» se encuentra asociado a una perspectiva literaria del hipertexto (Codina 2003), en contraposición al más común en la teoría del hipertexto, «nodo».

En el establecimiento de la coherencia hipertextual se introducen dos conceptos que surgen del estudio de la misma como proceso, la **coherencia local y global** (Storrer 2002). De acuerdo con la Lingüística textual, la primera se produce entre elementos consecutivos de un texto, o en un contexto más amplio, entre dos elementos semióticos consecutivos, tal como la relación existente entre una imagen y su descripción a través de una etiqueta «alt» en una página web. En la teoría del hipertexto este tipo de coherencia se relaciona con la **coherencia intranodal**, la que existe dentro de una misma página web a diferentes niveles (Engbresten 2000). Asimismo, ésta se corresponde con el concepto de **coherencia internodal**, la que se produce entre dos «nodos» o elementos que se procesan secuencialmente en un hipertexto, como la relación entre una página web que se activa con un hiperenlace y el nodo<sup>115</sup> en el que se encuentra.

La coherencia internodal es más flexible que la intranodal, ya que ésta puede producirse de manera más libre al no ser necesario explicitar en el texto la relación de coherencia, sino que se interpreta por parte del usuario. No obstante, es esencial que el receptor comprenda la relación entre el hipervínculo y el nodo que surja para que se mantenga la **continuación del sentido** (Beaudegrande y Dressler 1980: 84). En caso contrario, el receptor asumiría un fallo dentro del hipertexto al no interpretar relación de coherencia alguna. Otra área de estudio en este campo sería la semiótica del hipervínculo que tiene como objeto de estudio las relaciones semánticas que se producen en los hipertextos (Neumüller 2004).

La **coherencia global** define la unión entre la totalidad de los elementos constitutivos de un texto, y está mediada tanto por el tema global del mismo como por su función retórica en el contexto más amplio. En este sentido, la coherencia global desempeña un papel de mayor importancia en los hipertextos que en los textos lineales. Esta noción lingüístico-textual está relacionada con la **coherencia**

---

<sup>115</sup> Un «nodo» en la teoría del hipertexto no se corresponde con una página web. Por ejemplo, un nodo puede ser un anuncio en una página que activa un sitio web destino si el usuario actúa en el mismo.

**estructural** o **hipertextual**, que Engebresten (2000) identifica exclusivamente con la lógica que se refleja en la estructura textual global y gobierna el sistema de nodos o elementos textuales enlazados en un hipertexto. No obstante, nuestra concepción de coherencia hipertextual se define no sólo por la lógica subyacente a la estructuración, sino que supone una base global de comprensión que media en la recepción de un hipertexto en su totalidad. Por ejemplo, la cohesión global de un sitio web se produce mediante la inclusión de menús de navegación en todas las páginas como mecanismo cohesionador, o por presentación tipográfica y visual homogénea.

Dentro del establecimiento de la coherencia en los hipertextos, Storrer (2002) señala los siguientes problemas en la construcción de la coherencia: (1) el procesamiento textual discontinuo, (2) la falta de límites textuales fácilmente observables que conllevan *informational shortsightedness*<sup>116</sup> (Conklin 1987: 40), (3) o la falta de una secuencia de procesamiento fija. La resolución de estos problemas específicos de hipertextos conllevan la inclusión de pistas de coherencia, *coherence cues*, específicas de los hipertextos, como son la inclusión permanente de diferentes menús de navegación que sitúan al nodo procesado dentro de la estructura global<sup>117</sup>, los mapas web que muestran la estructura global del texto y pueden resolver la *informational shortsightedness*, o la homogeneidad visual y tipográfica que cohesionan visualmente las páginas localizadas, una de las principales indicaciones en estudios sobre globalización web (Yunker 2003; Hunt 1995). Por ejemplo, en los sitios web corporativos el emblema o logo de la empresa se sitúa en la esquina superior izquierda de todas las páginas (Nielsen y Tahir 2002; Adkisson 2002; Brinck 2006; Bolaños 2002), y refuerza la relación de coherencia internodal y estructural dentro de un mismo hipertexto.

---

<sup>116</sup> Esta se define como la falta de apreciación de límites claros por parte de los usuarios que conlleva el posible abandono de un sitio, ya que el usuario no aprecia un «todo global» cuya lectura deba completarse.

<sup>117</sup> En el contexto del diseño web este concepto se denomina «breadcrumbs» o la guía para conocer exactamente en qué lugar de la estructura jerárquica del sitio web se encuentra el usuario. Su ubicación convencional es en la esquina superior izquierda del marco de contenido de cada página.

No obstante, estas pistas para la coherencia específicas de los hipertextos pueden faltar o no estar presentes en la misma medida durante el proceso de localización web, con lo que los mecanismos de construcción de la obligatoria coherencia en el proceso de comprensión lectora de los traductores podrían diferir en parte.

### 3.4. El sitio web como texto unitario: hacia una definición

En los albores de Internet la unidad textual la constituían las páginas web, dado que el contenido web se concentraba en una página o pantalla única (Mata 2005). Estos inicios siguen estado presentes en ciertas investigaciones cuya unidad textual operativa es la «página web» (Santini 2007a, 2005a; Biber 2004; Rennau 2004). En nuestra opinión, esto hecho no se corresponde con el desarrollo actual de los textos web, ya que dada su evolución actual un sitio web por pequeño que sea consta de varias páginas enlazadas. Como se mencionó anteriormente, las páginas web no constituirían más que unidades de almacenamiento e información/contenido (Nielsen y Loranger 2006), y por tanto, subtextos.

La revisión de las particularidades de los *hipertextos* (Storrer 2002, y sus mecanismos de coherencia, nos lleva a considerar al **sitio web o hipertexto como texto unitario**, incluyendo tanto aspectos gráficos, visuales y tipográficos, así como los subtextos acompañantes (Göpferich 1995). En este caso, los entendemos como aquellos textos que el receptor no lee puesto que, aunque cada lectura de un hipertexto será única al seleccionar cada usuario su propio «camino» (Janoschka 2003). Según Göpferich (1995), los textos acompañantes son esenciales para llevar a cabo la función comunicativa del mismo. De hecho, una página de inicio no tiene sentido alguno si el usuario inicia su lectura con el propósito de encontrar cierta información precisa y el enlace no funciona. En este caso el texto no habría completado la función comunicativa por la que el usuario se dirigió al mismo. No

obstante, en ciertos géneros digitales, como los periódicos digitales o los *blogs*, estos subtextos constituyen *E-textos* (Storrer 2002), textos totalmente coherentes internamente y que constituyen una unidad, por lo que poseen una función pragmática por sí mismos.

Si reintroducimos los rasgos básicos que ya observamos en las definiciones de texto en Lingüística y Traductología (véase 3.1), el sitio web corporativo supondría un texto unitario ya que:

1. Un sitio web corporativo eso es una **unidad** o «todo unitario» (Halliday y Hasan 1976: 1). Un sitio web es una **unidad coherente**, a «coherent and cohesive unit» (Hatim y Mason 1990), ya que la unidad de coherencia global de tal texto se circunscribe exclusivamente al sitio web. En los sitios web corporativos, la demarcación textual cobra una gran importancia, quedando este aspecto implícitamente recogido en la existencia de subtextos que advierten al usuario al activar un hipervínculo externo. La siguiente captura de pantalla del sitio web de Hewlett-Packard demuestra la importancia de delimitación en estos textos:

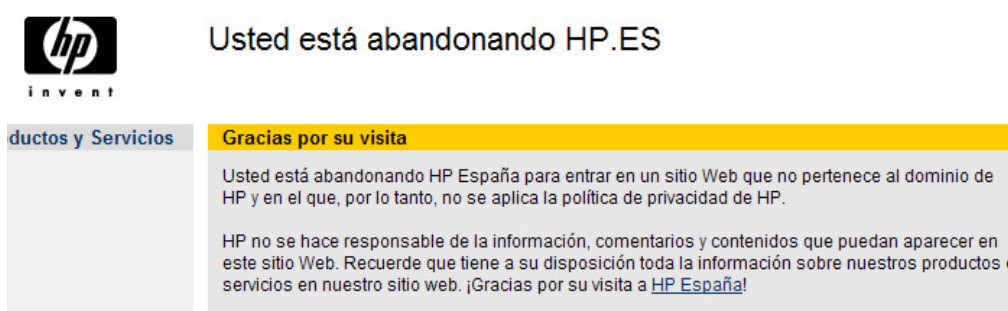


Figura 16. Aviso de abandono de un sitio web global como texto web unitario.

2. El carácter unitario de todo texto es una naturaleza inherente que se puede denominar **textualidad** (Beaugrande y Dressler 1981: 3). Un sitio web supone una unidad tanto de producción como de recepción, asociado con un dominio

único, [www.nombredeempresa.com](http://www.nombredeempresa.com). Los usuarios visitan un sitio web concreto en el que identifican un emisor único, una unidad temática, y una función textual determinada. Esta concepción unitaria se observa en la identificación común entre los posibles receptores que conforman su comunidad de discurso: los usuarios visitan el sitio web de Telefónica o de American Airlines.

3. Un sitio web corporativo constituye un todo **coherente** (Foltz 1996, Fritz 1999, Storrer 2002), tanto a nivel interno de cada página web, intranodalmente, como en el conjunto del sitio, internodalmente (Engebresten 2000). Cumple además ciertos criterios de textualidad necesarios (Beaudegrande y Dressler 1981: 3), como la coherencia o una intencionalidad reflejada en un conjunto de funciones que, como expondremos en el capítulo 8, prima la expositiva con una función secundaria exhortativa (véase 8.1.1).
4. El texto en el sitio web corporativo engloba no solo **elementos verbales**, sino que incluye igualmente tanto **elementos gráficos, visuales, tipográficos** (Göpferich 1995), como **estructurales, multimedia o sonoros**. Además, puesto que la funcionalidad en hipertextos es primordial (Shepherd y Watters 1999), hemos de agregar como parte integrante del texto las **etiquetas html** o xml (Santini 2005a), las **hojas de estilo** CSS o XSL<sup>118</sup> y los **scripts** (Kennedy y Shepherd 2005). Las etiquetas, que en la traducción de sitios web pueden ser separadas directamente los programas de traducción asistida, definen la función o la estructura de presentación de segmentos textuales, tal como su importancia con las etiquetas <H1>, etc., la posición o función del segmento en pantalla con <alt>, <title> o <select>, o la presencia de un *script* que indique interacción.
5. El sitio web corporativo tiene una **estructura jerárquica convencionalizada** (Nielsen y Tahir 2002; Adkisson 2002) que es **multilineal** (Fritz 1999; Janoschka

---

<sup>118</sup> *Cascading Style Sheets* o *XML StyleSheet Language* que presentan la configuración tipográfica de una página o sitio web.



2003). La estructura global del sitio web corporativo se subdivide en subtítulos de orden primario, las páginas web, o de orden secundario, como la publicidad inserta en una página concreta. A su vez, el sitio web global está configurado por una superestructura prototípica, que en estos sitios se puede asociar con las diferentes páginas que el usuario espera encontrar (Askehave y Nielsen 2005).

6. Los sitios web constituirían **textos complejos** (Reiss y Vermeer 1984), ya que pueden incorporar ejemplares de otros géneros en su estructura abierta. Por ejemplo, los usuarios esperan encontrar el manual de instrucciones en formato .pdf en el sitio web de la empresa.

### 3.4.1. Definición de texto en localización

De esta revisión de la concepción del texto en un sitio web corporativo surge por tanto nuestra definición operativa de texto:

Desde una perspectiva traductológica, el texto en localización constituye:

*Una unidad textual coherente en un medio interactivo que se desarrolla y se presenta al usuario de modo unitario.*

Desde una perspectiva traductológica, y debido a la necesidad de contar con una unidad textual operativa de límites reconocibles (que no fijos o necesariamente estables) en la que recaiga la coherencia global, un hipertexto en Internet lo constituye:

*Una unidad textual coherente que se desarrolla y se presenta al usuario de modo unitario en la WWW. La misma presenta una estructura multilineal de subtítulos enlazados por vínculos, conlleva una unidad temática, y una*

*unidad funcional que se refleja en una función dominante junto a una o varias funciones subsidiarias, y se expresa no sólo a través de aspectos lingüísticos sino gráficos, tipográficos e interactivos.*



## 4. LAS TIPOLOGÍAS TEXTUALES EN TRADUCTOLOGÍA

Los estudios traductológicos contrastivos que analizan, comparan o describen pares textuales se han de basar en criterios de selección objetivos y explicitables. Estas pautas guían tanto la recopilación de los textos que conforman el corpus como el método de análisis y, por consiguiente, la validez y replicabilidad de la investigación. En la búsqueda de la delimitación clara del nuestro objetivo, un estudio traductológico contrastivo de sitios web corporativos, hemos delimitado y definido qué constituye específicamente un «texto» en un sitio web corporativo. La siguiente cuestión inexorable e ineludible sería ¿constituyen estos textos una realidad clasificable? En caso positivo, ¿cómo se puede establecer una clasificación textual que garantice la comparabilidad textual de las dos secciones del corpus? En función de cómo se defina variarán tanto la perspectiva y método de análisis, como los criterios de recopilación del corpus comparable. Puesto que nuestro enfoque es traductológico, nos vemos abocados a revisar los conceptos clasificatorios más empleados en Traductología para la clasificación textual, el «género» y «tipo textual».

Preliminarmente y a grandes rasgos, el género y la tipología textual se corresponden con criterios externos o internos de caracterización textual, siendo ambos fundamentales en investigaciones empíricas que recopilen corpus en su metodología. Estas dos vertientes las recoge Pérez Hernández (2002) al sintetizar los posibles criterios que guían la selección textual en investigaciones de corpus, y los identifica con:

- Criterios externos: son esencialmente no lingüísticos; determinan el tipo de género, modalidad, origen y finalidad de los textos que han de incluirse, por lo que son los que aseguran que el corpus represente una amplia variedad de contextos situacionales.
- Criterios internos: identifican tipos diferentes de textos, de acuerdo con una serie de categorías lingüísticas (Pérez Hernández 2002: apartado 4.3.).

Ambos son por consiguiente de interés en las taxonomías textuales necesarias en las investigaciones empíricas de corpus, quedando justificado su análisis para los fines de la presente tesis.

A pesar de las diferencias existentes desde la aparición de estos conceptos, durante los últimos años ha aparecido en Traductología un consenso más amplio sobre el concepto de tipo textual (Reiss 1971; Reiss y Vermeer 1984; Köller 1979; House 1977; Arntz 1988; Hatim y Mason 1990,1997; Nord 1996; Sager 1997; Kussmaul 1997; López Rodríguez 2001), y género (Göpferich 1995; Agost 1996, 1999; Gamero 1998, 2001; Borja Albí 2003; García Izquierdo 2002, 2005; Bolaños 2002). Por consiguiente, la construcción de nuestro marco teórico está supeditada a acotar qué entendemos por tipo textual y género, para justificar a continuación qué criterios de clasificación y qué modelo de análisis se ajusta mejor a nuestros fines. Este aspecto condicionará tanto la composición del corpus (Swales 2004), como la máxima garantía comparabilidad, aspecto clave en investigaciones basadas en corpus comparables (Baroni y Bernardini 2003). En nuestra revisión no hemos incluido otros conceptos lingüísticos como el «registro» o el «discurso»<sup>119</sup>, que no se han introducido de manera importante como criterio clasificatorio en la investigación empírica en los Estudios de Traducción basados en corpus<sup>120</sup>. El registro, en un sentido más restringido, se asocia la variedad de lengua en ciertas profesiones (Trosborg 1997), y puesto que entraña categorías de clasificación

---

<sup>119</sup> Hatim y Mason (1990) identifican el discurso dentro de la dimensión semiótica del texto y lo definen como la actitud ideológica que transmite el texto, como el discurso racista, feminista etc.

<sup>120</sup> Catford (1965) ya introdujo el registro en Traductología, y no fue hasta Bell (1991) y Hatim y Mason (1990; 1997: 55) que el registro se sitúa como concepto esencial en esta disciplina siguiendo los principios de la lingüística sistémico funcional. No obstante, el registro se suele asociar exclusivamente con el área profesional, según la definición de Trosborg (1997: 5): «Registers comprise an open-ended set of varieties (or styles) of language typical of occupational fields».

bastante abstractas, su posible valor taxonómico sería escaso en un estudio de corpus formado por sitios web.

#### **4.1. Las clasificaciones textuales y la confusión terminológica y conceptual**

El interés por las clasificaciones textuales se introdujo en Traductología mayoritariamente desde la Lingüística y el Análisis del Discurso. En esta última disciplina, el debate sobre las clasificaciones textuales ha existido durante miles de años, remontándose a la retórica de Aristóteles<sup>121</sup>. En Lingüística existe desde los años sesenta, con gran cantidad de propuestas que gradualmente se fueron introduciendo en la teoría, práctica, y docencia de una incipiente Traductología. Reiss y Vermeer (1984) fueron pioneros en defender clasificaciones de validez y aplicación exclusiva en Traductología, con lo que se liberó a los investigadores de la necesidad de validar sus propuestas en otras disciplinas. Puesto que esta constituye nuestra perspectiva de estudio, nos centraremos primordialmente en analizar las propuestas de clasificación de tipo textual y-o de género desde nuestra óptica traductológica (véase 4.2.1.)

Las diferentes perspectivas de estudio en el contexto de la Lingüística quedan recogidas por Hurtado Albir (2001: 451-460), que tras señalar la necesidad e importancia de las tipologías textuales en el estudio y didáctica de la traducción, las agrupa entre aquellas que se basan en:

- a) La búsqueda de rasgos distintivos (Longacre 1983; Biber 1989; etc.)

---

<sup>121</sup> La Retórica durante siglos se ha asociado con los usos del lenguaje, y de manera más específica también se refiere a los modos del discurso que se llevan a cabo por medio de tipos textuales (Trosborg 1997: 19).

- b) Las clasificaciones de base contextual (Werlich 1975, 1976; Bronckart 1985; etc.)
- c) Las clasificaciones de base funcional (Adam 1991; Bustos 1996; etc.)
- d) Las clasificaciones por géneros (Swales 1990; Bhatia 1993; etc.)

En la presente tesis empleamos la terminología propuesta por la investigadora que disgrega las clasificaciones textuales en general, denominadas «tipologías textuales» o «clasificaciones textuales», de las diferentes taxonomías que se fundamentan en el «tipo textual», el «género»<sup>122</sup>, o la «clase textual»<sup>123</sup>. La autora recoge los problemas por la confusión conceptual y terminológica existente hasta finales de los años noventa, debidos sobre todo a la diversidad de perspectivas que revisan estos conceptos (Hurtado Albir 2001: 469-470). En nuestra búsqueda bibliográfica sobre tipologías textuales hemos constatado que hasta finales de los noventa los conceptos «tipo textual» y «género» son a veces intercambiables (Biber 1995), se evitan deliberadamente (Laviosa 1997; Stubbs 1996), o incluso se llegan a aplicar de forma distinta y con sentidos diferentes. Esta confusión conceptual llega a suponer su omisión deliberada para evitar problemas de orden metodológico o teórico tanto en Lingüística, «some authors distinguish between text type and genre. I will not» (Stubbs 1996: 11), como en Traductología:

The terms subject domain, subject field, text category and text genre are used interchangeably in this study. [...] I have deliberately chosen to avoid the words 'genre' and 'type' (Laviosa 1998: 566).

Varios estudios atribuyen la confusión inicial a que los términos empleados en el ámbito germánico para «tipo textual» equivalían a *textsorte* y *texttype*<sup>124</sup>

---

<sup>122</sup> Anna Trosborg también emplea el concepto de «text type» o «tipología textual» en un sentido amplio para referirse a cualquier tipo de texto distintivo, incluyendo esta noción el concepto de «género» (1997: viii).

<sup>123</sup> La clase textual, o *textsorte* del alemán (Reiss y Vermeer 1984; Göpferich 1995) corresponde con el actual concepto de género en Traductología.

<sup>124</sup> Proviene de los estudios de Reiss (1971, 1976, 1977).

(Kusssmaul 1997: 97), es decir, «clase de tipo textual» y «tipo textual»<sup>125</sup>; el *textsorte* o clase de tipo textual se asemejaría al concepto actual de género en Traductología, es decir, textos que se producen en situaciones comunicativas recurrentes (Hurtado Albir 2001)<sup>126</sup>. Otra fuente de confusión se debe a que el concepto más empleado en Traductología desde los años sesenta fue el tipo textual, y a su vez, en el que recae una mayor confusión terminológica, (Sánchez Trigo 2002: 123), sobre todo por que se ha empleado ampliamente como cajón de sastre que incluye diferentes instancias de variación textual (López Rodríguez 2001: 90).

La diferencia esencial entre ambos se resume en que el primero se define por una combinación de factores en los que predominan los externos y socioculturales, mientras que los tipos textuales se definen exclusivamente por sus características lingüísticas internas. Ambos se encuentran asimismo convencionalizados en distintas lenguas: el género en relación a su estructura externa y su proyección sociocultural y la tipología por a su estructura lingüística (García Izquierdo y Monzó Nebot 2003). Los géneros además son en principio ilimitados y en constante evolución (Miller 1984), mientras que los tipos textuales son conjuntos cerrados con un número limitado de categorías (Trosborg 1997: 15).

Durante los años noventa aparecieron ciertos estudios cuyo objetivo se centró en compartimentalizar y delimitar estos conceptos. En la introducción a su volumen sobre las tipologías textuales, Anna Trosborg (1997) se propone como objetivo clarificar la confusión conceptual existente en las diferentes disciplinas en torno al género, tipología textual, registro y discurso:

For some systemicists, genres are sometimes used in a broad sense to refer to register variation, such as journalistic language, legal language, scientific discourse etc. Other

---

<sup>125</sup> Traducción del alemán de Martín y Reina (1996) del libro de Reiss y Vermeer (1984).

<sup>126</sup> Reiss y Vermeer definen los «Textsorte» como «los tipos de actos de habla orales o escritos de carácter supraindividual y sujetos a actos comunicativos recurrentes, que han generado unos modelos característicos en el uso de la lengua y en la estructuración de textos precisamente debido a su constante repetición» (1984/1996: 154).



scholars mix genres with rhetorical types, naming expositions and argumentative texts as genres. How does the classification of genres work? How do genres relate to register and text types? (Trosborg 1997: viii).

El volumen de la autora supuso un gran paso hacia la desambiguación de los este conjunto de conceptos clasificatorios en Traductología. La perspectiva integradora de esta investigadora apoya su interrelación y la importancia de un análisis a múltiples niveles que combine estos conceptos clasificatorios, sobre todo en el desarrollo de estrategias que faciliten y señalen las posibles opciones y limitaciones al traductor:

Text typology involving genre analysis can help the translator develop strategies that facilitate his/her work and provide awareness of various options as well as constraints (Trosborg 1997: viii).

Hoy en día existe una clara separación entre estos dos conceptos en Traductología contemporánea (Trosborg 1997), aunque se sigue debatiendo la importancia de los mismos tanto en el proceso de traducción como en la formación de los traductores (Mayoral Asensio 2004; Borja Albí, A. 2005; Ezpeleta 2005; Trosborg 1997). No obstante, defendemos que el género y la tipología textual son esenciales tanto en el desarrollo profesional de la traducción como en su didáctica, puesto que un conocimiento inadecuado sobre el funcionamiento de los géneros, tipologías textuales y las diferencias interculturales puede conllevar traducciones inadecuadas (Trosborg 1997: 18).

Apoyamos así la complementación de ambas vías de análisis (ibid), ya que la confusión que ha existido es indicativa de la pertinencia de clases de clasificaciones diferentes y que se complementen (Hurtado Albir 2001). Por consiguiente, asumimos que tanto el «género» como el «tipo textual» suponen niveles de análisis y clasificación distintos y complementarios. En virtud de esta visión, es inevitable que adoptemos un enfoque multidisciplinar en el estudio de las tipologías textuales

que englobe los aspectos cognitivos, lingüísticos y sociales que repercuten en los procesos de creación y recepción textual (López Rodríguez 2001).

Pasamos a continuación a repasar la evolución y posibilidades de aplicación de ambos conceptos para justificar la metodología de la presente tesis.

## 4.2. Tipologías textuales

A mediados de los sesenta empezaron a aparecer desde la Lingüística y Análisis del Discurso ciertas propuestas de taxonomías textuales que influyeron en el surgimiento de la Lingüística textual y la introducción de estas clasificaciones textuales en Traductología. En un principio, estas tipologías textuales intentan abarcar todos los textos posibles, cubriendo todos los campos, como el técnico, jurídico, literario etc., así como los modos escritos y orales. A menudo, la crítica más extendida a las diferentes propuestas es su afán generalista, junto con la falta de flexibilidad al proponer categorías rígidas contrarias a la multifuncionalidad en la mayoría de los textos (Rabadán 1991). Hatim y Mason (1997) señalan este objetivo iluso de clasificar todo texto en compartimentos estancos:

[...] given this inevitable hybridization, no categories, no matter how rigorously worked-out, can be expected to be definitive (Hatim y Mason 1997: 129).

Los autores presentan en su obra una de sus mayores aportaciones al estudio de las categorizaciones textuales, la **multifuncionalidad** e **hibridación** de los tipos textuales<sup>127</sup>, ya que este aspecto es esencial en la validación de cualquier propuesta presentada:

The problem is that, however the typology is set up, any real text will display features of more than one type. This multifunctionality is the rule rather than the exception, and any useful typology will have to be able to accommodate such diversity (Hatim y Mason 1990: 138).

---

<sup>127</sup> Este aspecto ya fue señalado por Reiss (1971, 1981).

Por otro lado, una concepción flexible de las tipologías textuales es ineludible. Éste se presenta como un problema básico de varias taxonomías textuales caracterizadas por su rigidez clasificatoria (Arntz 1988). En su repaso de las tipologías textuales existentes, Rosa Rabadán (1991) también critica su rigidez y apunta que las tipologías de compartimentos estancos son: «una trampa metodológica que no responde a la realidad» (Rabadán 1991: 181). Esta corriente hacia clasificaciones rigurosas y estancas se cristaliza en la investigación de Isenberg (1983), de gran influencia en estudios posteriores<sup>128</sup>, en el que se detallan los requisitos fundamentales de cualquier tipología textual: homogeneidad, monotipicidad, rigor y exhaustividad. No obstante, como indica Hurtado Albir: «esas condiciones son quizás demasiado exigentes, tanto que [...], ninguna de las diversas tipologías propuestas las ha cubierto»<sup>129</sup> (Hurtado Albir 2001: 459). En los textos web, esta flexibilidad se convierte en un requisito esencial puesto que la mezcla e hibridación de tipos textuales en una misma página o sitio web es una de las características intrínsecas de estos textos (Santini 2005b). Las propuestas de clasificaciones más recientes abundan en la imposibilidad de establecer tipologías exhaustivas, ya que los tipos textuales y géneros evolucionan con las sociedades en los que se encuadran. De hecho, el universo textual que constituye la WWW es un ejemplo básico de esta dinamicidad y evolución en el siglo XXI.

La evolución de las tipologías textuales propuestas en Lingüística, Análisis del Discurso y Traductología, cuenta con un gran número de revisiones (Isenberg 1983; Bernárdez 1982; Bustos 1996; Fernández Antolín 2000; Hurtado Albir 2001; López Rodríguez 1999, 2001; Mayor Serrano 2002; Sánchez Trigo 2002), por lo que nuestro objetivo no será un repaso exhaustivo de las diferentes propuestas, sino más bien revisar aquellas las de mayor impacto en Traductología, para analizar y justificar nuestro marco teórico y metodológico.

---

<sup>128</sup> Como por ejemplo la investigación de Göpferich (1995).

<sup>129</sup> Hurtado Albir se apoya en esta cita en el estudio de Castellà (1992: 228).

### 4.2.1. Tipos textuales en Traductología

Las propuestas de tipos textuales en Traductología no han sido tan numerosas como las propuestas en Lingüística o el Análisis del Discurso, ya que durante los años sesenta y setenta debates como la equivalencia o la unidad de traducción centraron en mayor medida la atención de esta incipiente disciplina. Estas propuestas se vieron inevitablemente influenciadas por las primeras tipologías textuales en Lingüística (Gülich y Raible 1972; Sandig 1975; Petöfi 1979) o Análisis del Discurso, y se observa la influencia de diferentes fundamentos lingüísticos en la multiplicidad de perspectivas: sistémico-funcionales (Halliday 1973), análisis del discurso (Van Dijk 1977), los actos de habla (Searle 1969; Austin 1975), funcionalismo (Bhüler 1934; Jakobson 1960)<sup>130</sup>, o estilística (Crystal y Davy 1969).

El enfoque de clasificaciones por tipos textuales se centra en la investigación sobre la identificación, clasificación y caracterización de textos especializados en dos o más lenguas, y se basa en la premisa de que cuando se redacta o lee un texto, el emisor y receptor toman como base, consciente o inconscientemente, un modelo mental acumulado y compartido con otros miembros de su comunidad de discurso. Estos modelos no se refieren exclusivamente a la superestructura de Van Dijk (1980), que se suele asociar con el concepto de género y que repasaremos en el siguiente capítulo, sino que son mucho más amplios al referirse a otras características lingüísticas que son también prototípicas, como pueden ser el tono y el tenor, los actos de habla y como se ponen de manifiesto en cada tipo de texto los elementos metacomunicativos, etc. El estudio de la tipología textual es de gran interés en la didáctica de la traducción (López Rodríguez 2001: 9), en la labor

---

<sup>130</sup> López Rodríguez (2001: 52), señala la importancia de Jakobson en la formación del formalismo ruso, por lo que aboga por no categorizarlo como «funcionalista».

diaria de los traductores<sup>131</sup> y en la teoría de la traducción. Las tipologías textuales son propias de cada género textual y pueden presentar diferencias interlingüísticas e interculturales. Todo traductor necesita reconocer todos los rasgos prototípicos en las lenguas con las que esté trabajando para comprender mejor el texto y a su vez, producir un texto meta que se ajuste a las expectativas discursivas, tipológicas y de género del receptor. Estas expectativas se cristalizan en unas convenciones que éste espera encontrar de manera inconsciente (véase 5.5.1)

A grandes rasgos, las propuestas de tipologías textuales en Traductología se pueden dividir entre aquellas que parten del ámbito temático y profesional (Kade 1968; Wilss 1977; Delisle 1980; Snell-Hornby 1995), el enfoque mayoritario previo a mediados de los setenta (Hurtado Albir 2001), y aquellas que surgieron desde una óptica funcionalista. Estas últimas son mayoritarias a partir de los setenta influenciadas sobre todo por las aportaciones de Katrina Reiss (1968, 1971, 1976, 1977; Reiss y Vermeer 1984). A su vez, las propuestas de corte funcionalista se pueden dividir entre:

- Tipologías funcionales, por estar basadas en la función del texto, (Reiss 1971, 1976, 1981; Reiss y Vermeer 1984; Nord 1997), que parten de la función comunicativa y dividen los textos entre informativos, expresivos, operativos, multimedia y fáticos<sup>132</sup> (véase, 4.2.1.1 y 4.2.1.4).
- Tipologías basadas en la finalidad retórica o *rhetorical purpose* (Werlich 1975; Baker 1992; Hatim y Mason 1990 1997; Trosborg 1997; Downing 1996), que distinguen entre textos descriptivos, narrativos, expositivos, argumentativos y exhortativos (véase 4.2.1.3).

---

<sup>131</sup> Baker (1992: 222) señala su importancia en el proceso de traducción ya que el traductor necesita conocer la organización y convenciones que surgen de cada tipo textual. La autora no obstante no define qué entiende por tipología textual.

<sup>132</sup> Este último propuesto por Nord (1997b).

Estas dos propuestas no tienen por qué ser excluyentes, siendo válida e de interés una tipología a dos niveles. La propuesta de Trosborg (1997: 16-17) indica que las tipologías pueden considerar a la vez la función comunicativa y el propósito retórico.

Con el fin de establecer y justificar nuestro marco teórico, revisamos a continuación las propuestas que en nuestra opinión revisten una mayor importancia en la Traductología contemporánea: Reiss y Vermeer (véase 4.2.1.1), House (véase 4.2.1.2) Hatim y Mason (véase 4.2.1.3) y Nord (véase 4.2.1.4), si bien no podemos olvidar la existencia de otras propuestas de impacto en esta disciplina (Delisle 1980; Snell-Hornby 1995; Gömmlich 1993; Sager 1997; Kussmaul 1997; Roberts 1996; etc.).

#### **4.2.1.1. La propuesta funcionalista de Katrina Reiss (1971, 1976, 1981) y Reiss y Vermeer (1984)**

Indudablemente, Katrina Reiss fue la primera gran investigadora sobre tipologías textuales funcionalistas. Esta corriente comenzó en 1971 con su libro *Möglichkeiten und Grenzen der Übersetzungskritik*, considerado como el punto de partida del análisis investigador y traductológico en Alemania (Nord 1997b). Los presupuestos de Reiss y su terminología se fueron desarrollando y remodelando a lo largo de sus varios trabajos, y culminaron en su trabajo junto a Vermeer en la que se presentó la trascendental teoría del Escopo, o *Skopostheorie*, (Reiss y Vermeer 1984), una aportación que marcó la investigación y la didáctica en Traductología posteriores. Su primer estudio se dirigió hacia la crítica y evaluación de las traducciones y partía de la controvertida equivalencia traductora<sup>133</sup>, si bien la autora ya establecía excepciones a la misma en 1971 según

---

<sup>133</sup> Tanto en lingüística como en Traductología en los años sesenta y setenta las investigaciones sobre los tipos textuales se inició como complemento a los estudios que abogaban por el establecimiento de la equivalencia entre pares de textos.

el «encargo de la traducción» o cuando el propósito o función no fuera similar entre el TO y el TM (Nord 1997a: 9).

Su clasificación textual con fines evaluativos se apoyaba en las tres dimensiones del lenguaje que formuló Bühler (1965)<sup>134</sup>: la representación, la expresión y la apelación. Desarrolló así un análisis dual mediante la inclusión de un factor intratextual, la dimensión del lenguaje, que relacionó con un factor extratextual de orden pragmático, diferenciando en un primer momento entre tres tipos textuales que se relacionan con las dimensiones del lenguaje: el tipo informativo (en el que predomina el contenido), el tipo expresivo (predominio de la forma) y el tipo operativo (en el que predomina la apelación)<sup>135</sup>. A estos tipos de textos añadió otra categoría con posterioridad, los textos audio-media (Reiss 1977)<sup>136</sup>, textos que se desarrollan en el medio escrito pero representan en otro, como el caso de una ópera que se escribe para ser cantada. Consciente de la insuficiencia de esta primera definición, en 1981 define este tipo de textos como multimodal o multimedial, para incluir formas mixtas de textos que necesitan de varios medios para su transmisión, como pueden ser los cómics que dependen de la interrelación con la imagen para la transmisión. Con la inclusión del medio en la caracterización de su tipología esta se enriquece en sobremanera (López Rodríguez 2001), y aunque no deliberadamente, la proyecta hacia un futuro en el que los textos multimedia como los sitios web suponen una nueva realidad textual y traductológica. No obstante, esta categoría no constituye un compartimiento aparte puesto que estos textos multimedia son a su vez textos informativos, expresivos u operativos. El siguiente cuadro presenta la clasificación textual de la autora en el que hemos incluido géneros digitales:

---

<sup>134</sup> Newmark (1988) también emplea las funciones del lenguaje de Bühler (1934) en su clasificación.

<sup>135</sup> Esta terminología que usó posteriormente la introdujo la autora en 1976, en su trabajo de 1971 denominó estos tres tipos de textos como textos «contenido», «forma» o «apelativos» (Reiss 1971: 33).

<sup>136</sup> Reiss introduce muy someramente esta categoría ya en 1971 sin profundizar en ella (Reiss 1977: 49).

TIPO DE TEXTO	Informativo (predomina el contenido)	Expresivo (predomina la forma)	Operativo (Predomina la apelación)
<i>Función del lenguaje (Bühler, 1934)</i>	Representación	Expresión	Apelación
<i>Dimensión del lenguaje</i>	Lógica	Estética	Dialógica
CLASES DE TEXTOS (géneros)	Noticias de prensa, instrucciones de uso, patentes, documentos oficiales, textos técnicos <b>En sitios web:</b> descripciones de productos.	Biografías, novela, cuento, etc. <b>En sitios web:</b> blogs, sitios web literarios, sitios web de humor, logos de empresas.	Anuncios, discurso electoral, <b>En sitios web:</b> Páginas de búsqueda, anuncios en páginas, formularios de contacto, Términos legales,

Figura 17. Tipología textual de Reiss (1971, 1976, 1977). Adaptado de Hurtado Albir, (2001: 475).

En 1976 y en 1984 (Reiss y Vermeer 1984), Reiss abandona estas funciones del lenguaje como criterio taxonómico y se centra en la función del texto para tal fin. A su vez, a cada tipo textual le corresponden «clases de textos» o *textsorten*, que como mencionamos se corresponden con los géneros en la actualidad.

Las críticas a los primeros modelos tipológicos de Reiss se centraron, por un lado, en la rigidez de una clasificación que no acoge el carácter híbrido de los textos (Snell-Hornby 1995: 30), posición esta que cambió en estudios posteriores, y por otro, en la determinación de la estrategia o método de traducción dependiendo del tipo textual. Este último aspecto, la relación entre tipo textual y estrategia fue muy controvertido y criticado (Hurtado Albir 1996)<sup>137</sup>, a pesar de que lo suavizó en su trabajo con Vermeer, al señalar que solamente existirá una relación entre la estrategia a seguir y el tipo textual cuando los TM en su cultura comporten una función semejante a la del TO. Indica que en el resto de casos es necesario tener en cuenta factores extralingüísticos (1984: 202 y ss). En lo que respecta a la monofuncionalidad de los textos, Reiss abandonó en 1981 esta concepción con la

<sup>137</sup> Hurtado Albir indicaba que la relación existente se basa en las diferentes soluciones que proponen a los problemas de traducción dependiendo del tipo textual y no el método global (Hurtado Albir 1996).



incorporación de factores comunicativos en el análisis del texto, aunque el resultado aún supone un modelo demasiado rígido (Snell-Hornby 1995: 30).

La tipología textual de la autora se terminó de definir en su colaboración con Vermeer en la que adoptó el concepto de *Skopos* o «escopo», la función predominante de un texto determinado (Reiss y Vermeer 1984). En esta obra sostienen que la traducción es una clase de interacción humana y, por lo tanto, se encuentra determinada por el función de la misma, el *skopos*. Los autores señalan que la función del texto está determinada por el destinatario del texto, tanto en el TO como en el TM, y los textos están inmersos y conformados por las convenciones culturales tanto del TO como del TM. Reiss redefine la correlación entre el tipo de texto y el método de traducción para encajarlo en la teoría general de la traducción de Vermeer (1978). La autora, que partió del paradigma de la equivalencia en 1971, señala que puesto que la equivalencia funcional ya no es el objetivo primordial de la traducción de cada tipo de texto, éste no puede ser el criterio básico en la elección del método de traducción (Nord 1997: 10). No obstante, los autores tras la *Skopostheorie* señalan la posibilidad exista una relación de analogía o equivalencia entre el TO y TM, sobre todo cuando el escopo sea similar entre dos textos en sus contextos respectivos, un método que los autores denominan «traducciones imitativas» (Reiss y Vermeer 1984: 89).

De entre las aportaciones de mayor importancia para los trabajos posteriores en tipología textual, señalamos el interés por desvincular las propuestas de tipos textuales procedentes de la Lingüística de propuestas en el seno de la Traductología, cuyos objetivos no se han de corresponder necesariamente con los de la primera:

Mientras la Lingüística del texto centra su interés en establecer una tipología textual basada en el estudio de las particularidades y distintas modalidades o posibilidades descriptivas de los tipos de texto [...] a la Traductología, cuyo objeto de estudio es, entre otras cosas, el análisis de

las condiciones y las posibilidades de la traducción, le interesa [...] una clasificación según las categorías textuales (Reiss y Vermeer 1984/1996: 177).

Los autores por lo tanto no pretenden proponer «categorías textuales» que tengan una validez global en Lingüística general (Hurtado Albir 2001), sino que su propuesta es de interés científico exclusivamente para la Traducción, centrada en los textos que van a ser traducidos o se han traducido. Las propuestas y análisis contrastivo de la presente tesis surge con un enfoque traductológico desde el punto de vista del proceso de traducción, con lo que el mismo aspira a ser válido principalmente en esta disciplina.

A nuestro parecer, se observa un paralelismo entre estos principios funcionalistas con los estudios sobre la localización, ya que en este sector se prima la «abstracción» de la función de cada elemento textual en la fase de internacionalización, con el objetivo de que ésta pueda ser recreada en cualquier lengua y cultura (LISA 2004: 14). El mantenimiento de la función textual constituye por tanto uno de los pilares básicos en la localización de software, generalmente no presenta excepciones (Nord 1997a), y se refleja en gran medida en la localización web de elementos textuales basados en la interacción, tales como menús de navegación o cadenas textuales en *scripts*.

#### **4.2.1.2. La propuesta de Juliane House (1977)**

La propuesta de House (1977: 33 y ss.) se dirige hacia la evaluación y análisis de traducciones y se basa en las dimensiones situacionales propuestas por Crystal y Davy (1967: 66). La investigadora aplica estas dimensiones situacionales a la traducción y señala que los tipos textuales se encuentran limitados por éstas, que clasifica entre las del usuario (variables geográficas, sociales y diacrónicas, etc.), y las dimensiones de uso lingüístico (medio, relación emisor-receptor, etc.). Con la traducción en mente, y puesto que su tipología se puede denominar de

traducciones, añade otra dimensión más denominada la «actitud social», el grado de confianza entre los participantes. Estas dimensiones sirven a su vez para analizar el material lingüístico, y mediante la conjunción de este análisis lingüístico y situacional, se obtiene un «perfil textual» del TO (House 1981: 49). Este «perfil textual» se analiza con el fin de identificar en el texto origen aspectos que si se transfirieran directamente al TM podrían provocar problemas de comprensión en sus receptores. Posteriormente, el análisis del «perfil textual» del TM se desarrolla a continuación para compararlo con el del TO.

La propuesta de la autora sintetiza las funciones del lenguaje desarrolladas por Bühler, Jakobson y Halliday, entre otros, y las relaciona con las tipologías textuales. Distingue así los textos en dos macrofunciones siguiendo a Halliday, la ideacional y la interpersonal (House 1977: 185), que a su vez determinan la función de cada texto, definida como el uso que posee el texto en un contexto determinado o en una situación única. En los textos ideacionales predomina el contenido, como pueden ser los artículos científicos, los artículos de prensa etc., con una subdivisión entre los textos técnicos y no técnicos. En los textos interpersonales existe una relación entre el emisor y el receptor, pudiendo a su vez ser ficción y de no ficción. Los textos ideacionales, en los que predomina el contenido, no son específicos de la cultura origen y, consecuentemente, no disfrutan de un estatus independiente en la traducción, y los interpersonales pueden ser específicos de una cultura de origen. López Rodríguez (2001: 22) señala como crítica a esta división que estas dos macrofunciones pueden estar presentes en todos los tipos textuales, y éste es el caso de cualquier texto web, debido a la combinación de diferentes elementos textuales en un mismo sitio o página web. En una crítica a propuestas anteriores, la autora señala los textos no podrían clasificarse según funciones textuales fijas debido a su posible multifuncionalidad y que, por tanto, sería necesario aplicar varios métodos para una misma traducción (House 1981: 38 y ss).

La investigadora señala que según la función pragmática del texto origen y su nivel de dependencia de la cultura correspondiente, dependiendo del «perfil textual», el traductor ha de decidir entre el tipo de traducción a realizar: una traducción *overt* o *covert*, que en español se ha denominado traducción «patente» o «encubierta» (Rabadán 1991: 299)<sup>138</sup>. La traducción *overt* o «patente» es aquella que no obtiene el estatus de TO en la cultura meta, ya que el texto está específicamente ligado a la cultura y lengua de origen, «the ST is tied in a specific way to the source language community and culture» (House 1981: 189). La función del texto de partida ha de respetarse en esta traducción, puesto que el texto no puede desvincularse de la cultura de origen, siendo preciso un segundo nivel funcional para que se mantenga la función en la traducción. Así, la traducción es posible a pesar de estar ligada a la cultura origen, puesto que el TO puede potencialmente ser de interés general (House 1981: 189). Por otro lado, la traducción *covert* o «encubierta» goza en la cultura meta del mismo estatus que el original y, por consiguiente, adquiere condición de TO en el polisistema meta. En palabras de House, son traducciones en las que el TO «is not specifically addressed to a target culture audience, i.e. not particularly tied to the source language community and culture» (1981: 194). Las traducciones encubiertas presentan así un efecto de equivalencia pragmática entre los textos originales y meta. No obstante, la investigadora indica que para conseguir esta equivalencia se hace indispensable un segundo nivel de revisión, el denominado «filtro cultural» (1981: 196), que se desarrolla desde los niveles intra y extra textuales.

De entre sus mayores aportaciones, en varios estudios se ha señalado el reconocimiento de las relaciones emisor-receptor en su marco teórico y la incorporación de parámetros situacionales en el análisis de los textos origen y meta (López Rodríguez 2001: 33). Por cuanto a la posible aplicación a la localización,

---

<sup>138</sup> Estos mismos tipos se denominan *communicative translation* y *semantic translation* por Newmark (1981, 1988) y *dependent translation* y *derived translation* en la tipología de Sager (1997).

éste proceso supondría mayoritariamente un ejemplo de traducción *covert* o encubierta, puesto que los textos no necesitan ser recibidos como traducciones sino como producciones originales en la cultura meta (LISA 2003: 11). Como indicamos anteriormente, Anthony Pym (2004) apunta una gradual desaparición de la importancia del TO en la localización de software:

Software is anonymous technology speaking to anonymous markets. The text is not called to represent any previous text (Pym 2004a: 8).

Consecuentemente, en este tipo de localización primaría la traducción «encubierta» dada la menor relevancia del TO y la premisa de que estos textos se pretenden producir bajo una «neutralidad» cultural<sup>139</sup>. Los textos web en general<sup>140</sup>, a pesar de una mayor hibridación, suelen ser casos de traducciones «encubiertas», ya que los mismos se suelen establecer como textos originales en la cultura meta. No obstante, en gran cantidad de géneros web existen subtextos como pueden ser contratos legales, referencias literarias etc. que supondrían un tipo de traducción «patente». Esta hibridación textual a la que se enfrenta el traductor-localizador en el proceso de traducción web dificultaría la aplicación de esta tipología a un estudio contrastivo de textos web.

#### **4.2.1.3. La propuesta de Hatim y Mason (1990, 1997)**

Hatim y Mason (1990, 1997) presentaron una propuesta enfocada hacia la traducción desde una perspectiva pragmática y semiótica en la que se resalta la importancia del contexto sociocultural. Los autores señalan como punto de partida las deficiencias de las tipologías anteriores, ya que éstas suelen atender a un criterio taxonómico único, como el «campo del discurso» o la «función», que no aporta valor predictivo alguno y, consecuentemente, esta falta de concreción conlleva la posibilidad de una ingente cantidad de tipos. Frente a éstas, su

---

<sup>139</sup> Que como señalamos anteriormente, es ilusoria (Schäler 2002).

<sup>140</sup> En los sitios web pueden aparecer textos legales o literarios.

propuesta se centra en una **multifuncionalidad** textual, que según los autores suele ser la norma, «This multifunctionality is the rule rather than the exception» (1990: 138). Para adoptar esta multifuncionalidad, proponen un modelo exhaustivo de contexto que aúne valores pragmáticos, comunicativos y semióticos del texto tanto en su producción como en el intercambio comunicativo (véase 5.5). De esta manera, con el análisis de las tres dimensiones del contexto en la transacción comunicativa, aseguran que ninguna dimensión prevalezca o sea más importante que las otras, como es el caso de la dimensión pragmática en propuestas funcionalistas (Reiss 1970; Nord 1997b). Así para Hatim y Mason, junto con el concepto de intención las tres dimensiones deben estar presentes en el análisis de cómo se produce la interacción comunicativa.

El tipo textual se convierte en el marco conceptual según el cual se posibilita la clasificación textual partiendo de un foco pragmático identificado en cada interacción, y que a su vez está al servicio de un propósito retórico global:

In any particular juncture in an interaction, a pragmatic focus is identified. The focus is likely to subsume a set of mutually relevant intentions and will define the type of text currently evolving. This is the basis of what we will refer to as **text type – a conceptual framework which enables us to classify texts in terms of communicative intentions serving and overall rhetorical purpose**<sup>141</sup> (1990: 140).

Los propósitos retóricos, la intención global del productor de un texto concretado en la función del mismo, se encuentran en el contexto del texto. Siguiendo la propuesta de Werlich (1976: 19), basan su tipología textual en el foco contextual dominante, que los autores denominan, **foco tipotextual**: el aspecto del contexto que se considera función primaria de un texto y que determina el tipo textual (Hurtado Albir 2001: 481). El foco tipotextual sería por tanto la conjunción del foco contextual y el propósito retórico por el que se reconoce un texto como un ejemplar de un tipo textual determinado. Este foco tipotextual dominante, a pesar

---

<sup>141</sup> Énfasis propio.

de la multifuncionalidad de los textos, es único ya que cada texto puede servirse exclusivamente de un propósito retórico en un momento determinado.

Although we recognize multifunctionality as an important property of texts, we submit that only one predominant rhetorical purpose can be served at one time in a give text [...] Other purposes may well be present, but they are in fact subsidiary to the overall function of the text (Hatim y Mason 1990: 146).

Los autores señalan la presencia de posibles focos subsidiarios reconocibles en un texto concreto junto con el foco tipotextual dominante, y esta presencia conjunta se corresponde con el concepto de hibridación textual. Ésta, no obstante, se describe como un «continuo», y se señala la posibilidad de casos en los que sea difícil identificar el foco tipotextual dominante y subsidiario<sup>142</sup>.

Los autores, siguiendo de nuevo a Werlich (1976), distinguen entre tres tipos textuales básicos, el tipo textual **argumentativo**, el **expositivo** y el **exhortativo** (Hatim y Mason 1990: 153-158), que describimos a continuación:

- a) El tipo textual **argumentativo** se circunscribe a la evaluación de las relaciones entre conceptos. A su vez, Hatim y Mason distinguen dos tipos de argumentación (1997: 127): la **argumentación íntegra**, en la que se presenta una tesis inicial que se argumenta hasta el final, y la **contraargumentación**, en la que se presenta una tesis inicial de un oponente que se contrargumenta, para finalizar con la presentación de la tesis que defiende el emisor. A su vez, la contraargumentación la dividen en el equilibrio y el argumento «lopsided» o sesgado. En la primera se presenta el cambio entre la tesis y la contrargumentación de manera implícita o explícita sin conjunción alguna, mientras que en el argumento «lopsided» se anticipa la contraargumentación con una concesiva explícita (aunque, a pesar de, no obstante, etc.). Este tipo

---

<sup>142</sup> Se presenta un caso extremo de hibridación llamado «hibridación intertextual» en el que un texto se cambia a otro tipo para servir otro propósito retórico sin perder todas las propiedades del tipo original (1990: 147).

sería el único que no es propio del sitio web de estudio, aunque es dominante en otros géneros web como los foros de opinión.

b) El tipo textual **expositivo** surge del análisis de conceptos, o a partir de la síntesis de ciertos elementos que lo constituyen. Los autores proponen tres variantes que se solapan con tres de los focos contextuales propuestos por Werlich:

- La **exposición conceptual**: se refiere a conceptos
- La **descripción**: se centra en objetos o situaciones
- La **narración**: se centra en acciones o acontecimientos.

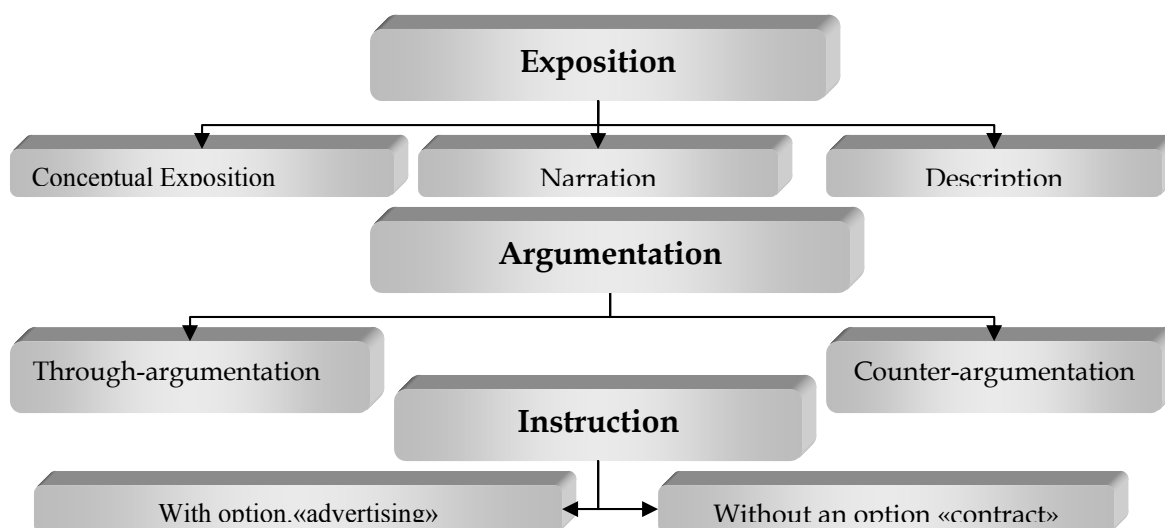


Figura 18. Tipología textual propuesta por Hatim y Mason (1990: 158).

Los autores indican la facilidad de reconocer la descripción y la narración, siendo más sutil la distinción entre la exposición conceptual y la argumentación íntegra. Para su reconocimiento, recurren a explicar las características de cada uno. Por un lado, la finalidad o foco del emisor en la argumentación es conducir la situación de manera que sea favorable para éste. Por otro, el fin de la exposición conceptual consiste en presentar un «seguimiento» de manera subjetiva de los aspectos de la escena, sin valoraciones del mismo.



c) Finalmente, los textos **exhortativos** o instructivos se basan en la formación de conductas futuras. El foco sería por tanto un intento de dictar por medio de la exhortación la manera de pensar o actuar del receptor. Los autores diferencian entre la **exhortación con alternativa**, como los anuncios web, los consejos al cliente, etc., y la **exhortación sin alternativa**, como los contratos de uso o los avisos legales en los sitios web. En el tratamiento de este foco tipotextual, los autores señalan la similitud entre la exhortación con alternativa y la argumentación, y señalan que Reiss (1976) los engloba en los textos operativos. Indican como características diferenciadoras la mayor organización lógica en el formato de la argumentación, frente a un menor grado de organización formal en la exhortación con alternativa. Señalan además que en caso de que se use la contraargumentación en formato publicitario, los receptores no reaccionan sopesando la validez de los argumentos sino que evalúan el texto según su atractivo global (1990: 158).

Por otro lado, Hatim y Mason relacionan los tipos textuales tanto con aspectos ideológicos (1990: 160-164), como con la existencia de factores cognitivos y psicológicos:

Our ability to recognize texts as instances of a type – exposition, argumentation, instruction – depends on our experience of previous instances of the same type, in other words, on our ability to recognize texts as signs (ibid: 2).

Estos factores determinan la habilidad de los hablantes para identificar ciertos patrones interiorizados de organización textual, así como la capacidad de anticipar cómo se va a desarrollar un texto. Este aspecto entronca con la adecuación psicológica de los tipos textuales presentada por Werlich (1976: 21), que incluyó la vertiente cognitiva en su tipología al asociar ciertas propiedades cognitivas

básicas<sup>143</sup> con los tipos textuales que propuso<sup>144</sup>. Igualmente, relacionan el procesamiento de las estructuras en los tipos textuales con los conceptos de *marcos, guiones o schemata* en términos de Beaugrande y Dressler (1981).

Por último, relacionan la competencia del traductor por cuanto a la sensibilidad hacia los focos tipotextuales y su importancia en el proceso traductológico (1990: 164).

#### **4.2.1.3.1. El campo, el modo y el tenor**

Dentro de la publicación de los investigadores, para nuestra tesis es de interés el tratamiento que realizan del campo, el modo y el tenor, que sitúan dentro de la dimensión comunicativa en su propuesta de contexto<sup>145</sup>. Partiendo de Halliday (1985), definen estas dimensiones como elementos definitorios del registro, que identifican con los usos de la lengua. A su vez, éste último queda definido por sus situaciones de uso a través de la interacción de estos tres factores:

...identifying the register membership of a text is an essential part of discourse processing; it involves the reader in a reconstruction of context through an analysis of what has taken place (field), who has participated (tenor), and what medium has been selected for relaying the message (mode) (Hatim y Mason 1990: 55).

El lenguaje de la web se caracteriza por la mezcla de campos temáticos, puesto que en un sitio web corporativo podemos encontrar avisos legales, descripciones técnicas de productos, noticias o publicidad de sus propios productos, es decir, los textos web se caracterizan por la presencia de varios campos específicos que se pueden mezclar en una misma página web.

---

<sup>143</sup> Percepción del espacio y el tiempo, la análisis y la síntesis, la evaluación de relaciones por medio de la extracción de similitudes, contrastes etc. y la planificación de conductas futuras (Werlich 1975, 1976).

<sup>144</sup> Descripción, narración, exposición, argumentación, instrucción (Werlich 1975, 1976).

<sup>145</sup> Los investigadores identifican tres dimensiones constitutivas del contexto: la comunicativa, la pragmática y la semiótica.

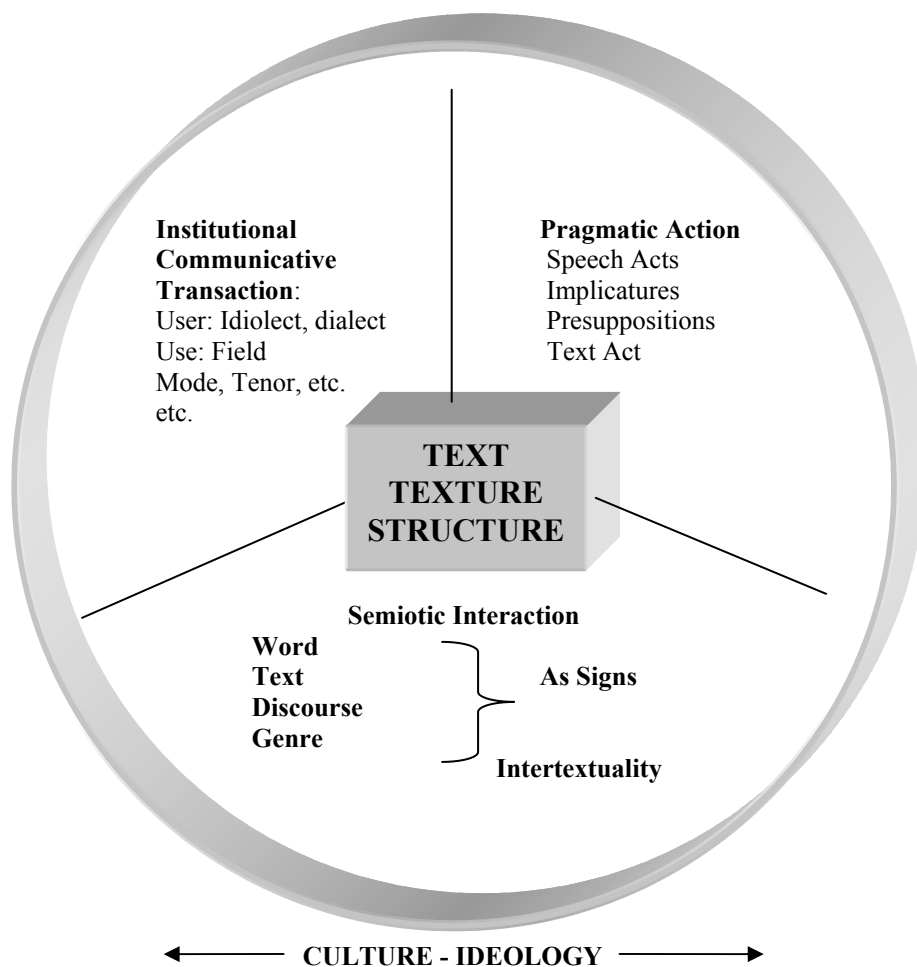


Figura 19. Esquema de las tres dimensiones del contexto (Hatim y Mason 1990: 58)

El modo digital se presenta con unas características generales que lo acercan más al modo hablado que al escrito (Posteguillo 2003: 35), debido sobre todo al uso de Internet como medio de comunicación en el que su naturaleza, la escrita, se combina con el discurso escrito, su representación (Giménez 2000: 239). Varios autores denominan esta tendencia *conceptual orality*, textos en modo escrito pero conceptualizados de manera oral (Janoschka 2003: 101-105), también denominado *informality* dentro de la Netlingüística de Posteguillo (2003). Ciertas características del texto web que lo acercan al oral emanan de estudios cognitivos de lectura en pantalla, que señalan aspectos como brevedad, claridad y la necesidad de ser conciso para que el texto sea efectivo (Brinck *et al.* 2004; Nielsen y Loranger 2006; McGovern *et al.* 2002). Es necesario indicar que el modo oral puede aparecer en un texto web en forma de clips de sonidos, vídeos o presentaciones habladas, pero

suele desaconsejarse por ralentizar la descarga del contenido (Nielsen y Tahir 2002).

De los elementos que componen el registro, Hatim y Mason (1997) identifican al tenor como el elemento más importante. En el análisis del tenor digital es necesario señalar el anonimato general que caracteriza los textos en localización, «software is anonymous technology speaking to anonymous markets» (Pym 2004a: 8). En el caso de los sitios web, la necesidad de identificar un emisor es criterio necesario en ciertos géneros<sup>146</sup>, y en nuestro caso, la empresa se presenta como emisor y se comunicaría con el receptor como tal. El tenor digital se caracteriza por una informalización o coloquialismo (Posteguillo 2003), aunque no obstante, en los sitios web corporativos se caracteriza por un grado de formalidad mayor que el que se encuentra en el ámbito web anglosajón (Jiménez 2005a; Rennau 2004; Posteguillo 2003: 101-102). Dada la importancia del tenor en la localización, al suponer una decisiones directas del traductor durante el proceso de traducción desde el inglés al español, este aspecto será objeto de un estudio cuantitativo en nuestro análisis extratextual del género (véase 8.1.2.3)

#### **4.2.1.4. La propuesta de Nord (1997a, 1997b)**

La tipología propuesta por Christiane Nord entronca con el funcionalismo y se enmarca dentro de la teoría del *skopos* (Reiss y Vermeer 1984), que la investigadora modifica dentro de su propio modelo. Esta propuesta supone una aplicación explícita de la teoría de la acción a la traducción, que siguiendo a Vermeer define como «an intentional, interpersonal, partly verbal intercultural interaction based on a source text» (Nord 1997b: 16-26).

---

<sup>146</sup> Dependiendo del género, muchos requieren que el autor se identifique, tales como las páginas personales, las páginas de contactos, los blogs, correos electrónicos.

Su propuesta se fundamenta en el rechazo implícito de la equivalencia en la que se basaban ciertos modelos tipológicos anteriores:

Linguistic translation theories hinge on the concept of equivalence, which has been one of the most ambiguous concepts in translation studies from the start (Nord 1997a: 43).

Tras un repaso de las diferentes acepciones de equivalencia en modelos anteriores, señala como principales deficiencias en los mismos (Nord 1997a: 43-45):

- a) La importancia que otorga a la estructura del TO, por lo que se olvida la relación intrínseca entre los aspectos extratextuales e intratextuales de la interacción comunicativa.
- b) La falta de consideración por los aspectos culturales.
- c) La falta de consistencia del modelo de equivalencia.
- d) La confusión que conlleva la existencia de criterios contradictorios para la selección de procedimientos de transferencia según el género o tipo textual.
- e) La aparente «universalidad» de este modelo que, por tanto, no da cabida a diferencias en los conceptos traductológicos específicos de cada cultura.
- f) El modelo excluye al TM que no satisfaga el criterio de equivalencia, por lo que no se le otorga el rango de «traducciones», como por ejemplo las adaptaciones, las traducciones filológicas, etc.
- g) El valor o valores del TO se consideran como el único baremo al que un traductor ha de subordinar su proceso de traducción, con lo que esta profesión se ve abocada a un bajo prestigio social.

La autora, tras exponer las deficiencias de este modelo de equivalencia, defiende como alternativa la combinación de «funcionalidad» más «lealtad» (Nord 1991: 28 y ss), basada en el escopo (Vermeer 1978; Reiss y Vermeer 1984). Según Nord, esta teoría supone un modelo que se ajusta a la realidad de la traducción y engloba las siguientes características necesarias (Nord 1997a: 45-47): es **pragmático**, ya que

señala se orienta hacia el TM y señala al receptor como al objetivo básico de las decisiones de traducción, **orientado hacia la cultura**, puesto que considera que la traducción como un proceso tanto verbal como no verbal que se desarrolla entre culturas, **coherente**, por que constituye un base metodológica y teórica que puede justificar las decisiones del traductor en cualquier forma de acción traductora, **exhaustivo**, que no sirve exclusivamente a un tipo de texto sino que pueda englobar la totalidad de los mismos, **antiuniversalista**, que frente a las concepciones que surgieron de la gramática estructural y el funcionalismo transformacional da cabida a las diferencias interculturales, **práctico**, para que pueda englobar la totalidad de las acciones transculturales o procesos de traducción, y por último, **experto**, para que otorgue al traductor el prestigio por su labor traductora y que posee consecuencias importantes en la enseñanza de la traducción.

En la Traductología contemporánea, la investigadora es reconocida por su modificación al modelo del escopo, que consiste en añadir a la primacía de la «funcionalidad», característica de la escuela funcionalista alemana, un compromiso con la «lealtad», definida como «this responsibility that translators have toward their partners in the translational interaction» (Nord 1997a: 48). En otras palabras, es la responsabilidad que vincula a todo traductor con el autor del texto original, los receptores del TM y el cliente que ha encargado la traducción, y que por tanto le compromete de forma bilateral con ambas partes, sin dejar a un lado las diferencias en cuanto a la concepción de la traducción que predomina en cada cultura (Hurtado Albir 2001: 535). La autora resuelve los problemas que observa en el funcionalismo radical con este modelo combinando «lealtad» y la «funcionalidad», ya que en estos se infravalora la relación existente entre el autor y el traductor (Nord 1997a: 48). Según este modelo, una buena traducción surge del compromiso entre el escopo y la lealtad hacia todos los participantes en el proceso de comunicación (López Rodríguez 2001).

La tipología de Nord surge en consonancia con su modelo de análisis, una perspectiva funcionalista, y parte de la definición de texto de Beaugrande y Dressler, «a text is a communicative occurrence» (1981: 3), un evento comunicativo que se produce en una «situación» determinada. La concepción de su función textual es una cualidad pragmática que asigna el receptor en una situación determinada, que no algo inherente en cada texto. La función del TM es además específica de la situación de recepción, por lo que el TM no ha de presentar necesariamente la misma función que la del TO: «it is an illusion that a target text should have *automatically* the "same" function as the original» (Nord 1997 a: 49).

Su modelo tipológico se desarrolla sobre las funciones del lenguaje de Bühler (1934) y Jakobson (1960), y de ellas selecciona cuatro como marco comparativo para todos los textos: la **referencial**, la **expresiva**, la **apelativa** del modelo de Bühler (1934) y la **fática** del modelo de Jakobson (1960). Identifica a su vez ciertas subfunciones principales que concibe como una lista abierta (Nord 2002: 5):

- a) La función **referencial** se refiere a la descripción de fenómenos u objetos del mundo. Incluye las subfunciones **informativa**, **metalingüística**, **instructiva** y **educativa**.
- b) La función **expresiva** se centra en la expresión de las emociones u opiniones del emisor sobre los objetos o fenómenos en cuestión. Incluye las subfunciones **emotiva** y **evaluativa**.
- c) La función **apelativa** se dirige a inducir al receptor a reaccionar de una manera determinada mediante la apelación a sus sentimientos, experiencias, sensibilidad, etc. Incluye las subfunciones **ilustrativa**, **persuasiva**, **imperativa**, **pedagógica** y **de publicidad**.
- d) La función **fática** tiene como objetivo establecer, mantener o terminar la acción comunicativa. Incluye las subfunciones *salutacional*, *small talk* o *peg function*.

La tipología presentada parte de la función del texto que, en principio, ha de definirse según el TO. De ahí identifica dos tipos de traducciones (Nord 1998: 51-52) la traducción documental (*documentary translation*) y la traducción instrumental (*Instrumental translation*). La primera es un tipo de traducción en la que el TM hace referencia al TO y, centrándose en ciertos rasgos que se reproducen, lo representa dentro de la cultura origen. El TM es por tanto un metatexto que se recibe como una traducción. Presenta como ejemplos la traducción literal o la traducción «exotizante» (*exoticing translation*), un tipo de traducción que produce un efecto «exótico» en los lectores meta al mantenerse elementos de la cultura origen que no causan este efecto a lectores del TO. En la traducción instrumental (*instrumental translation*), el TM constituye una recreación en la que el TO se usa como modelo y cumple una función textual como si fuera un texto original dentro de las condiciones del TM. En este caso, el lector procesa el texto tal como producción original, con lo que el texto se adapta a las normas y convenciones de género, tipo textual, registro etc. de cultura meta. Este tipo de traducciones se correspondería con el objetivo del sector de la localización para los textos localizados (LISA 2003).

La visión de las traducciones instrumentales es puramente funcional y dentro de éstas, la autora incluye las traducciones:

- a) «Equifuncionales», en las que se mantiene la función del TO en el TM, como en el caso de traducciones técnicas, manuales de ordenadores, o en nuestro caso, los menús de navegación que guían la interactividad en los sitios web;
- a) «Heterofuncionales» que se producen en el caso de que no se mantenga la función o la jerarquía de funciones del TO en el TM.
- b) «Homólogas», en este último caso se pretende crear el mismo efecto en el polisistema o contexto meta del que produce el TO en su contexto o polisistema



original. Esta última función se denomina «transformación semiótica» por Ludskanov o «transposición creativa» según Jakobson (Bassnet 1991: 18).

En principio, apreciamos que la tipología de traducciones de la autora no se relaciona de igual manera con los tipos textuales existentes en cada cultura. En las traducciones documentales se asume que se mantiene la función del TO en el TM y, por consiguiente, el mismo sigue las convenciones tipológicas y de género de la cultura de origen para conseguir traducciones «equifuncionales». Por el contrario, en el caso de las traducciones instrumentales, y puesto que se realiza una adaptación a las normas y convenciones de tipo textual, género etc. (Nord 1997a: 54), la búsqueda y contraste de las funciones en el TO y TM para realizar la traducción supone una adaptación de un tipo textual origen al tipo textual de la cultura meta.

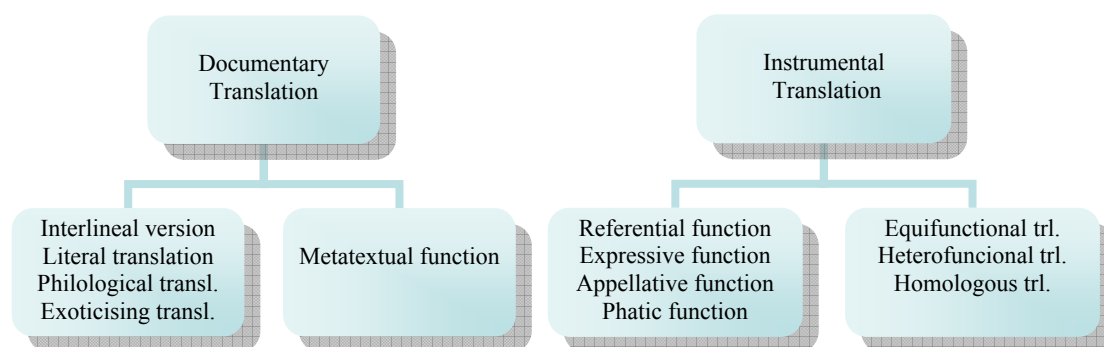


Figura 20. Tipología funcional de traducciones de Nord (1997 a: 52).

El modelo tiene implicaciones didácticas claras e importantes, por lo que va más allá al señalar que puesto que el objetivo del proceso de traducción consiste producir un texto «funcional» en la cultura meta, el proceso de traducción ha de comenzar por el análisis de la situación de recepción. A continuación es necesario conocer qué información ha de incluir el *escopo*, que en el contexto didáctico se

cristaliza en el «encargo de traducción» (Nord 1997b) o *translation brief*<sup>147</sup>. Las implicaciones didácticas de este «encargo» son importantes, puesto que a través del mismo se puede reconocer la tipología de textos y debe incluir: (1) la intención o intenciones del emisor, (2) el receptor o receptores, (3) el tiempo y lugar de la recepción textual, (4) el medio por el que se transmitirá el texto, y (5) el motivo por el que se produce y recibe.

La autora, consciente de las posibles deficiencias de un modelo que se centra casi unilateralmente en el concepto de equivalencia funcional, propone para finalizar una categorización de los posibles problemas que se presentan durante el proceso de traducción. La autora señala que éste debe ser de arriba-abajo, *top-down*, partiendo del encargo de la traducción y definiendo la función de la traducción. Esta jerarquía funcional de los problemas de traducción incluye elementos lingüísticos y semióticos que se descuidan en su tipología funcional. La autora presenta los posibles problemas que se pueden presentar al aplicar su modelo a diferentes tipos textuales, e incluye problemas pragmáticos, interculturales, interlingüísticos y los específicos de cada texto, como las metáforas, símiles, etc.

A modo de conclusión, Nord señala que su modelo funcional se fundamenta en un concepto de textualidad orientado hacia la comunicación y un concepto de traducción orientado hacia el texto meta, si bien basado en el principio de lealtad a los participantes de la comunicación (Nord 1991 a: 64). A nuestro parecer, una de las mayores aportaciones de la autora es la visión y análisis de la función textual que tiene repercusiones en los estudios posteriores. Sin embargo, la necesidad de incluir en su tipología una jerarquía de problemas de traducción sugiere la elaboración de dos tipologías complementarias; una se centra exclusivamente en la

---

<sup>147</sup> La inclusión del encargo de la traducción en la didáctica de la traducción se ha identificado como problemático puesto que los profesionales raramente reciben un encargo completo (Muñoz Martín 2002). No obstante la autora señala que el encargo puede ser o es deducido por el traductor profesional: «the TT function "skopos" is defined by the translation assignment give by the initiator or deduced by the (professional) translator» (Nord 1997 a: 56).

función y la segunda se presenta para la puesta en práctica de su modelo didáctico (López Rodríguez 2001). Por otro lado, y a pesar de la importancia y primacía del mantenimiento de la función de un texto determinado en los procesos de localización, la autora no define qué constituye un texto y cómo delimitarlo. Observamos por consiguiente un fin eminentemente didáctico al asumir que el escopo y encargo de traducción se corresponde a un texto único. Se asume además un tipo de traducción único para cada proceso de traducción, aspecto este ilusorio en un proceso multimodal como la localización.

Como resumen, la aplicación de este modelo a la práctica de la localización sería a nuestro parecer impracticable, puesto que en este entorno dinámico y colaborativo se traducen continuamente diferentes tipos de textos y subtextos. La localización se correspondería con un el tipo de traducción instrumental, aunque si aplicásemos este modelo a la localización web, se producirían cambios constantes entre traducciones equifuncionales, heterofuncionales y homólogas con la consecuente evaluación continua de la función del TO y la negociación hacia el TM.

En cualquier caso, y a pesar de que en la localización no sólo se produce un proceso *top-down*, sino que la negociación continuamente se produce *bottom-up*, la primacía que otorga la autora a un inicio de arriba-abajo argumentamos es esencial en la labor localizadora. En este contexto, el proceso suele incluir textos organizados según criterios informáticos y no comunicativos (véase, 3.1), por lo que un análisis global debe primar en el proceso de traducción en localización.

#### **4.2.2. Tipologías textuales en Internet y Localización**

Pese a las diversas propuestas de «tipos textuales» presentadas en Lingüística y Traductología, este hecho no queda reflejado en los incipientes campos de la Localización o *Netlinguistics* (Posteguillo 2003). Nuestra búsqueda bibliográfica

dirigida a identificar estudios que incorporen o desarrollen este concepto, tanto en Localización como en investigaciones sobre Internet, arrojó escasos resultados. En cualquier caso, la mayoría de investigaciones identificadas parten de un enfoque lingüístico (Biber 2004; Santini 2007a; 2005a, 2005b).

La falta de concreción terminológica y conceptual mencionada anteriormente (véase 4.1) se proyecta igualmente a las investigaciones en esta área. En localización el término «tipo textual» o *text type* se suele emplear simplemente como un parámetro de variación textual, definido según criterios específicos de cada investigador faltos de bases teóricas sólidas. En parte, este hecho se debe a que la dinamicidad propia de todo texto web y su enmarque en un medio interactivo dificultan la aplicación de tipologías textuales clásicas en Lingüística o Traductología.

Tras la revisión sobre tipos textuales, la flexibilidad y multifuncionalidad se presentan como criterios esenciales en cualquier categorización de unos textos web dinámicos erradicados en un medio interactivo. No obstante, este aspecto también aparece como el problema principal a la hora de establecer tipos textuales en la web (Santini 2005a, 2005b), dificultando la atribución de tipologías clásicas a estos textos (Tercedor 2005). Como resultado, aparecen textos web que, si se aplican modelos clásicos, se identificaría como inclasificables (Santini 2007a, 2005a).

#### **4.2.2.1. Tipologías textuales en Internet (Biber 2004; Santini 2005a, 2005b)**

Una de las vías actuales de investigación en tipologías textuales web se enmarca dentro de la perspectiva computacional léxico-semántica encabezada por Biber (1988, 1989, 1995, 2004). El autor presenta una perspectiva multidimensional de tipos textuales que se cristaliza en su modelo de corte funcional y retórico

mediante un análisis de *clusters* estadísticos. Éste se fundamenta en el análisis de elementos léxicos y semánticos que co-ocurren en los textos para, a partir de un análisis estadístico, asignar los textos según una serie de parámetros:

The typology developed here is based on sets of syntactic and lexical features that co- occur frequently in texts, rather than assuming sets of features defined on a priori functional grounds (Biber 1989: 5).

Por consiguiente, la tipología no parte de la dimensión funcional, sino que ésta se aplica una vez que el estudio lingüístico haya revelado ciertas características internas lingüísticas que resulten en una categorización lingüística previa (ibid: 5). No obstante, el investigador relaciona directamente los tipos textuales con el propósito retórico de su emisor:

By text type we refer to the purpose of the text, i.e. the reason for which a text has been written. Text types are related to the producer's intention towards the receivers (Biber 1989: 5).

En este contexto, el investigador aplica su modelo de análisis a los textos web a pesar de que intenta identificar una página web con un tipo textual determinado (Biber 2004), presupuesto que a nuestro parecer desvirtúa su propuesta ya que las páginas web no constituyen más que una unidad de contenido, que no unidad textual (véase 3.4). El autor identifica exclusivamente una página web con el texto extraíble en el cuerpo o área principal de la página web, obviando otros elementos textuales, gráficos, multimedia e interactivos en toda página web que contribuyen de igual manera al propósito comunicativo global. Consecuentemente, se deja a un lado la funcionalidad y el carácter icónico-verbal del texto web<sup>148</sup>. Argumentamos que un análisis exclusivo de los elementos textuales internos codificados en texto no sería suficiente para la clasificación textual web.

---

<sup>148</sup> En nuestro tesis, incluimos en el análisis las descripciones textuales de toda imagen o archivo en formato flash recogidas en las etiquetas <alt>, que representan su aportación al texto en su conjunto (véase 9.2.2).

Como resultado, el autor presenta cuatro «registros» textuales en Internet, concepto general que identifica como término general asociado con los aspectos de la variación en el uso, «the varieties of the language the we use in different situations are referered to as registers» (Biber *et al.* 1998: 2). Estos registros identificados son:

- Narración personal (*Personal, involved narration*)
- Argumentativo, persuasivo (*Persuasive, argumentative*)
- Consejos (*Advice??<sup>149</sup>*)
- Discurso abstracto o técnico (*Abstract-Technical discourse*)

Esta propuesta es preliminar y en la misma se aprecia la indecisión entre las posibles categorías al incluir el tipo «consejos». No obstante, se observa cierto paralelismo con las tipologías textuales basadas en el propósito retórico (véase 4.2.1.3), ya que el tipo expositivo, se correspondería con *involved narration* y «abstract-technical discourse», el argumentativo con el *advice* y el exhortativo con el «persuasivo».

Siguiendo una metodología similar, la investigadora Marina Santini (2007a, 2007b, 2005a, 2005b) recoge esta perspectiva de análisis computacional y presenta las deficiencias que tal análisis entraña, ya que los textos web son estructuras textuales no lineales, con elementos gráficos y una mezcla continua de funciones y propósitos retóricos. La investigadora recoge un corpus representativo de la totalidad de «páginas» existentes en la web, que analiza por medio de un análisis estadístico-computacional con el fin de realizar una identificación automática del tipo textual. En nuestra opinión, el punto de mayor interés para la presente tesis lo

---

<sup>149</sup> Biber hace propone esta categoría como posible y la cuestiona. En nuestro corpus, encontramos un bloque comunicativo que llamamos «Consejos» que aparece en un 9,30% de los sitios web originales en España, con lo que podemos corroborar la propuesta de Biber, aunque la describimos según elementos externos de género y no por cuanto a sus elementos intratextuales o tipología textual.

constituyen sus conclusiones referentes la dificultad de asignar tipologías clásicas a los textos web debido a:

- a) La existencia de gran número de elementos textuales codificados en imágenes que conllevan una dificultad añadida para aplicar técnicas de lingüística de corpus.
- b) La gran cantidad de encabezamientos presentes en los textos web.
- c) Los directorios o listados.
- d) La gran cantidad de texto tabular.

Coincidimos con la autora en la existencia de dificultades adicionales al procesar textos web. De hecho, en el afinamiento de nuestra metodología nos fuimos encontrando ciertos problemas que requirieron su adaptación o modificación continua. Ahora bien, como ya mencionamos, ciertos problemas identificados por la autora se deben al uso de las páginas web como unidad de análisis<sup>150</sup>, puesto que éstas dependen de la contextualización del sitio web global, y el análisis aislado puede producir resultados distorsionados. Finalmente, esta investigación se decanta, al igual que la nuestra, por el estudio de los textos web mediante el marco teórico que proporciona el género, ya que representan categorías más definidas y operativas con una mayor facilidad de aplicación a los textos en el nuevo medio, perspectiva esta con un mayor número de investigaciones en textos web.

#### **4.2.2.2. Tipologías textuales en Localización (Austermühl 2006; Reinke 2005)**

En Localización, la aparición de la «tipología textual» es casi inexistente en las investigaciones hasta la fecha ya que, como indicamos anteriormente, se ha

---

<sup>150</sup> Santini emplea tres corpus diferentes compuestos por páginas web identificadas con un único género por página. Los corpus utilizados son el «7-web-genre-collection» de 1400 páginas, el KI-04 compuesto por 1295 páginas, y el corpus web SPIRIT de 95 millones de páginas, aunque la investigadora sólo utiliza 1000 elegidas al azar entre el 40% compuesto por páginas «inclasificables» (Santini 2007 a: 191-197). Los corpus se pueden consultar en <<http://www.nltg.brighton.ac.uk/home/Marina.Santini/>>.

establecido en parte obviando el cuerpo de conocimiento de la Traductología (véase 2.9.2), y se han centrado mayormente en aspectos técnicos que se supone revierten una mayor complejidad o importancia que los «textuales». Además, la fase de control de calidad se realiza exclusivamente sobre el TM y su funcionalidad, sin necesidad de un análisis contrastivo con el TO. De ahí que el enfoque contrastivo crítico y evaluativo (Reiss 1971, 1976) o didáctico (House 1976; Nord 1997) con los que se ha asociado el estudio de las tipologías textuales en Traductología no quede reflejado en este campo.

El uso de «tipos textuales» se observa en mayor medida en ciertos estudios sobre la didáctica de la Localización, debido a la necesidad clasificar los mismos con objetivos pedagógicos, aunque el término se emplea idiosincrásicamente como cajón de sastre para señalar la variación textual (Byrne 2006; Reinke 2005; O'Hagan y Ashworth 2003; Auster mühl 2006). Al igual que sucede en otras disciplinas, esta confusión conceptual entraña que se evite introducir el concepto traductológico de «tipo textual». En este sentido, en su estudio sobre la selección textual para la docencia de la traducción, Uwe Reinke (2005) señala implícitamente como deja a un lado este concepto, «the notions of 'text type' [...] are rather intuitive and - depending on the [...] purpose, scope and other factors - may differ in granularity» (Reinke 2005: 15). Los únicos tipos textuales a los que se refiere en localización son los textos web y los textos en la interfaz de usuario, división esta intuitiva y de escasa repercusión en la docencia hacia la que se dirige el estudio.

Desde una perspectiva diferente, O'Hagan y Ashworth (2003: 52) presentan una tipología que heredan de Flanagan (1997), si bien sin definir expresamente qué entienden por tipo textual, en la que se dividen los textos multimedia en **comunicativos**, **referenciales** e **interactivos**, asociando cada tipo a un modo de traducción. Los textos **referenciales** los asocian con la localización web y la



traducción automática en la web, los textos **comunicativos** los identifican con la traducción de correos electrónicos y los textos **interactivos** con el nuevo modo de traducción que presentan en su libro, la *Transintepreting*, la traducción e interpretación automática o no de chats o de videoconferencias. Como crítica a esta propuesta, hemos de retornar a la ampliamente criticada asociación entre tipo de texto y un modo o estrategia de traducción únicos que mencionamos previamente (Rabadán 1991; Snell-Hornby 1995; Hurtado Albir 2001). Por otro lado, la relación exclusiva de un tipo textual, como el interactivo, con la traducción de chats, obviando otros textos interactivos como formularios web, interfaces gráficas en software, etc. nos lleva a calificar la propuesta de incompleta e impracticable en estudios contrastivos.

La siguiente propuesta de tipos textuales se propone dentro de la didáctica de la localización por el neozelandés Austermühl (2006). En la misma se expresa la importancia para los estudiantes de localización de los diferentes tipos textuales presentes en la localización, a pesar de que el uso de este concepto no se corresponde claramente con la noción traductológica ni define qué entiende por «tipo textual» en su propuesta. En su estudio es de interés la apelación por un estudio contrastivo de las diferencias en estos tipos textuales, tanto entre diferentes culturas como diferencias intralingüísticas en la realización de los mismos:

Scholars might want to look at and compare text types involved in website and software localization. Using a typology of software text types [...], analyses could focus on the textual characteristics, inter- and intracultural differences or technical constraints of these specific texts, which would be one way of preparing future translators [...] (Austermühl 2006: 79).

La propuesta del autor pretende ser exhaustiva por cuanto a la localización de software, pero no se atiende a unos criterios homogéneos en su clasificación. Así, aunque no quedan explicitados, nuestro análisis demuestra que los criterios de clasificación se basan en el medio, impreso frente a digital, codificación textual,

html frente a Xml, tipos basados en la función textual frente a clasificaciones basadas en la noción de género, mensajes de error frente a acuerdo de licencia, etc.

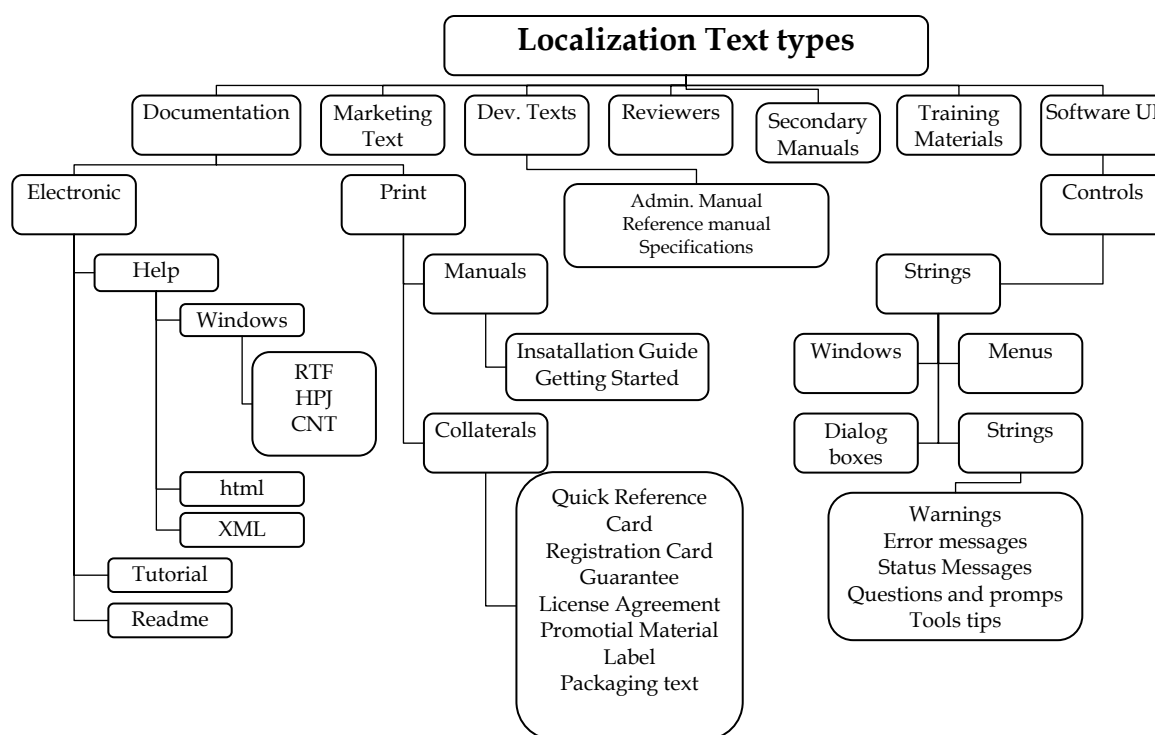


Figura 21. Tipología textual en localización de Austermühl (2006: 80).

La propuesta, por consiguiente, a pesar de ser exhaustiva y de representar un modelo útil en la docencia de la localización, no constituye más que un listado de los posibles subtextos con los que se ha de familiarizar todo estudiante de localización. El modelo de análisis se correspondería mejor con el concepto de «género», ya que los posibles tipos se han obtenido de manera inductiva de acuerdo con criterios externos y no exclusivamente lingüísticos. El género al que se referiría el autor sería el «producto de software»<sup>151</sup>, que a su vez se podría componer de los bloques o secciones textuales que el autor señala en su tipología.

<sup>151</sup> Byrne (2006) también realiza un estudio sobre un «tipo textual» en localización, la «documentación de productos de software», que según su análisis se correspondería con el concepto de género textual.

### 4.2.3. Los tipos textuales en el estudio comparativo de sitios web

Una vez definido y delimitado el concepto de «tipo textual» en Traductología, podemos concluir que constituye una base de análisis demasiado amplia y con un mayor grado de abstracción que los géneros textuales (véase 5). Los criterios para establecimiento de las diferentes tipologías expuestas se fundamentan en cada una en un único criterio distintivo, sea el foco contextual (Werlich 1975,1976), el foco tipotextual (Hatim y Mason 1990, 1997), o la función textual (Reiss 1971, 1976; Reiss y Vermeer 1984; Nord 1991, 1997a, 1997b). Se distinguen de los géneros textuales al definirse éstos a partir de varios parámetros, con lo que estos últimos constituyen categorías más homogéneas (Hurtado Albir 2001). Los tipos textuales suponen categorías de un mayor nivel de abstracción y engloban, por tanto, diferentes géneros textuales.

Los sitios web se caracterizan por la mezcla e hibridación de tipos textuales (Tercedor 2005), lo que dificultaría sobremanera la clasificación basada exclusivamente en una función o propósito retórico predominante; por ejemplo, un tipo textual apelativo o exhortativo se corresponde con los marcos de publicidad web o *banner ads*, con todos los formularios que han de exhortar al usuario a «actuar», en los términos legales de un sitio web etc., con lo que nuestra definición de un texto unitario, el sitio web (véase 3.4), no podría corresponderse con un tipo textual único o incluso con un tipo primario y otro secundario. La categorización textual por tipos textuales se hace por tanto impracticable como criterio taxonómico en nuestro estudio empírico dentro del nuevo modelo textual que representan los sitios web.

Frente al uso de tipos textuales en estudios comparativos, Hurtado Albir (2001) defiende el interés del género en estas investigaciones, sobre todo debido al mayor

nivel de abstracción de los tipos que se corresponden con categorías demasiado vastas:

Los tipos textuales configuran agrupaciones demasiado vastas y heterogéneas; es necesario, pues, contar con agrupaciones más concretas que sean operativas desde el punto de vista de la traducción[...] Los géneros comparten la situación de uso, y las formas convencionales que los caracterizan difieren de lengua a lengua y de cultura a cultura, por lo que es de sumo interés para la Traductología identificar los géneros propios de cada ámbito y conocer sus características para poder contrastarlas con vistas a la traducción y a su enseñanza (Hurtado Albir 2001: 488).

La autora reivindica por tanto la mayor concreción de un concepto de género que se define según criterios múltiples y que, en estudios contrastivos, entrañan consecuencias prácticas tanto en la práctica como en la didáctica de la traducción.

La mayor flexibilidad del concepto de género se debe además a que éstos en principio son ilimitados y en constante evolución dentro su contexto sociocultural (Miller 1984), mientras la tipología textual es un conjunto cerrado con un número limitado de categorías que varían según la propuesta (Trosborg 1997: 15). En el marco de la textualidad web, la flexibilidad y dinamicidad son esenciales, siendo esta naturaleza dinámica más aguda en Internet e informática que en otras áreas (Posteguillo 2003). El marco teórico del género otorga a esta dinamicidad un papel esencial (Swales 1990; Devitt 1993; Miller 1984), por lo que argumentamos que se ajusta mejor a nuestro objetivo, frente a una mayor estaticidad de las propuestas de tipos textuales.

Las dos perspectivas de análisis no son necesariamente excluyentes, aunque un análisis empírico basado en el género no ha de depender necesariamente de tipos textuales de base teórica y abstracta. Como manifestamos anteriormente en el apartado 4.1, la investigadora Anna Trosborg (1997) defiende un análisis a ambos niveles en contexto de la Traductología, sobre todo en el desarrollo de estrategias que faciliten y señalen las posibles opciones y limitaciones al traductor.

Desde nuestra perspectiva de análisis dentro de la textología contrastiva, coincidimos con Fernández Sánchez (2005) que, siguiendo a Adamzik (1998: 15), señala cómo los géneros textuales designan un concepto más aplicado a la práctica, ya que se refiere a una clase de textos que cuentan con una denominación empírica de uso habitual en el lenguaje cotidiano, como el sitio web de una empresa en nuestro caso. A su vez, esta denominación genérica es característica de los textos en los que se centra el análisis empírico de géneros textuales, que el investigador entiende como objetivo primordial de la textología (Fernández Sánchez 2004).

Una vez analizado el concepto de tipo textual y puestas de manifiesto las deficiencias que su uso como criterio taxonómico implicaría en nuestro estudio continuamos, en nuestra búsqueda de un marco teórico de clasificación y análisis que facilite nuestra investigación contrastiva, con una revisión del concepto de género.

## 5. GÉNERO TEXTUAL EN LOCALIZACIÓN

La noción de género, al igual que la de tipo textual, no es unívoca sino multidimensional al haberse estudiado desde una multiplicidad de perspectivas. En el presente capítulo definiremos este concepto mediante una revisión de aquellas propuestas y perspectivas que han articulado su visión actual en Traductología. Nuestro análisis revisará las aportaciones de interés para ayudarnos a seguir delimitando nuestro marco teórico y metodológico, el modelo de caracterización de género que aplicaremos y concluiremos con nuestra propuesta de definición del género objeto de estudio. Asimismo, al final del capítulo recogeremos la perspectiva prescriptiva de las guías de estilo web que, en nuestra opinión, constituyen un compendio de las distintas convenciones asociadas a los géneros web.

La introducción del concepto de género en Traductología no se produjo hasta principios de los ochenta bajo la principal influencia de dos corrientes: la textología contrastiva alemana (Fernández Sánchez 2004), y el Inglés para Fines Específicos (IFE). El objetivo principal de esta última disciplina es clasificar los textos con objetivos pedagógicos (Swales 1990; Bhatia 1993; Nwogu 1991, 1997), y constituye un área de estudio del género dentro de la Lingüística Aplicada. Otros análisis del género parten, dentro de esta rama de la Lingüística, de la perspectiva sistémica funcional y de la comunicación profesional en los Estados Unidos (Paltridge 1997). La Traductología comparte con los estudios de IFE un afán por describir los géneros de especialidad más recurrentes en su situación sociocultural, ya que tanto para la producción, traducción o enseñanza de los géneros textuales más comunes en los ámbitos de especialidad se requiere conocer las convenciones que rigen su funcionamiento.

Las clasificaciones por géneros textuales constituyen categorías más concretas que los tipos textuales, y puesto que los primeros comparten convenciones intra e extratextuales que varían entre lenguas y culturas, se han convertido en objetivo de estudios contrastivos en Traductología (Gamero 2001, 1998; Angost 1996, 1999; Borja 1998, 2000; Jiménez 1999; Bolaños 2002; Vázquez del Árbol 2002, 2006; García Izquierdo 2004; Fernández Sánchez 2004; etc.). Todo traductor debe reconocer convenciones propias del género al que pertenezca el TO y reflejarlas en las convenciones del mismo género en el TM, si bien cuando la traducción lo requiera (Nord 1997a). Este aspecto les confiere una gran importancia en el proceso de traducción.

El concepto de género se encuentra en la actualidad más delimitado tras la confusión conceptual y terminológica inicial (véase 4.1), con gran cantidad de estudios e incluso grupos de investigación, como el GENTT<sup>152</sup> (García Izquierdo y Nebot 2003), exclusivamente dedicados a su investigación, descripción y aplicación tanto a la didáctica de la traducción como a su práctica profesional.

### **5.1. El concepto de género**

El concepto de género posee una larga tradición en estudios literarios<sup>153</sup> y en su origen se observa igualmente la influencia de los trabajos de Aristóteles<sup>154</sup>. Según Paltridge (1997), el primer estudio de género como tal surgió a finales del siglo XIX desde los estudios de folclore (Grimm y Grimm 1881, cf. Paltridge 1997). Desde entonces, son múltiples las disciplinas desde las que se ha investigado tales como la antropología lingüística, la etnografía de la comunicación, el análisis de la conversación, la retórica, la teoría literaria, la sociolingüística o la lingüística aplicada. Esta multiplicidad de perspectivas se refleja en las diferentes

---

<sup>152</sup> GENTT-*Gèneres Textuals per a la Traducció*/ Géneros Textuales para la Traducción.

<sup>153</sup> De hecho el concepto de género se puede afirmar que surge de los estudios literarios.

<sup>154</sup> Los criterios de clasificación textual de Aristóteles dividían los textos en poesía, drama, etc.

concepciones teóricas del género a lo largo de los últimos cuarenta años, sobre todo por los diversos puntos de partida de cada análisis. Paltridge (1997: 4) resume las principales áreas de encuentro o desacuerdo entre las diferentes perspectivas en:

- a) La importancia cardinal de la estructura de un texto en descripción y definición.
- b) La relación existente entre género y su contexto, cultura, cognición, o audiencia.
- c) La importancia de las estructuras lingüística en su descripción.

En consonancia con esta variedad de perspectivas y objetivos de las investigaciones, en la actualidad se llega a cuestionar la necesidad de una teoría del género única, o incluso que puede existir una definición «universal» de esta noción. En este sentido John Swales (1990, 2002, 2004), considerado como uno de las principales influencias que motivó que este concepto se introdujera en Traductología<sup>155</sup>, defiende una mayor flexibilidad el estudio del género y la posible existencia de nociones diferentes en el seno de cada disciplina:

Is there a single one-size-fits-all or “best” genre theory for all circumstances? Or would be better off, in terms of applied linguistics, with a vision of genre that is somehow different from the one that might be most suitable for literary study, for rhetorical theory or for rhetorical history [...]? (Swales 2004: 3).

De igual modo, hemos de analizar si los estudios empíricos contrastivos de género, que como mencionamos han despertado un interés creciente en la Traductología en ámbitos de especialidad, han desarrollado un modelo propio de género adaptado a las necesidades teóricas, profesionales y didácticas de la traducción, para así evaluar su posible aplicación para nuestros fines.

---

<sup>155</sup> Su definición de género que propuso en su libro de 1990 se menciona en casi la totalidad de los estudios empírico de género en Traductología.



### 5.1.1. Definición de género

Como ya avanzamos en el epígrafe anterior, la definición de género es compleja y varía de acuerdo con la disciplina a la que se adscriba. A partir de los años sesenta se comenzó a esbozar el concepto de género desde la Lingüística sistémico funcional (Halliday 1985a; Halliday y Martin 1993), la nueva retórica (Bazerman 1988; Berkenkotter y Huckin 1995), el IFE (Swales 1990; Bhatia 1993; Dudley Evans y St. John 1998), y en menor medida desde la Traductología (Reiss 1971; Reiss y Vermeer 1984; Hatim y Mason 1990, 1997; Göpferich 1995; Gamero, 1998 2001; etc.).

Preliminarmente, se hace patente que los géneros responden a situaciones sociales convencionalizadas que imponen unas restricciones en la organización retórica recogidas en una conjunción de elementos intra e extralingüísticos. Éstos, a su vez, articulan una estructura textual que es reconocible y compartida por los miembros de una comunidad de discurso determinada. En su influyente libro desde el enfoque del IFE, John Swales (1990: 58) resume los aspectos comunes a las propuestas de género previas que asume en la suya propia<sup>156</sup>:

- a) El género existe dentro de una comunidad de hablantes que poseen objetivos comunes, a la que el investigador denomina «comunidad de discurso»<sup>157</sup>.

---

<sup>156</sup> La siguiente sería la definición completa de género del investigador (1990: 58): «A class of communicative events, the members of which share some set of communicative purposes. The purposes are recognised by the expert members of the parent discourse community, and thereby constitute the rationale for the genre. This rationale shapes the schematic structure of the discourse and influences and constraints choice of content and style. Communicative purpose is both a privileged criterion and one that operates to keep the scope of genre as here conceived narrowly focused on comparable rhetorical action. In addition to purpose, exemplars of a genre exhibit various patterns of similarity in terms of structure, style, content and intended audience. If all high probability expectations are realised, the exemplar will be viewed as prototypical by the parent discourse community».

<sup>157</sup> La definición de Swales de comunidad de discurso es restringida en cierta manera (1990: 24-27):« (1) A discourse community has a broadly agreed set of common public goals. (2) A discourse community has mechanisms of intercommunication among its members. (3) A discourse community uses its participatory mechanisms primarily to provide information and feedback. (4) A discourse community utilizes and hence possesses one or more genres in the communicative furtherance of its

b) Remarca, ya desde sus primeros trabajos (Swales 1981), la importancia de la «acción social» y la «intención comunicativa» común para poder definir un género determinado. Para el investigador el género constituye un «evento comunicativo» que comparte «un conjunto de propósitos comunicativos». Bhatia (1993: 13) matizó la propuesta de Swales al primar la función comunicativa, añadiendo que cualquier cambio significativo en el propósito comunicativo, ocasionaría un género diferente (1993: 13).

c) Incide en los aspectos estructurales, que se sitúa entre una de sus mayores aportaciones: «exemplars of a genre exhibit various patterns of similarity in terms of structure, style, content and intended audience» (Swales 1990).

De entre las diferentes corrientes, esta propuesta se encuentra en el grupo que engloba este concepto desde una perspectiva integradora de los elementos intra y extratextuales (Gamero 2001: 50). Este enfoque integrador se ha identificado en Traductología como la de mayor interés (Swales 1990; Bhatia 1993; Brinker 1988; Trosborg 1997 etc.), frente a otras visiones que se centran en el contexto social y comunicativo (Berenkotter y Huckin 1995; Kress 1993; Bazerman 1988; Myers 1990; Miller 1993), o aquellas que se centran en analizar exclusivamente las características internas de los géneros (Tarone *et al.* 1988; Salager-Meyer 1993, 1994, 1998).

Si nos adentramos en las definiciones de este concepto, la primera que recogemos surge de la denominada escuela estadounidense de la «Nueva Retórica» (Miller 1984; Berkenkotter y Huckin 1995), una corriente que lo identifica con la tipificación de la acción social y retórica:

---

aims. (5) In addition to owning genres, a discourse community has a threshold level of members with a suitable degree of relevant content and discourse expertise».

Genres refer to a conventional category of discourse based in large scale typification of rhetorical action; as action, it acquires meaning from situation and from the social context in which that situation arose [...] (Miller 1984: 37).

En esta corriente el género se presenta inmerso en un contexto social y por tanto refleja la sociedad en la que se incluye:

Because it is impossible for us to dwell in the social world without repertoires of typified social responses in recurrent situations; from greetings to thank you to acceptance speeches and full-blown, written expositions of scientific or scholarly investigations. We use genres to package our speech and make of it a recognizable response to the exigencies of the situation (Berkenkotter y Hucking 1995: 7).

Para Berkenkotter y Hucking (1995: 4), la perspectiva en la que se enmarca su concepto de género incluye: el dinamismo, la cognición situada, la forma y contenido, una dualidad situada<sup>158</sup> y una pertenencia a una comunidad.

En este mismo enfoque de reconocimiento de la estructura textual y la función comunicativa se sitúa la llamada Escuela Australiana del género (Martin 1985; Reid 1987). Los investigadores de esta corriente coinciden en el estatus del género como acción social y señalan la existencia de etapas en la estructuración de los géneros. Sin embargo, esta perspectiva parte del marco teórico de la Lingüística sistémico funcional (Halliday 1973, 1978) que le confiere un mayor estatismo a su propuesta. Según Martin (1985) el género se define como:

Genres are how things get done, when language is used to accomplish them. They range from literary to far from literary forms: poems, narratives, expositions, lectures, seminars, recipes, manuals, appointment making, service encounters, news broadcasts and so on. The term genre is used here to embrace each of the linguistically realized activity types which comprise so much of our culture. Its meaning extends far beyond its use in literacy studies to refer to different kinds of verbal art, though each of these does remain a genre in the usage here (Martin 1985: 250).

---

<sup>158</sup> Una dualidad entre la estructura social en la que se sitúan los hablantes y las estructuras textuales que reproducen: «As we draw on genre rules to engage in professional activities, we constitute social structures (in professional, institutional and organizational contexts) and simultaneously reproduce these structures» (Berkenkotter y Hucking 1995a: 4).

La diferencia esencial entre la Escuela Norteamericana de género y la Australiana, a pesar de que ambas enfatizan el aspecto social y la estructuración de los géneros, se fundamenta en la defensa de la dinamicidad y flexibilidad de los géneros de la primera (Miller 1984; Berkenkotter y Hucking 1995; Yates 1989; Bazerman 1988, 1991, 1995), frente a la mayor estaticidad de la segunda debido a sus principios sistémico funcionales. Esta flexibilidad y dinamicidad ha sido defendida por gran cantidad de autores posteriores (Devitt 1991; Fortanet *et al.* 1999; Todorov 1990; Swales 2004), ya los géneros deben de ajustarse y modificarse según la evolución y exigencias del contexto sociocultural al que deben su existencia (Berenkotter and Huckin 1995a: 24).

El campo de los lenguajes de especialidad confieren una valor esencial al cambio en los género, una dinamicidad que los hace evolucionar a la par de la sociedad que los usa (Swales 1990; Devitt 1993; Fortanet *et al.* 1998). Esta naturaleza dinámica se agudiza de manera especial en los géneros digitales (Posteguillo 1997, 2002, 2003). La flexibilidad en la estructura del género es un aspecto esencial en nuestra concepción, ya que esta noción es central dentro del universo textual web en constante evolución (Kennedy y Shepherd 2005), de ahí que la revisemos en mayor profundidad con posterioridad (véase 5.6).

Aunque en menor medida que en Lingüística, desde la Traductología se han presentado varias definiciones de género (Reiss 1971; Reiss y Vermeer 1984; Hatim y Mason 1990; Bell 1991; Göpferich 1995; Trosborg 1997; etc.), si bien no siempre se emplea el término «género», sino *textsorte* o «clase de texto» (Reiss y Vermeer 1984; Göpferich 1995), o *text form* (Bell 1991). En nuestras labores de investigación, la definición de mayor repercusión en estudios traductológicos es la presentada por Hatim y Mason (1990, 1997; Hatim y Munday 2004: 340):

Genres are conventionalized forms of texts which reflect the functions and goals involved in particular social occasions as well as the purposes of the participants in them (Hatim y Mason 1990: 69).

Los géneros se presentan como actos de comunicación que se expresan por medio de formas textuales convencionalizadas; a su vez, estos textos reflejan las funciones de una situación social determinada y recurrente. En palabras de García Izquierdo (García Izquierdo *et al.* 2002), este concepto suma los aspectos más importantes en la noción de género: rasgos formales, *conventionalized forms of texts*, cognitivos, *purposes of the participants* y socioculturales, *particular social occasions*. Hatim y Mason desarrollaron esta definición a partir de la propuesta del sociolingüista de la escuela australiana Kress (1985: 19), y por tanto hemos reseñar la influencia de la Lingüística sistémico-funcional en su concepción. En cualquier caso, si aplicamos preliminarmente esta definición al género *sitio web corporativo*, éste supondría una forma textual convencionalizada (Nielsen y Tahir 2002; Adkisson 2002; Krug 2005) que refleja los propósitos de una empresa que precisa comunicarse de manera efectiva con sus clientes en una situación social recurrente, la necesidad de comunicación empresa-clienta antes, durante y tras la venta, a su vez sujeta a las limitaciones que impone el medio, Internet.

Con un enfoque integrador de los aspectos intra y extratextuales, aún dentro del modelo sistémico funcional de Halliday, otra propuesta de impacto sería el modelo de Trosborg (1997: 11), que propone definir el género a través de criterios múltiples<sup>159</sup>, usando las variables de campo, modo y tenor. En primer lugar, el campo se referiría al componente ideológico del contenido lingüístico. El tenor se centraría en el componente interpersonal que se define por las funciones comunicativas en la relación emisor-receptor. Por último, el modo se referiría al componente que representa el medio. Uno de los pilares del modelo de Trosborg es la multidimensionalidad e hibridación genérica (Trosborg 1997), postura ésta que

---

<sup>159</sup> En respuesta a la propuesta de Bhatia (1993) que centraba su propuesta de género en la existencia de un propósito o finalidad comunicativa común.

secundamos ya que en los géneros digitales ambas son esenciales debido a su constante evolución. Ésta se materializa en una mezcolanza de géneros, tipologías y registros (Tercedor 2005).

Por su afán sintetizador, y desde la misma perspectiva que la anterior, señalamos la definición de Hurtado Albir (2001):

Los géneros son agrupaciones textuales que comparten una situación de uso determinada, con emisores y receptores particulares, que pertenecen a un mismo campo y-o modo, generalmente con una misma función (o funciones) y tono textual, y que tienen características textuales convencionales, fundamentalmente en cuanto a su superestructura y ciertas formas lingüísticas fijas (Hurtado Albir 2001: 497).

Si nos centramos en los géneros web, el investigador Erickson (1999, 2000) define el concepto de género en el contexto de su estudio sobre las comunidades virtuales en la red o Web<sup>160</sup> como:

A patterning of communication created by a combination of the individual (cognitive), social, and technical implicit in a recurring communicative situation. A genre structures communication by creating shared expectations about the form and content of the interaction, thus easing the burden of production and interpretation (Erickson 1999: 2).

Como reflexión final sobre la definición de género, volvemos a señalar que tras dos décadas de investigación sobre géneros textuales, John Swales (2004), cuyos trabajos fueron decisivos en la introducción del mismo en Traductología, no se muestra partidario de buscar una definición única al concepto de género:

In the light of all the various writings on genre over the last decade, I am less sanguine about the value and viability of such definitional depictions (Swales 2004: 61).

Nos hacemos eco de esta llamada de atención, que señala cómo la adopción de definiciones estáticas puede dificultar o impedir el análisis de nuevos géneros o de aquellos por investigar. Consecuentemente, la flexibilidad necesaria en toda

---

<sup>160</sup> Que el autor identifica con un género.

investigación sobre géneros textuales se debe extender a cualquier propuesta de definición que lo encuadre. En consonancia con esta concepción, en la presente tesis caracterizaremos el género de estudio basándonos en los rasgos del modelo de Gamero (1998, 2001) (véase 5.5), que modificaremos para adaptarlo a los textos web, y mantendremos una concepción flexible y no necesariamente universalista.

### **5.1.2. Las clasificaciones por géneros: los subgéneros y supragéneros**

Los géneros raramente existen de manera independiente, sino que forman cadenas, redes o estructuras jerarquizadas (Swales 2004). Los analistas del género han propuesto la existencia de diferentes super y subcategorías que responden a diferentes niveles de agrupaciones que, aunque comparten ciertos parámetros de caracterización, no constituyen géneros determinados y concretos. Nuestra tesis se encuentra inmersa en el mundo híbrido y evolutivo de los géneros digitales y, consecuentemente, se hace imperativa la adopción de esta perspectiva (Bhatia 2002).

En primer lugar, las investigaciones sobre géneros textuales han propuesto la existencia de **subgéneros**: tipos de textos secundarios frente a los primarios, los géneros en sí. Ésta constituiría una de las mayores aportaciones<sup>161</sup> de Bhatia (1993), que los define como subcategorías de géneros que presentan un propósito comunicativo y unas estrategias para llevarlos a cabo ligeramente diferente de los géneros de los que parten (Vázquez del Árbol 2006). Bhatia señala que los subgéneros surgen según ligeras modificaciones en los propósitos comunicativos: «minor changes or modifications in the communicative purpose(s)» (1993: 14). Biber (1988) ya indicó, con anterioridad a Bhatia, que los criterios de clasificación entre géneros pueden variar<sup>162</sup> «some genres have several sub-classes which are quite different from one another» (Biber 1988: 170). Esta variación puede

---

<sup>161</sup> Según Connor (1996) que parte desde el campo de la Retórica.

<sup>162</sup> Biber señala el tema y el propósito del hablante como criterios diferenciadores de géneros.

potencialmente resultar en la aparición de subgéneros agrupados según distintos criterios lingüísticos. La concepción de subgénero de Silvia Gamero es más amplia y se refiere a «variantes que surgen dentro de un género» (1998: 157). Esta concepción sería más apropiada a nuestros objetivos, puesto que defendemos que la subdivisión en subgéneros no ha de fundamentarse ineludiblemente en un criterio único, tal como la función textual señalada por Bhatia (1993), a pesar de que este parámetro sea predominante<sup>163</sup>. Un ejemplo sería el género *blog* de reciente aparición (Santini 2005; Herring *et al.* 2004), sobre el que ya se ha desarrollado un subgénero al producirse un cambio en el modo predominante del escrito al audiovisual, el *videoblog*.

En los estudios de género se ha propuesto además el concepto de **supragénero** (Izquierdo y Nebot 2003), una agrupación de géneros de similares características y en un mismo ámbito, si bien no se corresponde con un género delimitado y concreto. En la presente tesis emplearemos el término «supragénero» por ser más intuitiva y clara frente a otras denominaciones desde diversas perspectivas como «etiqueta supragenérica», «pre-género» (Swales 1990, 2004), «macrogénero»<sup>164</sup> (Martin 1992) o «colonia de géneros» (Bhatia 2002)<sup>165</sup>. Según Swales (2004), este concepto supone una macroagrupación, una especie de guía intuitiva que nos orienta a la hora de clasificar los textos. Ahora bien, el investigador prescribe ser cuidadosos al manejar este concepto puesto que no se corresponde con géneros determinados. Bhatia (2002) describe este concepto como «a constellation of closely related and overlapping genres, sometimes within but often across discourse communities», con lo que su concepción supondría géneros de

<sup>163</sup> Dentro de los sitios web el «campo» puede ser un factor de división entre los mismos (Bolaños 2002).

<sup>164</sup> Para una reflexión sobre la diferencia entre «supragénero» y «macrogénero» en el contexto del proyecto de investigación GENTT de la Universidad Jaume I, consultar el artículo de García Izquierdo y Mozón Nebot (2003).

<sup>165</sup> No se deben confundir estas macroagrupaciones de géneros con el concepto *genre system* de Bazerman (1994), ya que éste se refiere al conjunto de géneros interrelacionados de que dispone una comunidad de discurso especializada determinada, como sería el conjunto de géneros especializados a disposición de los miembros de una comunidad de discurso en el ámbito jurídico o médico.



características similares que no llegan a constituir un género único y que pueden ser usados por comunidades de discurso diferentes.

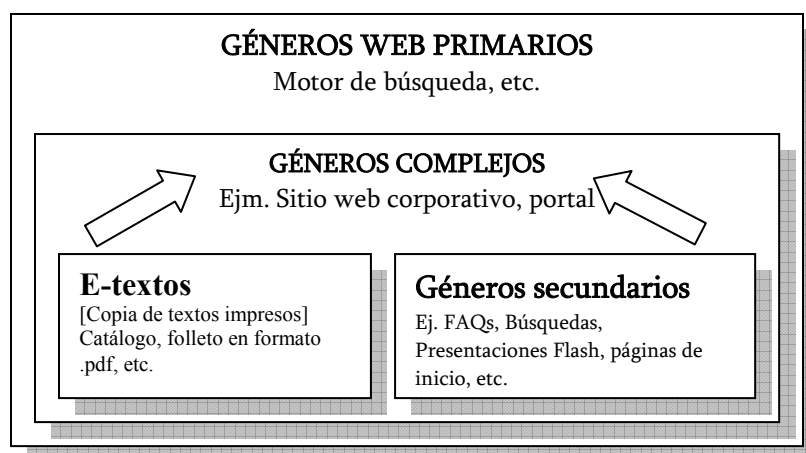


Figura 22. Relación de entre géneros web complejos y secundarios.

Aparte de la existencia de subgéneros y supragéneros, existen otras propuestas de variación genérica de interés. Los australianos Martin (1995) y Hanks (1996: 242-243) introdujeron el concepto de «géneros complejos», definidos como géneros textuales que pueden contener en sí mismos ejemplares de otros géneros textuales. Los autores proponen como ejemplo una misa en la que se pueden incluir salmos, oraciones o sermones. Martin (1995: 24) se refiere a la relación entre géneros complejos y secundarios como *genre embedding*<sup>166</sup>. En sintonía con esta propuesta encontramos un paralelismo con los conceptos de textos simples, complejos y complementarios<sup>167</sup> de Reiss y Vermeer (1984). Los textos «complejos» se definen como aquellos en los que se insertan otros textos, con lo que se solaparía con los «géneros complejos» presentados anteriormente. Consideramos necesaria esta distinción para imbricar nuestra investigación con la principal corriente de análisis sobre géneros digitales iniciada por Crownston y Williams (1997) y Shepherd y Watters (1998, 1999). En estos estudios se presentan ciertas páginas web como

<sup>166</sup> Bhatia introduce también el concepto de *genre embedding* para señalar la posible inclusión de un género en otro (Bhatia 1997).

<sup>167</sup> Estos textos dependen enteramente de un texto simple o primario para su realización, como una crítica literaria de un libro o una reseña.

géneros independientes en lugar de sitios completos, tales como las *Frequently Asked Questions* (Crownston y Williams 1997) o las presentaciones Flash (Paolillo *et al.* 2006). Éstas se consideran en nuestra tesis «bloques comunicativos» integrantes de un género «complejo» único (véase 5.3.1), ya que defendemos el carácter unitario e integral de todo sitio web corporativo.

De hecho, estos géneros secundarios nunca existen *per-se*, sino que se presentan incorporados al propósito comunicativo del sitio global en el que se encuadran. Estos ejemplares, que a menudo constituyen e-textos (véase 5.5.6.1), se imbrican en los sitios web como «género complejo», y los consideramos géneros secundarios web ya que dependen de un «género complejo» o «género primario» para su realización. En este sentido, defendemos que esta doble vía de clasificación es positiva y enriquecedora.

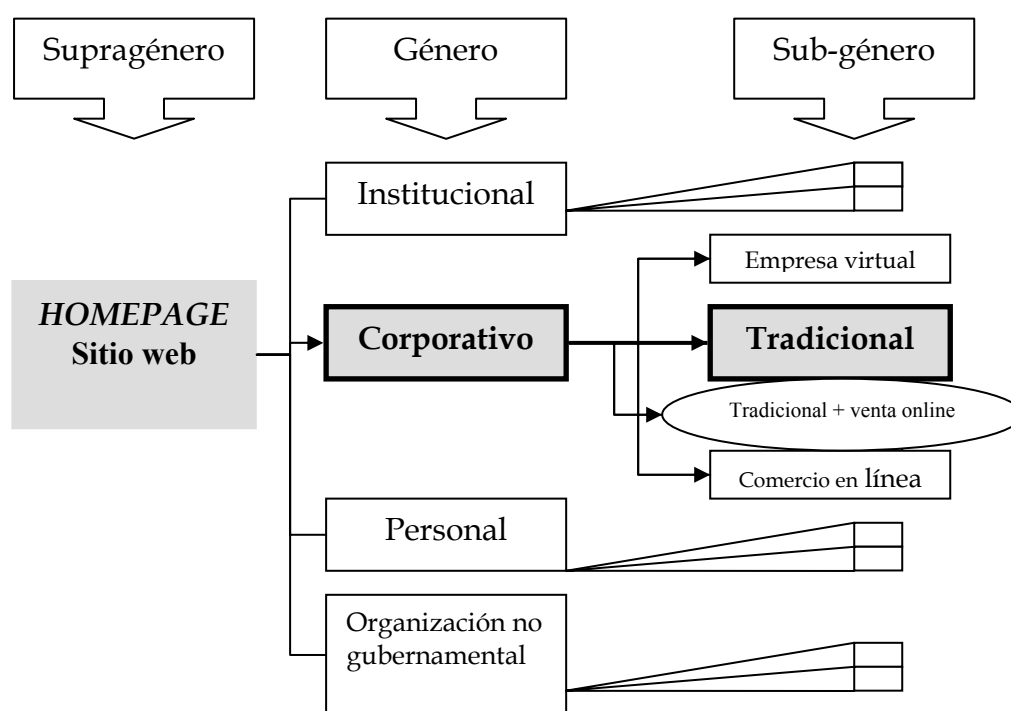
### 5.1.2.1. Propuesta de situación del género objeto de estudio

El objeto de nuestra investigación sería el género *sitio web corporativo*, o *corporate homepage* en inglés (Shepherd y Watters 1999). Éste constituiría un género primario o «complejo» al poder incorporar géneros secundarios, tales como un contrato web o un libro de instrucciones en formato .pdf (véase 5.1.2). Este género se puede agrupar dentro del «supragénero» denominado *homepage*<sup>168</sup>, que aglutina los géneros: sitios web **corporativos**, los **institucionales**, los **personales** (Kennedy y Shepherd 2005), así como los sitios web de **organizaciones no gubernamentales** (Nielsen y Loranger 2006). A su vez, y tras el exhaustivo análisis durante la etapa preliminar de selección textual para componer nuestro corpus, identificamos que el género *sitio web corporativo*, mediante cambios en la función o conjunto de funciones y los elementos de la situación comunicativa, se puede

---

<sup>168</sup> Este término anglosajón no se refiere a la «página de inicio» o «home» como bloque comunicativo en los sitios web, sino que en términos amplios constituye una plataforma de comunicación interactiva entre un sujeto o entidad reconocida como tal por la sociedad (empresa, ONG, institución, etc.), con una audiencia que necesita básicamente información sobre la misma, o servicios y productos ofrecidos por el emisor, y cuyo propósito son principalmente expositivos o exhortativos.

dividir en subgéneros como (1) los sitios web corporativos clásicos, (2) los basados en la venta en línea, o (3) los sitios web que de empresas con presencia exclusiva en la WWW. Igualmente, y dada la ya aludida hibridación genérica, identificamos un (4) subgénero que combina el sitio web corporativo tradicional con la venta en línea de productos o servicios, y que hemos situado entre ambos. Este género se descubrió en la investigación piloto al descargar sitios web con un número de páginas superior a 30.000 páginas web, que se corresponden con el catálogo de productos para su venta en línea<sup>169</sup>. El siguiente gráfico presenta nuestra propuesta de situación del género objeto de estudio, e incluye los subgéneros característicos y su agrupación bajo el supragénero *homepage*<sup>170</sup>:



**Figura 23. Propuesta de situación del género objeto de estudio, el sitio web corporativo tradicional.**

<sup>169</sup> Otros sitios web cuya función no es la venta en línea de productos, pueden incorporar el catálogo completo de productos pero en formato .pdf. Este formato es indicativo de la falta de adaptación al medio de este e-texto.

<sup>170</sup> Hemos decidido dejar en inglés el término anglosajón «homepage» (Askehave y Nielsen 2005), existente en los estudios de géneros digitales en inglés, que no se corresponde con el bloque «página de inicio o principal» en los sitios web. Los autores reconocen igualmente la posible confusión terminológica de este término anglosajón con «página de inicio» o *page d'accueil* en francés.

Esta propuesta no pretende ser exhaustiva en lo que respecta a la totalidad de los géneros o subgéneros bajo el supragénero *homepage*, ya que dentro los sitios web institucionales se engloban los sitios web de ayuntamientos, diputaciones o delegaciones de turismo, o sitios web de ONG bajo el género *sitio web de organizaciones no gubernamentales*.

## 5.2. El género y la noción de prototipo

En el análisis de género la noción de prototipo es central, puesto que libera al investigador de la necesidad de precisar todos los rasgos de caracterización al definir un género determinado (Swales 1990; Paltridge 1997). Esta noción de prototipo surge de los trabajos del filósofo Wittgenstein (1963) y de la psicóloga Eleanor Rosch (1973, 1975, 1977), cuya semántica de prototipos recoge cómo los principios de categorización están condicionados por factores socioculturales, a la vez que demuestra como los sujetos poseen una imagen mental prototípica con la que comparan los ejemplares que se les presentan. Asociado al concepto de prototipicidad se encuentra la noción de «fuzziness» (Labov 1973), entendida como la inexistencia de unos límites claros en toda categorización. Este aspecto ya se observó en las primeras clasificaciones de géneros digitales y hemos de reclamar su importancia en nuestra investigación. Por ejemplo, Shepherd y Watters (1998) señalan que su taxonomía es «difusa» dado que continúan apareciendo nuevos géneros digitales y fusionándose los existentes.

Son muchos los investigadores que han defendido la importancia del carácter prototípico del género, (Swales 1990, 2004; Snell-Hornby 1995: 51; Paltridge 1997; Corpas 2003; etc.). El género, como patrón que surge de la experiencia de los sujetos en unas situaciones sociales específicas y recurrentes (Hurtado Albir 2001), aparece en la encrucijada de elementos intra y extratextuales. Si la configuración pragmática que confieren estos últimos coincide entre dos o más textos, los sujetos

los identificarán como ejemplares de un mismo género, aun cuando difieran en mayor o menor grado en su configuración intralingüística. En este sentido, Swales (1990) remarca el predominio en la formación del prototipo de los factores extratextuales, tales como el propósito comunicativo, la forma, la estructura y las expectativas de la audiencia, que a su vez «operate to identify the extent to which an exemplar is *prototypical* of a particular genre»<sup>171</sup> (Swales 1990: 52). Así, una vez identificado como ejemplar de un género determinado, éste se apreciará como un ejemplo prototípico por la comunidad de discurso a la que pertenezca: «the exemplar will be viewed as prototypical by the parent discourse community» (1990: 58). En esta aproximación al análisis de género las características se describen como rasgos de un género particular, que no aspectos mediante los cuales se pueda definir. En este mismo sentido, Paltridge (1997: 56) apunta que en el proceso de asignación de un género a una categoría específica predominan, dentro de la noción de prototipo y frente a las características internas, ciertos aspectos extratextuales de los eventos comunicativos, como los pragmáticos y perceptuales.

### **5.3. El género como estructura textual**

La noción de género como estructura textual convencional surgió de los estudios de folclore, en los que se consideraba el único aspecto definitorio (Paltridge 1997). Esta importancia de la estructura textual se mantuvo en los estudios desde una perspectiva sistémico funcional (Martin 1989; Kress 1985; Hasan 1989). No obstante, desde la perspectiva de la denominada «Nueva Retórica» norteamericana se analiza la estructura textual en relación con el entorno sociocultural del texto (Miller 1984), con lo que ésta deja de ser un elemento determinante en su definición. En su lugar, la flexibilidad y la posibilidad de evolución se convierten en elementos esenciales.

---

<sup>171</sup> Énfasis original.

Los géneros son el resultado de una situación comunicativa recurrente y se organizan según una estructura retórica formada por «etapas» o «movimientos» (Swales 1990; Nwogu 1991, 1997). Este concepto surge del marco teórico de los *schemata* (Rhemerlhart 1975, 1980), y se puede definir como:

A text segment made up of a bundle of linguistic features (lexical meaning, propositional meanings, illocutionary forces, etc.) which give the segment a uniform orientation and signal the content of discourse in it (Nwogu 1997: 122).

Estos segmentos suponen subtextos de diverso tipo que articulan la estructura discursiva de un género. A su vez, estos movimientos se pueden a su vez subdividir en «submovimientos» o «subfunciones» de orden menor (Nwogu 1997), también denominados «unidades significativas» y «subunidades significativas» (Gamero 1998). Estas unidades representan el contenido de cada una de las partes del texto (Paltridge 1997: 111) y pueden poseer una función específica que contribuya a la función global de un texto. El conjunto de elementos constitutivos de la estructura de un género dado componen su superestructura y macroestructura<sup>172</sup>, y su orden puede ser bastante estricto, como ocurre en el género *artículo de investigación médico* (Swales 1990, 2000, 2004; Paltridge 1997; Nwogu, 1991, 1997):

The moves and their constituent elements assign functions to segments of information which together in the research paper constitute the overall semantic macrostructure of such texts (Nwogu 1997: 135).

En el presente estudio denominaremos estas unidades «bloques comunicativos» y «secciones comunicativas» siguiendo el trabajo de Gamero (1998, 2001), frente a otras denominaciones para estos subtextos como «movimientos o etapas», *moves-step*<sup>173</sup> (Swales 1990; Nwogu 1997), *keys*, (Paltridge 1997), *triad-key* (Vázquez Marruecos 2002, 2005, 2006), o «subtextos de orden primario, secundario o

---

<sup>172</sup> Dependiendo del objeto de estudio.

<sup>173</sup> Askehave y Nielsen (2005) en su análisis de la estructura del sitio web corporativo, emplean el concepto de «movimiento» o «move» de Swales y Bhatia.

terciario» (Fernández Sánchez 2004), ya que en nuestra opinión estos términos se ajustan mejor a los textos web y facilitan la comprensión de este concepto.

Una de nuestras adaptaciones básicas de esta visión teórica a los géneros digitales consiste en que los *bloques* y *secciones* pueden identificarse con las diferentes páginas<sup>174</sup> integrantes de sitio web (Askehave y Nielsen 2005). No obstante, esta estructura no sería lineal y fija, como puede suceder en un artículo de investigación biomédico (Nwogu 1997; Paltridge 1997), ya que el orden temporal de recepción de movimientos o bloques en todo proceso de lectura hipertextual lo dirige el usuario (Fritz 1999). Por consiguiente, en esta adaptación de la concepción de género desaparece la noción de movimiento o bloque como secuencia jerárquica temporal en la producción y recepción textual (Bhatia 1993: 30). La siguiente cita de Askehave y Nielsen (2005: 127) es clarificadora en lo que respecta a esta adaptación:

Due to the fact that the hypertextual nature of the web documents forms is an integral part of the generic properties and communicative purpose of these texts, one must not overlook the media characteristics when describing web-mediated texts as genres (Askehave y Nielsen 2005: 136).

La estructura tampoco es estricta, puesto que dependiendo de cada género pueden existir bloques con mayor o menor grado de prototipicidad. Así un usuario, como miembro de la comunidad de discurso que accede a un sitio web corporativo, esperará encontrar ciertos bloques en forma de hipervínculos, como una página de contacto o información sobre la empresa, junto con otros opcionales como noticias sobre la empresa o catálogos de productos. En ciertos géneros web la superestructura se puede visualizar en los mapas web, que a su vez pueden ser

---

<sup>174</sup> No obstante y como indicamos anteriormente, en la teoría del hipertexto los «nodos» constituyen las unidades en la que se subdividen los hipertextos. Sería incorrecta la identificación exclusiva de un nodo con una página web determinada (Codina 2003), aunque en nuestra opinión un nodo tampoco podría identificarse en un texto unitario ya que depende de las relaciones estructurales y de cohesión para su correcta comprensión por parte del usuario.

jerárquicos o conceptuales (Engbresten 2000: 222). Como ilustración de este aspecto, presentamos un mapa web que refleja esta estructura jerárquica del texto:



Figura 24. Superestructura de un sitio web representada en el mapa del sitio web jerárquico (Engbresten 2008). [www.lifescan.es](http://www.lifescan.es)

En la presente tesis defendemos la plausibilidad de identificar este mapa web con una representación visual de la superestructura textual que subyace tras todo sitio web corporativo. La superestructura depende de cada género y constituye uno de los aspectos de mayor relevancia en los estudios el género, sobre todo en aquellos altamente convencionalizados (Göpferich 1995b; Gamero 1998, 2001). El conocimiento de la superestructura de un género y las convenciones asociadas a su conjunto, o a cada bloque comunicativo, es esencial durante el proceso de traducción (Gamero 2001; Vázquez del Árbol 2006), de ahí que dada su importancia en la práctica y didáctica de la traducción pasemos a revisar este concepto.



### 5.3.1. La superestructura textual en los géneros web

En proceso de análisis intralingüístico de un género textual, la superestructura se sitúa como la categoría de mayor rango jerárquico y constituye el punto de partida sobre el que analizar el resto (Göpferich 1995). Las investigaciones en este campo definen la superestructura como el esquema estructural prototípico de un género compuesta de varios bloques textuales que difieren en su rango y que están organizados en un orden con distinto grado de fijación dependiendo del género en cuestión (Hurtado Albir 2001: 495). La definición de Göpferich (1995) remarca este carácter de patrón estructural convencional:

La secuencia convencionalizada en que se suceden de forma jerarquizada, pero flexible dentro de unos límites, elementos del texto invariables a nivel temático y funcional destinados a desarrollar su tema (Göpferich 1995: 127)<sup>175</sup>.

Este carácter de estructura textual más o menos fijada se refleja asimismo en la definición de Hurtado Albir (2001: 495): «un esquema estructural típico de un género que se compone de varias partes con un orden más o menos fijo». El estudio descriptivo de Gamero (1998, 2001) concluye que la superestructura es el mayor elemento cohesionador del género técnico objetivo de su estudio, los manuales de instrucciones. La superestructura se identifica como el elemento más característico en géneros altamente convencionalizados (ibid), es dependiente del género al que pertenezca el texto (Göpferich 1995), y llega a constituir el principal elemento cohesionador de géneros técnicos (Gamero 1998). Coincidimos con estas investigadoras en que el resto de categorías convencionalizadas de cada género deben analizarse en el marco de las subunidades textuales que conforman la superestructura y, consecuentemente, organizaremos nuestro análisis posterior de

---

<sup>175</sup> Traducción de Fernández Sánchez (2004: 249).

este modo. La superestructura puede además ser una guía con la que diferenciar subagrupaciones o subgéneros en el seno de un mismo género (ibid).

Siguiendo a Gamero (1998, 2001), denominaremos de manera diferenciada a los segmentos textuales convencionales que se ven dispuestos en una organización recurrente propia de cada género. En primer lugar se encuentran los **bloques comunicativos**, que constituyen subtextos de primer orden y representan una función determinada dentro de la configuración textual global. Como veremos, los sitios web corporativos presentan un bloque comunicativo que denominamos de «contacto», cuya aparición es un requisito básico dado el estatus de plataforma de comunicación entre usuarios y empresas de estos textos (Lommel 2007: 5). En segundo lugar en el rango jerárquico se encuentran las **secciones comunicativas**. Éstas poseen una función específica dentro de la función global del bloque comunicativo en el que se integran. Como ejemplo, las secciones comunicativas «historia» y «localización» de carácter opcional se enmarcan dentro del bloque con información sobre la empresa. Su función subsidiaria es expositiva en descripción y se integra dentro una función global de este último bloque, dirigido a consolidar el estatus y reputación del emisor. En este bloque identificamos además un foco secundario exhortativo, puesto que se pretende influir en la conducta del usuario mejorando su valoración subjetiva con el fin ulterior de fomentar el establecimiento de una relación comercial.

Las investigaciones contrastivas de género apuntan que la superestructura puede presentar importantes diferencias entre un mismo género en culturas diferentes (Gamero 2001; Mendiluce 2005), o entre diferentes subgéneros dentro de una misma lengua (Vázquez 2006: 155; Fernández Sánchez 2004: 907). La primera constituye una de nuestra hipótesis de partida que aplicaremos a un géneros digital y verificaremos en nuestro estudio (véase 9.1.1). La segunda vía de estudio centrada en la variación superestructural entre subgéneros en misma lengua sería

posible a partir de nuestro corpus, por ejemplo, entre subgéneros de grandes, medianas y pequeñas empresas, o entre sitios web de empresas en diferentes áreas económicas, pero dejaremos a un lado esta posible vía de investigación para el futuro.

La pertinencia de analizar intratextualmente los sitios web corporativos a partir de la superestructura se fundamenta en el carácter altamente convencionalizado del género en cuestión (Kennedy y Shepherd 2005), por lo que los usuarios los reciben con un modelo mental difuso que incorpora sus elementos textuales convencionalizados (Vaughan y Dillon 2006). Este carácter convencional de la superestructura se hace patente en los bloques textuales recomendados para los sitios web corporativos en los manuales de usabilidad (Nielsen y Loranger 2006; Nielsen y Tahir 2002), en las guías de estilo web (Price y Price 2002; Mill 2005; Bly 2002; etc.), así como en investigaciones previas sobre la superestructura de este género<sup>176</sup> (Robbin y Stylianou 2003; Liu *et al.* 1997; Bolaños 2002; Adkisson 2002; etc.).

En una concepción alternativa, la superestructura textual web se ha identificado en investigaciones previas con los elementos de la estructura de presentación de toda página web, como el cuerpo de la página, los menús de navegación o el encabezamiento (Bolaños 2002a, 2002b). Estos compartimentos se encuentran convencionalizados y se corresponderían con la estructura intranodal, la estructura interna de cada página. Sin embargo, hemos de reseñar que estos elementos constituyen compartimentos recurrentes abiertos a distintos contenidos proposicionales, de ahí que creamos más productivo identificar el concepto de «superestructura» con el esquema general global recurrente cuyos contenidos proposicionales reconoce el receptor. Pasamos a describir dos de estos elementos

---

<sup>176</sup> Aunque la mayoría de estas investigaciones se centra en los «contenidos» de los sitios web corporativos y se menciona implícitamente el concepto de superestructura.

textuales básicos en la estructura de presentación de toda página web: el texto en el cuerpo de la página web y los menús de navegación.

### **5.3.1.1. El cuerpo de la página web: el texto de «contenido» frente al texto «interfaz»**

El cuerpo de cada página web constituye el contenido diferenciado de cada página web dentro del sitio global. Este subtexto constituye el contenido diferenciado de cada página, resumido en su título, y otorga a cada página su carácter como unidad de contenido y almacenamiento. Price y Price (2002: 70-73) distinguen dentro del texto visible en la estructura de superficie entre el texto «interfaz» y el texto de «contenido». El texto de «interfaz» es todo subtexto de carácter metatextual cuyo objetivo es articular el hipertexto global. Estos metatextos se suelen presentar en menús de navegación o en resúmenes de contenidos a la derecha de la pantalla. El texto «interfaz» está marcado tipográficamente y funcionalmente mediante las convenciones propias de los hiperenlaces, como su distinto color, el subrayado, o la transformación del puntero del ratón a su paso sobre el mismo. El carácter hipertextual de los sitios web conlleva la repetición del mismo texto «interfaz» en la mayoría de páginas, con lo que los análisis léxicos automáticos pueden quedar sesgados por la alta recurrencia de ciertos elementos léxicos pertenecientes a este tipo de textos (ej. texto en menús de navegación).

Por su parte, el texto de «contenido» se corresponde con el cuerpo de toda página web, y supone el conjunto de contenidos proposicionales diferenciados a los que se les otorgó el carácter de unidad de contenido y almacenamiento en el contexto de la WWW (Nielsen y Loranger 2006; McGovern *et al.* 2002).



Figura 25. Distribución textual en página web. De [www.microrent.es/index.html](http://www.microrent.es/index.html).

Los experimentos en usabilidad concluyen que la organización textual óptima dirigida a la lectura en pantalla se compone de: (1) cabecera, (2) pie de página, y tres columnas, (3) el menú de navegación a la izquierda, (4) el cuerpo de la página en el centro, y (5) otra columna a la derecha con contenidos resumidos del sitio. En esta estructura óptima, la columna central contiene el texto de «contenido», y se reserva para el texto «interfaz» la cabecera, la columna izquierda y la derecha. La página web en la Figura 25 incluye los elementos de navegación básicos en la cabecera, junto con el logotipo de la empresa en la esquina superior derecha siguiendo una de las convenciones más extendidas (Nielsen y Loranger 2006; Adkisson 2002; Bolaños 2002b), el menú de navegación secundario en la columna izquierda, y elementos resumidos del sitio web en la columna derecha. Finalmente, el contenido específico de esta página web se encuentra en la columna central.

### 5.3.1.2. Los menús de navegación

Los menús de navegación representan un elemento obligatorio en todo sitio web, principalmente por su labor estructuradora en hipertextos y por su función contextualizadora y cohesionadora. A su vez, representan un mapa hipertextual del

sitio web global que permiten al usuario contextualizar cada página con respecto al texto. Éste facilita su interpretación al situar cada página en el modelo mental textual global del usuario (Price y Price 2002: 264). Su carácter metatextual y estructurador hace plausible su identificación como indicadores directos de la superestructura textual, por lo que desempeñarán un papel fundamental en nuestro análisis superestructural posterior.

En nuestro estudio desempeñan un papel fundamental en la diferenciación entre géneros digitales que dependen de Internet para su realización, los género *noveles* que presentaremos posteriormente (véase 5.5.6.1), de los géneros *replicados*, géneros impresos que se cuelgan en Internet sin funcionalidad alguna, tales como un folleto en formato .pdf (Shepherd y Watters 1998). Hoy en día podemos afirmar su presencia se ha convencionalizado, de hecho, su presencia forma parte integrante del prototipo genérico de la mayoría de géneros web (Price y Price 2002: 255). En la estructura interna de cada página, estos menús pueden aparecer preferentemente en la parte izquierda, en la parte superior o en ambas, así como en la columna de la derecha aunque en menor medida (Bolaños *et al.* 2005).



Figura 26. Menú de navegación superior. [www.mahou-sanmiguel.com/inicio.html](http://www.mahou-sanmiguel.com/inicio.html)

Según Jeney (2007) el objetivo de todo menú es:

Effective menus greet visitors with a visual representation of the main topics or areas of the site and set up expectations. [...] also provide the categories and scope of information available on the site (Jeney 2007: 52).

Para conseguir este objetivo, las guías de estilo recomiendan que los menús sean extremadamente claros, breves concisos y descriptivos, con un número de elementos limitado (Garrand 2001), sin que pueda haber ambigüedad alguna (Jeney

2007: 54). No obstante, la brevedad es relativa, puesto que es preferible una construcción más larga si así se evita toda ambigüedad (Price y Price 2002: 237). Su función es eminentemente descriptiva y, por consiguiente, se ha de evitar el humor, el sarcasmo o la ironía. Estos recursos discursivos dificultan la navegación al tener que ser interpretados por el usuario (Price y Price 2002: 238). En el caso de hipertextos de gran extensión, se recomienda su jerarquización para facilitar la construcción de modelos textuales que se almacenan en la memoria a largo plazo. Esta estrategia reduce la demanda de atención al facilitar la percepción de la estructura global textual (Spyridakis 2000).



**Figura 27. Pie de página.**www.azud.com\es\index.html

Finalmente, los textos «interfaz» o de navegación se pueden situar en los pies de página, unos subtextos en los que tipográficamente se ha convencionalizado un tamaño de fuente más pequeña. En esta área suelen aparecer subtextos obligatorios en toda página que el usuario no suele leer (Brink *et al.* 2002), como avisos legales, información de contacto, el registro de derechos de reproducción<sup>177</sup> o los mapas web.

## **5.4. Estudios contrastivos empíricos basados en géneros**

La metodología de análisis empírico de géneros contrastivos con fines descriptivos, didácticos y prácticos era casi inexistente en el panorama hispánico de los años

---

<sup>177</sup> Bolaños (2002; 2003) señala como elementos recurrentes en los pies de página: «información sobre la fecha de la última actualización, el registro de los derechos de reproducción o la resolución idónea de visualización. En ocasiones aparece también el navegador recomendado y el nombre del creador y responsable del sitio web en cuestión» (Bolaños 2003: 46). En contraste con esta investigación, la llevada a cabo por Renau (2004: 323) sobre sitios web corporativos recoge que los pies de página mostraron mayoritariamente la dirección del *webmaster* y la dirección y correo de la propia empresa.

noventa (Gamero 1998: 437), si bien gozaban de mayor tradición en otros ámbitos como el germánico. No obstante, podemos afirmar que en la actualidad constituye una corriente en alza (Angost 1996, 1999; Gamero 1998, 2001; Borja Albí 1998; Hurtado Albir 1999; Bolaños 2002; Bolaños *et al.* 2005; Vázquez del Árbol 2002, 2006; Fernández Sánchez 2004, etc.). Esta escasez de estudios contrastivos se acentúa en el ámbito de la localización multimedia<sup>178</sup>, a pesar de los claros beneficios de este enfoque descriptivo. Entre sus posibles aplicaciones, esta perspectiva permite conocer mejor esta nueva modalidad, así como los tipos de traducción asociados a la misma con el consiguiente impacto en su didáctica (Hurtado Albir 2001: 505).

El método de análisis de estos estudios se fundamenta en la textología contrastiva de Hartman (1980) y en la estilística de Vinay y Dalbernet (1964). La idoneidad de este método en estudios de géneros como unidad de comparación recae en un doble nivel:

- a) Extratextualmente, ya que una definición funcional y pragmática del género asegura la máxima comparabilidad. Ésta es esencial en los estudios de corpus comparables como el presente (Baroni y Bernardini 2003).
- b) Intratextualmente, ya que el modelo de análisis que se realiza es común en las dos secciones del corpus o entre lenguas. En nuestro caso supondría un mismo análisis de las dos secciones del corpus comparable en una lengua (Baker 1995), el original y el traducido, que argumentamos presentan características diferenciadas (Baker 1998).

Bajo esta perspectiva el análisis es principalmente sincrónico, y se realiza mediante corpus paralelos bilingües compuestos por textos que han surgido en situaciones

---

<sup>178</sup> Con la excepción de Alicia Bolaños (2002).



comunicativas similares pero en el seno de culturas diferentes (Reiss y Vermeer 1984: 194). Estos investigadores defienden el uso de textos originales en estas investigaciones, puesto que las traducciones tienden a mostrar fenómenos de asimilación respecto a las convenciones de género habituales en la sociocultura de partida (Ibid 1984: 195). No obstante, no existen hasta la fecha investigaciones empíricas amplias como la presente que demuestren esta hipótesis en localización. De ahí que coincidamos con Hurtado Albir en reclamar la importancia de estudios empíricos de carácter contrastivo sobre géneros, ya que «sólo la estadística permitirá asentar regularidades fiables» (Hurtado Albir 2001: 505). De igual manera, la investigadora señala la pertinencia de realizar estudios contrastivos como el nuestro:

La formación de corpus formados por originales y traducciones es también de interés para describir la manera que funcionan las traducciones y para analizar si al traducir se siguen o no las convenciones estructurales y lingüísticas propias de la lengua de llegada (ibid).

En este sentido, nuestros objetivos son triples, ya que (1) analizamos descriptivamente las convenciones ya establecidas en un género digital determinado en nuestra cultura, y (2) analizaremos qué tipo de huella deja el proceso de traducción/localización en un género representativo de importancia traductológica, para (3) comprobar si las convenciones de género establecidas de la cultural de partida o de la meta se muestran en las traducciones.

Hasta la fecha, los estudios contrastivos se han dirigido esencialmente hacia géneros de especialidad y pueden, o no, suponer una aplicación directa a la actividad traductora. A su vez, los análisis de género se han realizado por textos globales (Arntz 1990; Göpferich 1995a; Gamero 2001; Vázquez del Árbol 2002, 2006; Fernández Sánchez 2004), corriente en la que se engloba el presente trabajo, o análisis por subtextos principalmente dentro del entorno germánico (Oldenburg 1992; Engberg 1997).

### 5.5. El modelo de género de Silvia Gamero (1998, 2001)

Silvia Gamero propone un modelo de caracterización del concepto de género centrado en el género técnico, que partiendo de las definiciones propuestas en el campo de las lenguas para fines específicos (Swales 1990; Bhatia 1993; Fortanet *et al.* 1998), recoge varios factores que intervienen a la hora de caracterizarlo y definirlo. En su propuesta el género se define a través de cinco parámetros básicos que aplicaremos a nuestra caracterización del género *sitio web corporativo*:

- a) Rasgos convencionales (véase 5.5.1)
- b) Función textual (véase 5.5.2)
- c) Elementos de la situación comunicativa (véase 5.5.3)
- d) Influencia del contexto sociocultural (véase 5.5.4)
- e) Los elementos intratextuales (véase 5.5.5)

Puesto que nos centremos en géneros digitales, nuestra propuesta añade otro rasgo, (f) la funcionalidad (véase 5.5.6), ya que constituye una característica esencial en la evolución y caracterización de los cibergéneros (Shepherd y Watters 1999). De hecho, la funcionalidad se identificó como el principal criterio clasificatorio en las primeras propuestas de caracterización de géneros digitales desde la disciplina de las ciencias de la información (Crownston y Williams 1997; Shepherd y Watters 1998) (véase 5.5.6.1). Este criterio ha comenzado a incluirse en investigaciones empíricas recientes como dimensión clasificatoria de géneros web (Bolaños 2002; Bolaños *et al.* 2005; Rennau 2004; Paolillo *et al.* 2007; Jiménez 2006).

En este modelo se apunta que no todos los parámetros son igualmente decisivos a la hora de caracterizar un género (Gamero 2001: 51). Por otro lado, la flexibilidad ha de extenderse igualmente a las clasificaciones textuales y rasgos que caracterizan a cada género (Hurtado Albir 2001: 506), de ahí que en nuestro

modelo de caracterización incorporaremos esta necesaria perspectiva flexible para dar cabida a la inexorable futura evolución de los géneros digitales.

Pasamos seguidamente a repasar cada uno de estos rasgos que asumimos en nuestra caracterización del género.

### **5.5.1. Rasgos convencionales**

El carácter convencional está presente en la totalidad de estudios de género, ya que las convenciones constituyen una guía que facilita la comunicación y la comprensión, y en Internet, la funcionalidad (Nielsen 2000). Para definir que constituyen las «convenciones de género», se hace inevitable abordar en primer lugar el concepto de «convención» (véase 5.5.1.1), distinguiéndolo además del concepto de «norma» empleado en Traductología (véase 5.5.1.3).

#### **5.5.1.1. Convención**

La definición de convención presentada por Gamero (1998: 50)<sup>179</sup>, se basa en la propuesta de Lewis (1969), que desde la psicología la define como aquellas regularidades observadas en la conducta humana en situaciones de cooperación:

A regularity R in the behavior of members of a population P when they are agents in a recurrent situation S is a convention if and only if it is true that, and it is common knowledge in P that, in almost any instance of S among members of P,

- (1) Almost everyone conforms to R;
- (2) Almost everyone expects almost everyone else to conform to R
- (3) Almost everyone has approximately the same preferences regarding all possible combination of actions;
- (4) Almost everyone prefers that any one more conforms to R on condition that almost everyone conforms to R;
- (5) Almost everyone would prefer that any one more conform to R', on condition that almost everyone conforms to R'

---

<sup>179</sup> La investigadora sigue los estudios de Göpferich (1995).

Where R' is some possible regularity in the behavior of members of P in S, such that almost no one in almost any instance of S among members of P could conform both to R' and R. (Lewis 1969: 78).

En una matización a esta definición, Göpferich (1995; 153-154, cf. Fernández Sánchez 2004: 64) apunta que esta definición de Lewis no hace referencia a cómo surge una convención determinada, si bien este aspecto está presente en una definición previa del autor. Según éste, las mismas surgen por la existencia de un problema de coordinación<sup>180</sup> en una situación recurrente que se resuelve mediante un equilibrio coordinativo<sup>181</sup>, un especie de acuerdo entre las partes. En resumen, las convenciones aparecen un contexto comunicativo (S) determinado cuando se cumplen estos tres requisitos:

- Intervienen varias personas (miembros de una población P).
- Las situaciones (S) se repiten.
- Las personas persiguen una intención que sólo puede conseguirse si todos se comportan de la misma forma (R).

En el marco de la Traductología las convenciones se han definido como:

Implicit or tacit non-binding regulation of behavior, based on common knowledge and the expectations of what others expect you to expect from them (etc.) to do in a certain situation (Nord 1991: 96).

De entre las posibles convenciones, en su propuesta funcionalista Reiss y Vermeer (1984: 180 y ss.) se centran exclusivamente en las «convenciones de género». Nord (1997b, 2003) critica esta importancia que los autores otorgan exclusivamente a las convenciones de género, y presenta otras convenciones de importancia como las

---

<sup>180</sup> Un «coordination problem» según Lewis suponen situaciones de decisiones interdependientes de dos o más agentes en las que predomina la coincidencia de interés y en la que existen dos o más coordinaciones de equilibrio apropiados (1969: 24).

<sup>181</sup> Un «coordination equilibrium» para Lewis supone una combinación de distintas maneras de actuar por parte de varias personas dentro de una situación determinada en la que nadie saldrá ganando si cualquier agente hubiera actuado de otra manera (1969: 14).

convenciones de estilo (Nord 2003), las de conducta no verbal, y las de traducción. No obstante, defendemos que los géneros constituyen un marco práctico de estudio de las diferentes convenciones a diferentes niveles, intratextuales, de estilo, etc. (Gamero 2001).

El papel que desempeñan las convenciones en este entorno aparece en la totalidad de los estudios cognitivos de usabilidad Web (Nielsen 2000; Nielsen y Tahir 2002; Nielsen y Loranger 2006; Krug 2006; Brinck *et al.* 2002; etc.). En una investigación sobre páginas de inicio corporativas, Nielsen y Tahir (2002: 37), ponen de manifiesto que las convenciones de sitios web son de primordial importancia tanto desde el punto de vista de la producción textual como desde su recepción:

[...] over time, we expect more and more **conventions** to emerge.[...] by the time a user arrives at your homepage for the first time, that user will already be carrying a large load of mental baggage, accumulated from prior visits to thousands of other homepages. [...] by this time, users have accumulated a **generic mental model** of the way homepages are supposed to work, based on their experiences on these other sites<sup>182</sup> (Nielsen y Tahir 2002: 37).

En el caso específico de los géneros web, los autores señalan que las expectativas del receptor no sólo se reducen a la estructura discursiva, tipología textual y género, que se engloban bajo este *generic mental model*, sino cómo «funcionan» estos textos<sup>183</sup>. El carácter convencional de esta interacción constituye la razón por la que consideramos apropiado añadir funcionalidad de los géneros digitales como dimensión adicional (Shepherd y Watters 1999).

En esta misma línea se sitúan los resultados del estudio empírica de Vaughan y Dillon (2006), en la que investigaron el impacto de la falta de conformidad a las convenciones de género establecidas en la comprensión, navegación y usabilidad. El

---

<sup>182</sup> Énfasis propio.

<sup>183</sup> En el trabajo más reciente de Nielsen mantiene esta misma concepción de un model mental genérico acumulado y compartido por la comunidad de usuarios que se corresponde con nuestra visión de género (Nielsen y Loranger 2006: 78).

instrumento utilizado se compuso de dos sitios web, uno ajustado a las convenciones del género y otro que las transgredía deliberadamente. Los resultados que obtuvieron apuntan a una mejor comprensión, y a mejoras en la usabilidad, actuación y navegación en aquellos sitios web que presentan las convenciones establecidas en el género. Dentro de los posibles niveles de análisis, el experimento se centró en el nivel estructural y semántico. Las implicaciones para la presente investigación son claras, ya que apoyan nuestra hipótesis de partida referente a la necesidad de reflejar las convenciones establecidas tanto en el contexto original de producción textual como en la localización.

### 5.5.1.2. Convenciones de género

Este concepto emana del anterior en virtud de un mayor nivel de abstracción del primero. Las convenciones asociadas a los géneros son primordiales ya que éstas pueden presentar diferencias entre culturas o lenguas diferentes (Nord 1997b: 54; Gamero 2001), de ahí la necesidad de que el traductor conozca las convenciones propias de la cultura origen y refleje en la traducción las convenciones de la cultura meta<sup>184</sup>. Éstas son esenciales tanto en la producción textual como en la recepción y, consecuentemente, en la traducción (Nord 1997b: 52). Según Gamero (2001: 58-59) no existe ninguna propuesta sistemática de modelo de análisis de estas convenciones de género, si bien recoge los diferentes aspectos convencionales objeto de investigaciones previas.

A continuación presentamos una síntesis de los aspectos que consideramos de mayor importancia dentro del estudio de las convenciones de género:

a) La necesaria existencia de **alternativas** a la convención. La existencia de una convención requiere que exista una alternativa posible, puesto que de otra

---

<sup>184</sup> El encargo de traducción puede no requerir que se sigan las convenciones en la cultura meta (Nord 1997a).

manera no podría existir una convención como tal (Göpferich 1995). La investigadora presenta los conceptos de **alternativa** y **variante** para ilustrar este aspecto. La *alternativa* sería una forma de expresión que no se considera convencional, pero que puede cumplir la misma función comunicativa y es correcta desde un punto de vista gramatical y estilístico. Por otro lado, la *variante* es la gama reducida de posibilidades aceptadas dentro de una convención determinada. El traductor, en los casos de géneros que permiten una mayor variación, ha de conocer además las posibles variantes, ya que el grado de prototipicidad de los mismos depende del género en cuestión. Ahora bien, en casos como los textos no especializados, la variación estilística puede ser un requerimiento (Gamero 2001: 54)<sup>185</sup>. Como ejemplo, en nuestro estudio preliminar (Jiménez 2006) identificamos que en inglés se encuentra altamente convencionalizado el uso del término *Contact us*, presente en el 89% de sitios web comerciales anglosajones (Nielsen y Tahir 2002), mientras que la investigación mostró que el término más convencional en el mismo género en España sería «Contacto», con un 49,42% del total, aunque encontramos 22 alternativas para la denominación de este bloque o página web, todas ellas posibles y correctas aunque con distintos grados de convencionalización.

b) El establecimiento de las convenciones es **arbitrario**. Las convenciones se adoptan al azar y el mecanismo por el que se adoptan, siempre que las posibles alternativas permitan cumplir con éxito la misma función comunicativa, es totalmente arbitrario (Lewis 1969: 70). Como nos comenta Gamero (2001):

El mecanismo que hace que una conducta determinada se convierta en una convención es absolutamente arbitrario, siempre que exista una gama de conductas posibles y que todas permitan cumplir con éxito la función comunicativa (Gamero 2001: 53).

Esta arbitrariedad en el establecimiento de las convenciones es precisamente la responsable de que éstas difieren de una lengua, cultura o *locale* a otra (Gläser

---

<sup>185</sup> Cuanto más especializado sea el género mayor será la convencionalización e importancia de las convenciones de género. La traducción jurídica se presenta como un ejemplo clásico en este sentido (Borja 1998).

1970: 29). Siguiendo los estudios previos, hemos introducido la metodología de la Lingüística de corpus para cuantificar y definir empíricamente los posibles rasgos convencionales. Asumimos así que un elemento característico en un género que pudiera eventualmente considerarse una convención, se distinguirá por su utilización más frecuente. Éste elemento realizará su función comunicativa de forma más efectiva frente a otros elementos que representan las necesarias «alternativas», ya que sin estas variantes no podríamos considerar que exista una convención (Lewis 1969). En este sentido, coincidimos con Gamero que sitúa en 50% el porcentaje necesario (Gamero 1998: 314)<sup>186</sup>.

c) **Competencia activa y pasiva** de las convenciones de género. La importancia de las convenciones en la práctica y docencia en traducción se basa en parte en la distinción entre competencia pasiva y activa de las mismas (Gläser 1990: 72). La competencia activa se refiere a la capacidad de los hablantes de producir y reconocer géneros, como puede ser una carta o un correo electrónico informal. La competencia pasiva se refiere a la capacidad de los receptores de reconocer ejemplares prototípicos de un género, reconocer sus posibles variantes, etc., si bien carecen de competencia activa para producirlos. Existen géneros de especialidad, como una patente, en los que los hablantes no pertenecientes a la comunidad de discurso específica no poseen la competencia activa ni pasiva de los mismos (Gamero 2001: 53). Dentro del conjunto de géneros digitales se pueden encontrar claras diferencias en el caso de la competencia genérica, ya que el conocimiento de las convenciones textuales de gran parte de usuarios de Internet puede ser activo, como en el género *blog* o el correo electrónico, o pasivo, como en el caso de los sitios web corporativos, institucionales o portales, que requieren de los conocimientos expertos de un redactor profesional (véase 8.1.4).

---

<sup>186</sup> Fernández Sánchez siguiendo a la autor también lo sitúa en 50% (Fernández Sánchez 2005: 71), aunque señala que otros investigadores lo sitúan en el 70% (Hoffmann 1991, cf. Fernández Sánchez 2005).



La vertiente didáctica de nuestra investigación se asiente en que los traductores han de adquirir una competencia activa de géneros de especialidad, en la que se incluye el conocimiento de las convenciones de género (García Izquierdo 2002; Ezpeleta 2005). Las convenciones de géneros específicos requieren de una sensibilización y aprendizaje sistemáticos que permitan a un hablante desarrollar de forma consciente su actuación textual con respecto a un género determinado (Gläser 1990: 27). Ésta puede verse favorecida, sobre todo con finalidad traductiva, por medio de análisis contrastivos como la presente tesis (Fernández Sánchez 2004: 69). En caso contrario los futuros traductores, sin una formación específica en las particularidades discursivas y textuales de un género web, pueden presentar un déficit en la competencia activa necesaria. Este tipo de deficiencia presenta un paralelismo con el concepto de «déficit de tipología textual» propuesto por Hatim y Mason (1997):

Within a given language and across languages, the various forms of a given type may not be equally available to all users – a factor we may refer to as **text type deficit**<sup>187</sup> (Hatim y Mason 1997: 133).

La competencia pasiva también desempeña un papel de importancia en los sitios web, ya que el reconocimiento de las convenciones tanto del emisor como del receptor tiene un lugar esencial en la comunicación web (Vaughan y Dillon 2006). Desde el punto de vista cognitivo de la usabilidad del texto web, Nielsen y Tahir, (2002) señalan la labor facilitadora de las convenciones en esta comunicación:

The less you have to wonder about the meaning of shared features, the more you can think about the new features. (Nielsen y Tahir 2002: 54)

Los autores repiten aquí la prescripción más recurrente en guías de usabilidad y de estilo web: seguir las convenciones ya establecidas en los sitios web a varios

---

<sup>187</sup> Énfasis del autor.

niveles<sup>188</sup> (Nielsen y Loranger 2006: 78; Jeney 2007: 3; McAlpine 2001: 220; Adkisson 2002; etc.).

d) **Evolución y flexibilidad** de las convenciones. La evolución y flexibilidad de las convenciones supone reconocer su variabilidad tanto en el espacio, como en el tiempo (Fernández Sánchez 2004: 70-76). Según Göpferich (1995), la variabilidad en el espacio se refiere al carácter culturalmente específico que ya apuntamos de las convenciones (Reiss y Vermeer 1984). De mayor interés en nuestro estudio son las tres posibilidades de variabilidad en el tiempo (Göpferich 1995):

- Cambios en la mentalidad imperante dentro de una sociocultura por cuanto al tema u objeto de estudio.
- Los cambios pueden surgir por una alteración en el problema de coordinación (véase 5.5.1.1) que una convención contribuía a resolver.
- Las interferencias debidas a errores de traducción pueden suponer la paulatina asimilación de lo que inicialmente se trataba de un calco o préstamo.

En los géneros digitales existen gran cantidad de convenciones a todos los niveles, y se espera que más convenciones sigan surgiendo, sobre todo en géneros como los corporativos o institucionales (Shepherd y Watters 1999). Por consiguiente, hemos de prestar atención a la aparición y desarrollo de convenciones digitales (Vaughan y Dillon 2006), ya que la evolución de los cibergéneros supone un proceso en continua expansión y evolución (Kennedy y Shepherd 2005). De ahí surge la necesidad de realizar nuestro estudio de forma sincrónica: a medida que cambien

---

<sup>188</sup> Los manuales de desarrollo, localización y estilo web no suelen mencionar expresamente el concepto «convenciones de género», a pesar de que sus estudios se centran en géneros determinados (Nielsen y Tahir 2002), o incluyen capítulos para diferentes géneros, supragéneros web como por ejemplo sitios web comerciales, informativo o de entretenimiento en la guía de estilo web de Jeney (2007).

las funcionalidades de la web, seguirán surgiendo nuevas convenciones, tales como la aparición de formularios interactivos en las páginas de contacto<sup>189</sup>.

e) Convenciones en la **identificación y traducción** de los géneros. Las convenciones de género contienen un valor especial por ser señas en la identificación de un género determinado. Reiss y Vermeer (1984: 189) otorgan una función triple a las mismas: (1) son señas que permiten reconocer un género textual, (2), activan las expectativas del receptor por cuanto al texto, (3) son señales que coordinan la comprensión del texto. Las convenciones de género son de importancia tanto en la producción textual como en su recepción (Nord 1997b: 53), y por ende, en la traducción (Kussmaul 1997; Nord 1997a, 1997b, 2003; Göpferich 1995:168). Reiss señala que, de cara a la traducción, es de interés la decisión de mantener, sustituir o introducir las convenciones de la lengua meta. Más en concreto, la adaptación a las convenciones de la cultura de llegada no es automática en su visión funcionalista, sino que depende de la finalidad de la traducción y las normas de la cultura de llegada (Reiss y Vermeer 1984: 194). En el contexto de la localización se pretende que los TM sean recibidos como originales en las lenguas meta, el tipo de traducción se situaría en una *traducción comunicativa* (ibid: 194), en el que se incluyan las convenciones del género en la lengua y cultura de llegada por cuanto a patrones lingüísticos y estructurales, que podríamos aproximar a la *traducción instrumental equifuncional* de Nord (véase 4.2.1.4).

### **5.5.1.3. Convenciones y normas**

Las convenciones suponen un acuerdo tácito entre los participantes en el proceso comunicativo y por tanto suponen «unwritten rules» (Lewis 1969: 159). Frente al margen de actuación más o menos amplio de las convenciones, las normas son de

---

<sup>189</sup> Presentes en el 31,4% de los sitios web del corpus original (véase 9.1.2.2).

obligado cumplimiento (Toury 1980; 1998; Chesterman 1993) y suponen una posible sanción por su inobservancia. Toury define el concepto de norma como:

the translation of general values or ideas shared by a group -- as to what is conventionally right and wrong, adequate and inadequate -- into performance instructions appropriate for and applicable to particular situations, specifying **what is prescribed and forbidden**, as well as what is tolerated and permitted in a certain behavioral dimension<sup>190</sup> (Toury 1998: 16).

La no conformidad con las convenciones de género puede por su parte inhibir el funcionamiento de la comunicación, pero no conlleva sanción alguna. Ambas tienen en común que «son culturalmente específicas y no pueden transferirse sin más de una lengua a otra ni tampoco de una cultura lingüística a otra» (Nord 2003: 27). Si ponemos como ejemplo el caso de los sitios web corporativos, el incumplimiento de una convención, como sería la inclusión del menú de navegación a la derecha de la pantalla, podría dificultar la finalidad comunicativa, aunque no impedirlo (Vaughan y Dillon 2006). La infracción de una norma ortográfica, como los errores encontrados en sitios web localizados (véase 11.4.4), entraña por un lado el entorpecimiento del proceso comunicativo y, por otro, la posible asociación de la empresa emisora con una falta de rigor y profesionalidad que repercute negativamente en la relación comercial empresa-cliente (Jeney 2007).

### 5.5.2. La función textual

La importancia de las convenciones se conjuga también con el hecho de que el emisor o emisores<sup>191</sup> del texto poseen una intención tras su elaboración (Beaudegrande y Dressler 1981: 7). La función textual se presenta como un rasgo extratextual esencial en la caracterización de géneros (Swales 2004; Bhatia 1993; Göpferich 1995; etc.). Ésta es una cualidad pragmática que se asigna al texto por

---

<sup>190</sup> Énfasis propio.

<sup>191</sup> Como se comentó anteriormente, el autor suele ser múltiple y anónimo en nuestro género. En otros géneros web como un blog o página personal la identificación del emisor es de esencial importancia y define el mismo.

parte del receptor en una situación particular, que no algo inherente en el mismo (Nord 1997: 49).

La identificación de la función o conjunto de funciones textuales en el proceso comunicativo y en el proceso traductor es fundamental, siendo la multifuncionalidad una de las características inherentes a un gran número de géneros (Hatim y Mason 1990; Gamero 2001: 55; Nord 1997: 49)<sup>192</sup>. Esta multifuncionalidad se amplifica en los géneros digitales y se cristaliza en textos que pueden presentar dos funciones, una primaria y otra secundaria<sup>193</sup> (Hatim y Mason 1990; Gamero 2001: 37). Como ejemplo, los sitios web corporativos se pueden disgregar entre los que poseen una función primaria expositiva con una función secundaria exhortativa, [www.ono.es](http://www.ono.es), o los sitios web corporativos de comercio *online*, [www.elcorteinglés.es](http://www.elcorteinglés.es), en los que la función textual o foco tipotextual primario sería exhortativo y el secundario expositivo (Bolaños 2002)<sup>194</sup>; la primera para la descripción de productos y procesos de compra, y la segunda para que el usuario realice la transacción. Esta diferencia en el conjunto de funciones textuales supone la existencia de dos subgéneros web diferenciados. La división entre función primaria y secundaria se podría tomar también en el sentido de Göpferich (1995), que al clasificar los textos técnicos subraya cómo la función secundaria se produce cuando un receptor usa el texto para un fin distinto al que fue escrito, aunque en nuestra opinión este particular no sería significativo en los textos web.

En el presente estudio emplearemos las funciones textuales desde el enfoque pragmático-semiótico de Hatim y Mason (1990, 1997), que los autores denominan focos tipotextuales (véase, 4.2.1.3). Como ya repasamos, los autores distinguen

---

<sup>192</sup> Nord expresa este aspecto con las siguientes palabras: «When I refer to “function”, in this context [el texto], I mean “function or set of functions” because texts are rarely intended for one function only. Various functions usually form a hierarchy of functions, subfunctions, etc. » (Nord 1997: 9).

<sup>193</sup> No es de extrañar encontrar dos géneros o subgéneros en los que se encuentran dos funciones pero éstas están invertidas.

<sup>194</sup> La autora se centra en el subgénero sitio web corporativo de empresas de informática y analiza exclusivamente las páginas de inicio, que en el presente estudio consideramos un subtexto o bloque comunicativo dentro del género global.

entre la función *argumentativa*, *la expositiva* y *la exhortativa* (Hatim y Mason 1990: 153-158).

En el contexto de los sitios web cobra gran importancia la relación entre las convenciones de género y las funciones o focos tipotextuales que aparecen. La propia naturaleza del medio hace necesario que el traductor de textos web conozca las convenciones textuales que se han desarrollado en la lengua meta y las plasme en la traducción. De otra manera, un texto web localizado que no se ajuste a las convenciones del género corre el riesgo de resultar no funcional (Vaughan y Dillon 2006). Este resultado sería contrario al objetivo del proceso de localización consistente en recrear la función o funciones que se le asignan al texto. Como nos indica Pym (2004a) los textos localizados no han de representar TO alguno, sino un texto funcional en la cultura o *locale* de destino, «the localized document is not called to represent any OT translation but a functional text in the target locale or culture» .

### **5.5.3. Elementos de la situación comunicativa**

Los géneros digitales se encuentran en una situación comunicativa determinada que se encuadra por su emisor y receptor, como por el campo, el modo y el tenor digitales. Cada uno de estos factores es primordial en el proceso de traducción y está relacionado con la producción y la recepción del mensaje. Sin embargo, hemos de señalar que no se pueden emplear como criterios únicos en la caracterización de géneros textuales (Hurtado Albir 2001), si bien útiles en la identificación de posibles subgéneros de un mismo género<sup>195</sup>.

---

<sup>195</sup> El campo se puede utilizar como elemento clasificador de subgéneros dentro del género sitio web corporativo en investigaciones sobre sitios web corporativos de empresas de informática (Bolaños 2002; Bolaños *et al.* 2005), sitios web de empresas azulejeras (Rennau 2004), sitios web de bancos (Fernández 2003), etc.

Los sitios web corporativos se enmarcan dentro de un nuevo modelo de comunicación que siguiendo las investigaciones de Anna Janoschka denominamos Modelo de Comunicación de Masas Interactivo<sup>196</sup> (Janoschka 2003: 96 y ss.). La investigadora apunta que una de las particularidades de esta situación comunicativa es la combinación de elementos de comunicaciones de masas (MC) y de comunicación interpersonal (IPC) presentes en este nuevo medio. Esta situación queda reflejada en la siguiente gráfica, donde queda patente cómo el mayor flujo comunicativo del mensaje *online* (MO) se produce unidireccionalmente desde el emisor al receptor en esta comunicación de masas, representado en azul. Los elementos blancos en la gráfica se refieren a la comunicación interpersonal, que se puede realizar entre grupos o entre usuarios, decantándose por denominarlos «usuarios» (U). Las flechas blancas representan el flujo bidireccional que se puede entablar a través de la interactividad que proporciona el nuevo medio, Internet, entre los participantes en la situación comunicativa.

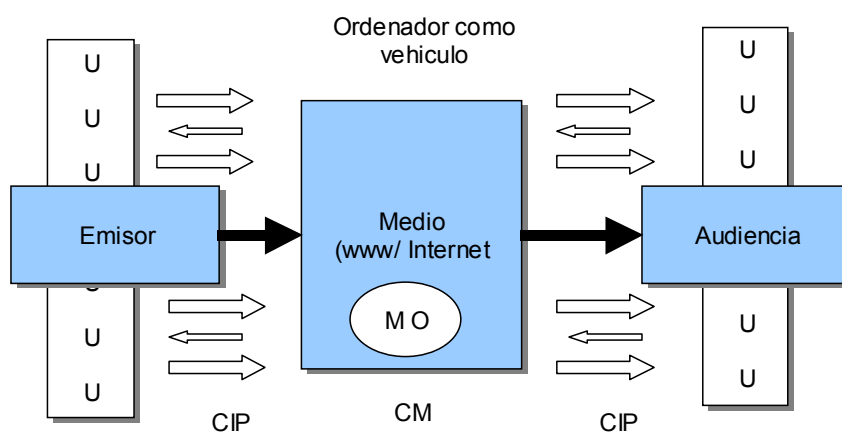


Figura 28. Modelo de comunicación de masas interactivo. Adaptado de Janoschka (2003).

La identificación del emisor y el receptor dentro de los sitios web corporativos es un rasgo de caracterización definitorio en el propuesta de situación presentada anteriormente (véase 5.1.2.1 ). En el género *sitio web corporativo*, la comunicación se entabla entre una empresa y sus clientes, aunque los resultados obtenidos en la

<sup>196</sup> Traducción propia del inglés *Interactive Mass Communication*.

presente tesis indican que esta no es la única situación comunicativa existente (véase 8.1.2). En los sitios web institucionales, el intercambio comunicativo se entabla entre una institución con un miembro de la sociedad en la que se sitúa, como sería un ciudadano en un sitio web de un ayuntamiento.

En los textos web el receptor es heterogéneo, siendo difícil identificar un perfil único<sup>197</sup> (Krug 2005). En este contexto, el receptor suele considerarse un «usuario», posición no exenta de críticas (véase 2.4). Los posibles receptores en los textos web se pueden dividir en cuatro tipos de audiencia que distinguimos siguiendo a Jeney (2007: 26-27):

- **Audiencia primaria:** la audiencia a la que se dirige implícitamente el sitio web.
- **Audiencia supervisora:** que consistiría del iniciador o el cliente que encarga el texto web y a los que responde directamente el redactor: el director de proyecto, el ejecutivo del departamento de finanzas, o el gerente .
- **Audiencia periférica:** los usuarios que no se espera que visiten un sitio web pero que no obstante hay que tener en cuenta.
- **Audiencia secundaria:** se compone de las personas que han de actuar según el texto escrito, como los miembros del equipo legal tras la escritura de los términos legales.

Los traductores-localizadores se encontrarían dentro de la audiencia periférica, puesto que como señala Muñoz Martín (1995: 45), «los mediadores no son nunca los destinatarios de los textos o discursos originales y se constituyen en la audiencia de los mismos (es decir, no son destinatarios sino receptores fortuitos)».

---

<sup>197</sup> Existen no obstante autores que reivindican que es inapropiado dirigirse a una audiencia global y es necesario ajustar el texto web a un destinatario medio (Price y Price 2002: 4).



El medio en el que se encuadra la situación comunicativa ha revolucionado el proceso de comunicación entre la empresa y clientes en lo que respecta a la publicidad, ventas, distribución o atención al cliente. La última publicación de la LISA (Lommel 2007: 5) señala cómo el carácter revolucionador de Internet en el contexto comunicativo las empresas se debe a (1) su bajo coste, (2) su alcance mundial, y (3) su estatus como plataforma mundial de marketing y publicidad, de ventas, de distribución y (4) de asistencia al cliente:

- **A low-cost, worldwide advertising and marketing platform.** By setting up a web site, even small companies can advertise themselves, their products and their services to an audience that potentially extends around the world. In addition, they can use their web sites to directly gather information on customer needs in real time.
- **A low-cost, worldwide sales platform.** The rise of e-commerce solutions enables companies to sell goods direct from their web sites. Payment is typically made by credit card or by debiting an existing offline account, and alternative payment means suitable for specific markets have emerged.
- **A low-cost, worldwide distribution platform.** Any product that can be made available as electronic data (e.g. software, books and other text-based information, music, films and so on) can also be directly distributed via the web. Companies and their customers now expect materials to be available immediately at any time of the day anywhere in the world. This level of expectation is made possible only through the Internet. In the case of physical products such as hardware, cars, etc., worldwide logistics providers now offer efficient international distribution. Shenzhen (China) is now closer in many ways to Los Angeles (U.S.) than Detroit (U.S.).
- **A low-cost, worldwide support platform.** Companies serving a newly acquired international customer base can minimize support issues by providing registered customers with information and contacts via their web sites. Moving support to web sites can drastically reduce support costs (Lommel 2007: 5).

Como último aspecto dentro de la situación comunicativa se sitúa la «funcionalidad», que por su relevancia en las investigaciones sobre géneros digitales las revisaremos por separado (véase 5.5.6).

#### **5.5.4. Influencia del contexto sociocultural**

Los géneros se encuentran inmersos en un entorno sociocultural específico y se desarrollan por una necesidad comunicativa que se repite en su seno (Hatim y

Mason 1990; Berenkotter y Hucking 1995). Consecuentemente, su estudio ha de situarse dentro del contexto sociocultural que los haya generado. Si bien la mayoría de géneros web son comunes al conjunto de países industrializados, existe la posibilidad de que ciertos géneros sean exclusivos de una cultura determinada, como el *haiku* japonés. El contexto sociocultural interviene de tres maneras determinantes:

- a) Es determinante en las convenciones propias de cada género. Como indicamos anteriormente, las convenciones de género dependen de cada lengua y cultura (Nord 1997b: 54; Gamero 2001), por lo que pueden no ser extrapolables a otras.
- b) Es determinante a la hora de crear nuevos géneros, puesto que si en un contexto sociocultural no existe una necesidad comunicativa, no se desarrollará un género.
- c) Es determinante en la evolución de cada género, puesto que se adapta a la evolución y cambios en el seno de la sociedad.

Los géneros digitales empezaron a desarrollarse en el entorno sociocultural estadounidense (Schäler 2002), de ahí que se sobrevalore a menudo la influencia de este contexto sociocultural y lingüístico en los mismos (Bolaños 2002; Schäler 2002). No obstante, y de forma paralela a otros géneros impresos, existen diferencias entre géneros digitales web en culturas diferentes (Tercedor 2005). Esta tendencia se acentúa en los géneros más recurrentes como los sitios web corporativos. En este aspecto, existen menores diferencias interculturales en las convenciones de género en los programas de software en comparación con los géneros web debidas a la menor producción original de los primeros en España. Defendemos que existen características de los géneros web que se han convencionalizado dependiendo de cada cultura y lengua dada la abundante producción original de este tipo de textos, una de las hipótesis de partida del

presente estudio y principio básico que pone en valor este parámetro (Gamero 2001: 55).

Como ejemplo, en los sitios web anglosajones nos encontramos con la convencionalización de las *FAQs* o «Preguntas más Frecuente», que se han constituido en un género secundario<sup>198</sup> digital en sí (Santini 2005a). Sin embargo, hemos constatado que en España este género secundario se encuentra únicamente en el 3,4% de los sitios web del corpus. En este contexto sociocultural, los sitios web españoles presentan mayormente una opción de ayuda o atención al cliente, recayendo la responsabilidad de asistencia en mayor medida sobre el emisor del texto, en lugar de las «Preguntas Más Frecuentes», que corresponden a una opción más individualista de ayudarse por sí mismo (Singh y Pereira 2005).

### **5.5.5. Elementos intratextuales**

Generalmente, los modelos de caracterización del género otorgan una mayor importancia a los factores extratextuales que a los intratextuales. Las convenciones de género se presentan como uno de los aspectos de mayor relevancia de estos estudios en Traductología y éstas afectan exclusivamente a los elementos intratextuales. De esta manera, los elementos convencionales que se activan frente a las posibles alternativas son exclusivamente intratextuales, puesto que cualquier cambio en la función comunicativa o en la situación comunicativa resultará en un género diferente. Un ejemplo claro lo constituye un folleto impreso de una empresa y su sitio web, géneros diferentes no sólo por el medio, sino por la funcionalidad que proporciona y que permite, por ejemplo, búsquedas avanzadas o vistas de productos en 3D<sup>199</sup>.

---

<sup>198</sup> En estos estudios se considera un género digital en sí. No obstante en la presente tesis hemos argumentado que constituyen géneros secundarios ya que dependen de un género «complejo» o primario para su realización (véase 5.1.2).

<sup>199</sup> Ciertos investigadores han señalado, erróneamente a nuestro parecer, que un sitio web corporativo es una variante del género folleto informativo de una empresa (Fernández 2003).

Coincidimos con Gamero (1998, 2001) al señalar que no existe un modelo claro y definitivo de análisis de los rasgos convencionales intratextuales, de ahí que la mayoría de los estudios se centren en uno o dos parámetros de caracterización. En su revisión bibliográfica se presentan entre otros: la superestructura, los actos de habla, la sintaxis, el léxico, el tono o tenor, los títulos, la progresión temática y la cohesión, elementos metacomunicativos y metalingüísticos, la sintaxis, o los elementos no verbales (Gamero 2001: 58-59). No obstante, no todos los parámetros revierten la misma importancia en todos los géneros o pueden ser igualmente relevantes para el traductor. El estudio de la autora concluye que la superestructura aparece como el elemento más convencionalizado y cohesionador de géneros técnicos. Su importancia se debe además a que los rasgos convencionales pueden ser diferentes entre bloques o secciones de un mismo género. Podemos ilustrar este punto mediante el tono en los sitios web corporativos españoles, en donde descubrimos una primacía del tratamiento de «usted» frente al más informal «tú». Así, ciertos bloques dirigidos al conjunto de usuarios, como «contacto» o «regístrese», reflejan la preferencia general por un tratamiento formal sobre el informal. No obstante, en un avance de nuestros resultados, el estudio empírico del tono discursivo en los diferentes bloques muestra que en el bloque «trabajo», dirigido a aquellos interesados a formar parte de la empresa en cuestión, el tratamiento de «tú» sería mayoritario (véase 8.1.2.3).

Si nos centramos en los estudios sobre los géneros web corporativos, en éstos se han incorporado los siguientes parámetros:

- Elementos metadiscursivos (Alejo González 2005)
- Macroestructura (Bolaños 2002b)
- Correlación imágenes-texto (Tercedor y Jiménez 2007; Bolaños *et al.* 2005)

- Léxico (Bolaños *et al.* 2005)
- Coloquialismos (Bolaños *et al.* 2005)
- Verbos (Bolaños *et al.* 2005; Renau 2004)
- Pronombres (Bolaños *et al.* 2005; Renau 2004)
- Adjetivos (Bolaños *et al.* 2005; Renau 2004)
- Tono (Bolaños *et al.* 2005)
- Sintaxis (Pollard 2006)
- Estructura retórica (Ruiz 2003)

Siguiendo a Gamero (1998, 2001) y Göpferich (1995), nuestro análisis intratextual comenzará con la superestructura y posteriormente se analizarán los diferentes parámetros de análisis intratextual que seleccionaremos.

### **5.5.6. Funcionalidad**

La funcionalidad supone uno de los elementos que caracteriza a los géneros digitales y la base de clasificación de la primera investigación de impacto sobre los mismos (Crownston y Williams 1997). Dada la convencionalización de las distintas posibilidades de interacción (Shepherd y Watters 1999), la hemos incorporado a nuestro modelo de caracterización del género digital. Por ejemplo, los usuarios esperan que los resultados de una búsqueda incorporen hipervínculos hacia la página en cuestión o que al pulsar sobre una dirección de correo electrónico se abra directamente un correo en blanco<sup>200</sup>.

La funcionalidad del texto web se refiere a las posibilidades de interacción que ofrece el nuevo medio, Internet (Shepherd y Watters 1999) y se ha caracterizado como el principal motor que genera nuevos géneros (Shepherd y Watters 1998). Siguiendo la propuesta de Janoschka, (2003: 115), definiremos la funcionalidad

---

<sup>200</sup> Que se consigue por medio de la inclusión de la etiqueta html <a mailto:>.

descomponiéndola entre la **interacción** y la **interactividad**<sup>201</sup>. Por un lado, la interacción se refiere a la comunicación entre el emisor del mensaje y el receptor (Janoschka 2003: 115). En los sitios web corporativos la **interacción** se produce a través correos electrónicos, chats de ayuda o en foros, con la posibilidad de recibir una respuesta o *feedback*, que a su vez definimos como la acción comunicativa iniciada por el emisor para recibir una reacción específica del receptor. Por su parte, la **interactividad** se define por la comunicación con el medio (Janoschka 2003; Jucker 2003). Esta comunicación se refiere al uso de funciones de búsqueda, la lectura selectiva de la página, la activación de hipervínculos, la activación de menús de navegación o cualquier otra forma de participación activa por parte del usuario (Janoschka 2003: 116).

La funcionalidad que posibilita la interacción máquina-usuario se ha investigado en profundidad desde los albores de los textos digitales, destacando los trabajos de Apple con el libro *Machintosh Human Interface Guidelines* (Apple Computers Inc. 1992). En un principio, se intentó recrear los textos impresos en la WWW, pero ciertas particularidades del medio digital como la lectura en pantalla o la funcionalidad hicieron que los cambios a los géneros existentes fueran de tal magnitud que surgieron nuevos géneros, tales como los periódicos en línea a partir de los periódicos tradicionales (Yates y Orlikowski 1992). La funcionalidad supuso así el principal elemento que contribuyó a la evolución de los géneros digitales y su distanciamiento de los impresos (Crownston y Williams 1999).

Los investigadores Shepherd y Watters (1999) realizaron una investigación pionera sobre la funcionalidad de páginas de inicio en la que descubrieron seis géneros digitales primarios<sup>202</sup>. Su propuesta de caracterización plantea definir los géneros

---

<sup>201</sup> Ambos términos son sinónimos en español, pero no obstante hemos decidido seguir la propuesta terminológica de la autora por economía y claridad frente a otras posibilidades más complejas como «interacción maquina-usuario» presente en los pioneros estudios de Apple (1992) y «interacción emisor-usuario».

<sup>202</sup> De objetivos limitados puesto que la investigación se incluyó sólo 97 páginas de inicio.

web por medio de tres únicos rasgos: (1) el contenido, (2) la forma y (3) la funcionalidad. A pesar de su gran repercusión en estudios posteriores, en nuestra opinión resulta demasiado vaga puesto que no definen, por ejemplo, que constituye el «contenido».

Cybergenre	Content	Form	Functionality
<b>Homepage</b>	Information about company-person-institution	Introduction Hierarchical images Animated images	Browsing e-mail
<b>Brochure</b>	Product and services	Shallow hierarchy High impact visual	Browsing e-mail
<b>Resource</b>	Subject-specific information	Hierarchical Images Video Audio	Browsing E-mail Search Discussion Interaction
<b>Catalogue</b>	Products and services	Hierarchical Images	Browsing e-mail ordering and inquiry search on-line ordering on-line enquire
<b>Search engine</b>	Categories of sites Urls	Query box List of site Virtual document	Browse Search

Figura 29. Resumen de contenido forma y funcionalidad de los seis cibergéneros descritos por Shepherd y Watters (1999).

Dentro de esta clasificación, los sitios web corporativos se situarían dentro del supragénero *homepage*, que los autores acertadamente dividen entre los personales, los institucionales y los corporativos. En líneas generales, mencionan que los sitios web presentan como funcionalidad la posibilidad de navegar mediante enlaces e interacción, sobre todo por medio del correo electrónico. Varias investigaciones previas han aplicado la propuesta de Shepherd y Watters a diferentes géneros digitales (Schmitt-Isler 2000; Bolaños 2002; Alamán 2003; Fernández 2003; Rennau 2004), manteniendo las posibles variables asociadas a la funcionalidad presentadas por estos investigadores, como búsqueda, correo electrónico, discusión, o compra en línea.

En estudios más recientes queda patente la evolución de la funcionalidad en géneros digitales, llegando a incorporarla como variable contrastiva en estudios empíricos. Por ejemplo, en la investigación empírica del género web secundario «presentación Flash» de Paolillo *et al.* (2007) se recoge un modelo empírico y cuantitativo de caracterización de la funcionalidad por medio de la presencia de los siguientes elementos: (1) botones interactivos, (2) menús, (3) acciones con el teclado o ratón, (4) controles de pausa y *play*, (5) posibilidad de incluir o excluir subtítulos en presentaciones Flash.

En resumen, la importancia de la funcionalidad en el entorno web es tal que se considera el motor de evolución, aparición y convencionalización de los géneros digitales (Kennedy y Shepherd 2005), por lo que pasamos a describir su importancia clave a la hora de clasificarlos.

#### **5.5.6.1. La funcionalidad como criterio clasificador de géneros digitales en Shepherd y Watters (1998)**

Shepherd y Watters (1998) realizaron una clasificación de géneros digitales partiendo de la existencia de un «género digital» inicial, en la que se considera a la funcionalidad como el parámetro esencial tras su convencionalización y evolución. Esta clasificación se menciona y apoya en gran número de investigaciones (Santini 2007a, 2005a; Rosso 2005; Rennau 2004; Posteguillo 2003; Rehm 2002; Bolaños 2002; Bolaños *et al.* 2005; Fernández 2003; *inter alia*). Los investigadores dividen el universo textual digital entre los géneros impresos que dieron el salto a la web sin adaptación, los géneros «existentes» o *extant*, y aquellos géneros que se crearon específicamente en la web, los «nuevos» o *novel* que no existían en el mundo impreso. Este aspecto es de interés puesto que en las diferentes perspectivas sobre el género se suelen crear a partir de los existentes. Todorov ya indicó en 1990:



Where do genres come from? Quite simply from other genres. A new genre is always the transformation of an earlier one, or of several: by inversion, by displacement, by combination (Todorov 1990: 15).

En Internet, un género como el *videoblog* no existía con anterioridad, ya que depende de la funcionalidad que proporciona el nuevo medio. La Web muestra cómo la evolución el medio puede repercutir en la creación de nuevos géneros dependientes del mismo.

A continuación, los investigadores dividen los géneros «existentes» entre en géneros «reproducidos» o *replicated*, aquellos textos impresos que simplemente se «cuelgan» en la red, tales como los artículos de investigación. Estos géneros se adaptan a la red, puesto que requieren de conversión a formato .html o .pdf, pero textualmente no presentan características diferenciadas, sólo el medio de transmisión es diferente. En general, estos géneros no suelen funcionar en la web puesto que no se adaptan a las particularidades del medio (Nielsen y Tahir 2002), de ahí que su análisis no tenga interés en nuestra investigación. Por otro lado, los géneros «variantes» o *variant* se basan en los géneros existentes como los periódicos, pero incluyen ciertas funcionalidades dadas por el nuevo medio. En la actualidad, la funcionalidad de la mayoría de periódicos en línea los convierte en un género *novel*.

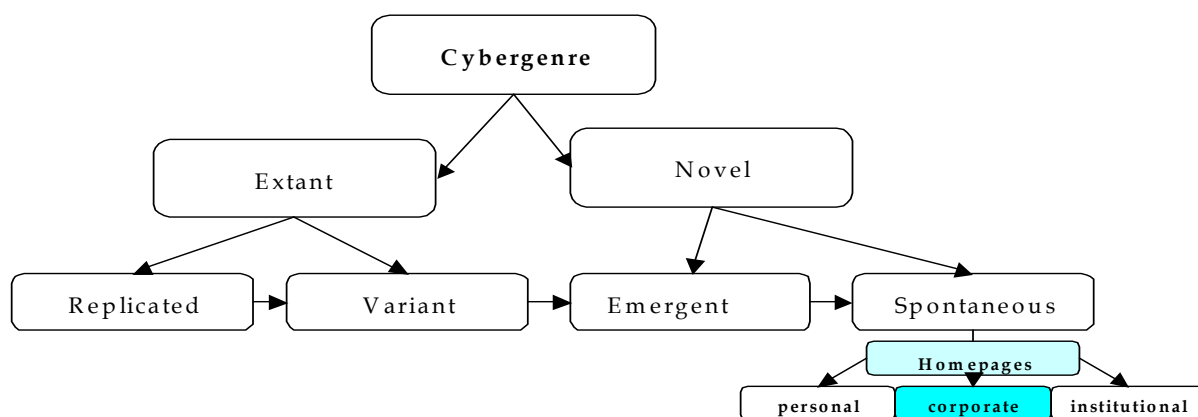


Figura 30. Los cibergéneros. Adaptado de Shepherd y Watters, (1998), y ubicación del objetivo de estudio, combinado con Kennedy y Shepherd (2005).

Los géneros *novel* son aquellos que no existen en otro medio; según Shepherd y Watters es el nivel de funcionalidad el que los distingue:

the primary distinguishing feature of these genres is a level of functionality that makes it fully dependent for its existence on the new medium (Shepherd y Watters 1998: 100).

Estos géneros se pueden dividir entre los «emergentes», *emergent* y los «espontáneos», *spontaneous*. Los géneros «emergentes» surgen de la evolución de un género *variant* y se establecen como género diferenciado. Un ejemplo son los periódicos en línea, que muestran una estructura y funcionalidad en constante evolución con la incorporación gradual de elementos multimedia como vídeos, archivos mp3, gráficos Flash animados, foros de discusión, chats, etc. Por su parte, los géneros *spontaneous* son géneros que no aparecen en ningún otro medio, como las *homepage* o los sitios web de las radios que mezclan un portal de noticias con la posibilidad de escuchar programas de radio o de descargar *podcast*.

En su clasificación añaden otro aspecto esencial en el análisis de géneros web: la persistencia o dinamicidad de la información. En su terminología, los géneros pueden ser «persistentes», *persistent*, o «virtuales», *virtual*. La diferencia principal se erradica en que todo género «persistente» posee versiones definitivas y estables de las páginas almacenadas, siendo posible en cualquier momento acceder a estas versiones estables por distintos usuarios. Los géneros «virtuales» se basan en la creación del contenido por medio de servidores web dinámicos con lenguajes de programación como Java, ActiveX, CGI, Pearl Shockwave, etc.<sup>203</sup>, por lo que usuarios diferentes pueden obtener páginas web distintas aunque accedan a un mismo momento a un sitio web determinado. Esta diferenciación entre géneros se corresponde en localización con los sitios web «dinámicos» o «estáticos» (véase 2.8).

---

<sup>203</sup> Que producen páginas web con extensiones .jsp, .aspx o .do etc. Para una lista exhaustiva de los diferentes formatos web, consúltese [www.fileinfo.net](http://www.fileinfo.net).

Los autores admiten que su clasificación presenta un grado e indeterminación debido a la naturaleza cambiante de estos géneros. Este aspecto está en sincronía con la flexibilidad y dinamicidad que presentan las investigaciones teóricas sobre género presentadas anteriormente (Swales 2004; Gamero 2001; Hurtado Albir 2001). En la WWW la constante evolución de la tecnología supondrá el rápido desarrollo y adaptación de los géneros digitales.

Si aplicamos este marco al género objeto de estudio, éste constituye un género **novel** y **espontáneo**, puesto que no existente en el mundo impreso, y a su vez no fue el resultado de la evolución de un único género anterior. Este género complejo incorpora bloques textuales que pueden constituir géneros textuales **existentes**, como catálogos de productos en formato .pdf, o **noveles**, como los formularios de contacto, y sus realizaciones pueden ser **persistentes** o **virtuales**.

Si comparamos nuestra caracterización con otras investigaciones previas sobre este género, principalmente desde una perspectiva lingüística, hemos de defender que supera la simple evolución de ciertos géneros impresos con los que se ha comparado, tales como el folleto publicitario (Fernández 2003) o el catálogo comercial (Renau 2004: 43). La funcionalidad y textualidad de este nuevo género superaría una simple evolución de estos géneros impresos, apareciendo un nuevo género que depende enteramente de su realización en un medio específico, Internet.

## **5.6. Los géneros web: una constante evolución**

Si nos centramos en definir qué constituye un género web en general, consideramos que la simplicidad de Santini (2005a: 2) es apropiada: «Web genres are genres that are used on the web», aunque debemos matizarla para excluir

aquellos géneros producidos para su impresión: «Web genres are genres that are used **exclusively** on the web».

En varias investigaciones encontramos un consenso más o menos amplio sobre los diferentes géneros que han aparecido en la web, siempre teniendo en cuenta que no existen unos límites definidos a los que asignar cada género web existente (Shepherd y Watters 1999), y que los mismos, al igual que todos los géneros (Swales 2004), se encuentran en constante evolución (Santini 2005a; Vaughan y Dillon 2006)<sup>204</sup>. Como mencionamos anteriormente, Miller (1984) señalaba que los géneros no se prestan a su clasificación debido a su continua evolución, para la investigadora «genres change, evolve and decay» (1984: 153). Esta corriente norteamericana defiende la dinamicidad de los géneros (Miller 1984; Yates 1989; Bazerman 1988, 1991, 1995), aludida en gran cantidad de estudios posteriores (Devitt 1991; Fortanet *et al.* 1999).

Como mencionamos en el apartado anterior, son tres las categorías de género que se encuentran consolidadas en la web: *Reproduced-Replicated*, *Novel-Spontaneous*, *Adapted-Variant*. En cada una de estas clasificaciones se encuentran ciertos géneros que han migrado a la web, como los periódicos o los manuales en línea, géneros adaptados puesto que se adaptan a las funcionalidades que proporciona la web (Crownston y Williams 1999). Entre los géneros que se han identificado (Shepherd y Watters 1998 1999; Crownston y Williams 1997, 1999; Santini 2005a) se encuentran:

- *Homepages*. Un género novel entre los que se encuentran las personales (Dillon y Grushowski 2000), las institucionales (Rehm 2002), las web de organizaciones no gubernamentales (Nielsen y Loranger 2007), y las corporativas

<sup>204</sup> Swales indica la consecuente necesidad de revisar los estudios sobre géneros específicos debido a esta constante evolución (Swales 2004: 31).

- Los folletos
- Los sitios de recursos
- Los catálogos web
- Los motores de búsqueda
- Los juegos
- Los contactos personales (Santini 2005a)
- *Quick information delivery* (Santini 2005a)
- Las *Hotlists*
- Las presentaciones Flash (Paolillo *et al.* 2007)
- Los blog (Miller y Shepherd 2004)

Para una revisión de los diferentes estudios en géneros digitales desde la perspectiva lingüística, como los web de informática, de universidades, la tesis de la investigadora Renau (2004: 39-44) ofrece una detallada revisión.

Si nos adentramos en el sector de Internet, observamos que es común clasificar los géneros web de acuerdo a un criterio de clasificación único, como sería el campo para la empresa más importante en la investigación de mercado en Internet, Nielsen-Netratings. Ésta presenta una categorización cuyo propósito básico es el muestreo de visitantes en cada género web.

Genre	Impressions
Email	21,937,663,000
General Community	7,354,519,000
Portals & Search Engines	3,842,844,000
General/National News	1,882,931,000
Finance	1,679,209,000
Sports & Recreation	1,127,700,000
Local/Regional	1,100,824,000
Entertainment	988,483,000
Games	729,501,000
Weather	544,230,000
Other	3,680,868,000
Total	44,868,772,000

Figura 31. Clasificación de géneros según la empresa Nielsen-Netratings<sup>205</sup>.

<sup>205</sup> [http://www.adrelevance.com/intelligence/intel\\_dataglance.jsp?flash=true&sr=36808](http://www.adrelevance.com/intelligence/intel_dataglance.jsp?flash=true&sr=36808)

El criterio de clasificación que conduce a esta división entre géneros web, basado en el campo temático, supone la separación entre portales y motores de búsqueda, deportes, negocios, servicios de noticias, correos electrónicos, etc. Esta clasificación se asemeja a ciertas propuestas de tipologías textuales que han sido objeto de críticas por incluir el campo temático como criterio único de clasificación. Por consiguiente, podemos concluir que no son de interés desde un punto de vista traductológico (Hatim y Mason 1997: 128).

### 5.7. Esquema del modelo de caracterización de género

Como resumen de nuestro marco teórico de caracterización del género, a continuación presentamos el siguiente esquema que supone un resumen del modelo de género de Gamero (2001: 60) adaptado a nuestros objetivos, al que añadimos la funcionalidad dentro del contexto comunicativo como segundo determinante en el establecimiento de convenciones.

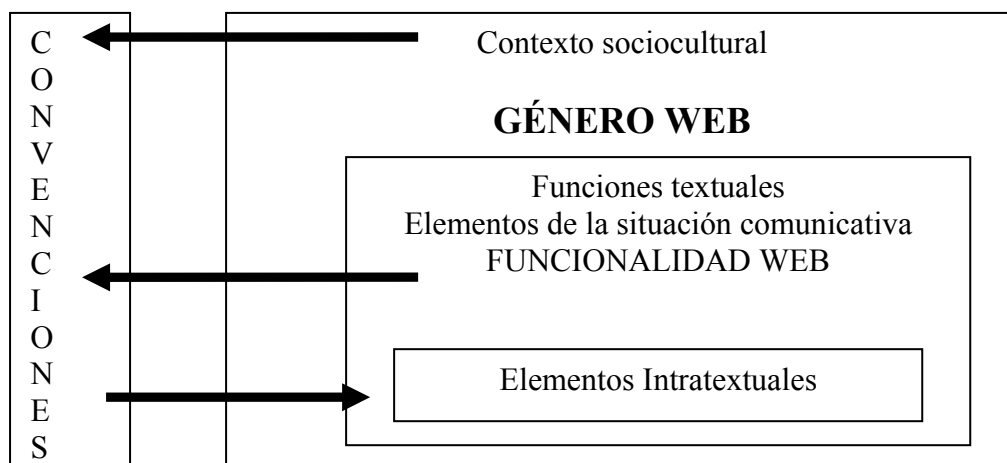


Figura 32. Esquema de modelo de caracterización de géneros digitales, adaptado de Gamero (2001: 60).

Según este modelo, el género se enmarca dentro de un contexto sociocultural determinado que determina, junto con la funcionalidad propia del medio, las convenciones de un género determinado; a su vez, éstas sólo pueden mostrarse en los elementos intratextuales. Los otros dos elementos extratextuales que caracterizan el género son las funciones textuales y el resto de elementos de la situación comunicativa.

### **5.8. Propuesta de caracterización general del género *sitio web corporativo***

Si tenemos en cuenta y aplicamos los factores de caracterización adaptados de Gamero (2001), podemos definir el género de sitio web corporativo tradicional como:

*Un prototipo textual convencionalizado que se repite en una sociedad concreta con el fin de que la comunicación se lleve a cabo de la forma más eficaz posible; como prototipo, este género presenta una función expositiva con una función secundaria exhortativa, un emisor que se asocia con una empresa y un receptor indeterminado que necesita información o interactuar con la misma. La comunicación se realiza a través un medio, Internet, con la funcionalidad que ésta proporciona, y presenta un modo escrito que en su realización se acerca al hablado, pudiendo estar presente el medio oral; este género posee unos bloques o secciones y unas características textuales, tipológicas, tipográficas, y funcionales que se encuentran convencionalizadas en diferente grado. Los receptores son capaces de reconocer estas características convencionalizadas a partir de la forma externa, el medio y la situación de uso, pero para adquirir una competencia activa en este género, por ejemplo como traductores, es*

*necesario realizar un esfuerzo consciente de aprendizaje y tener conocimientos básicos de los lenguajes de programación en los que se encuentran inmersos los textos.*

## **5.9. Los géneros digitales desde la perspectiva prescriptiva de las guías de estilo web**

Como hemos indicado previamente los géneros web presentan unas características específicas, fruto de su dependencia de un medio determinado que ha condicionado su evolución y distanciamiento de los géneros textuales impresos (Crystal 2001; Posteguillo 2003; Shepherd y Watters, 1998). Si situamos esta evolución en su contexto sociocultural y socioprofesional, podemos observar cómo este proceso de producción textual, originalmente encomendado a redactores técnicos o periodistas, fue dando cabida a ciertas peculiaridades específicas dadas por el medio (Díaz Noci y Salaverría 2003; Price y Price 2000). A pesar de la naturaleza eminentemente descriptiva de nuestra investigación, hemos de reconocer la importancia de revisar la perspectiva prescriptiva que aportan las guías de estilo especializadas<sup>206</sup>. Éstas constituyen un compendio de las distintas convenciones establecidas que ejercen una influencia en la labor de redactores y traductores-localizadores. Su existencia e influencia se situaría dentro de las restricciones socioculturales en el contexto de producción de los textos localizados (Baker, 1999). En estas guías se recoge el carácter diferenciado de los textos web, de ahí que constituyan una primera aproximación a las diferencias entre textos impresos y digitales. En éstas se señala unánimemente la importancia de una formación específica en este tipo de redacción, e incluso se incluyen capítulos específicos sobre la adecuación de géneros impresos para su distribución por Internet (Bly 2002).

---

<sup>206</sup> A pesar de que su naturaleza prescriptiva, las guías de estilo en sí son eminentemente descriptivas puesto que simplemente recogen las distintas recomendaciones en sus páginas (Nobs 2005).



Dentro de esta perspectiva prescriptiva, en adición a las guías de estilo web específicas hemos de señalar el indudable valor de los manuales de estilo técnicos, ya que los manuales de estilo generales, como los periodísticos<sup>207</sup>, no se ajustan necesariamente a las peculiaridades de los textos web<sup>208</sup>. Si intentamos demarcar el objetivo de los manuales de estilo técnicos especializados, éste queda patente en la siguiente cita de Martínez de Sousa (2007):

Un manual de estilo tiene la función de elegir, para una publicación o una entidad determinada, las grafías que autoriza y los usos que rechaza, así como parte de documentación genérica que pueda resultar útil en función del tipo de escrito de que se trate, con las explicaciones oportunas cuando sea necesario [...] (Martínez de Sousa 2007: 28).

En el contexto de la localización se ha descrito la función de los manuales de estilo como:

To create documents appropriate for the end-user, for meeting company and country standards, and for maintaining geographic and cultural suitability (Lingo Systems 2004: 29).

Toda investigación que profundice en el análisis textual y genérico en este ámbito ha de analizar los criterios generales que ofrecen este tipo de publicaciones. Su impacto en el desarrollo del proceso de redacción original y el proceso de localización es innegable. No obstante, hemos de señalar que las recomendaciones de estos manuales no se reflejan automáticamente en el conjunto de género presentes en la Internet (Gamero 2001).

Pasamos a resumir los criterios recurrentes prescritos para la redacción de géneros web, que hemos recopilado de las distintas publicaciones españolas (Díaz Noci y

---

<sup>207</sup> Como el Libro de estilo del País, el libro de estilo de la Agencia EFE o el libro de estilo de El Mundo.

<sup>208</sup> No obstante, en las publicaciones estadounidenses se señala a *Chicago Manual of Style* como la referencia básica en la redacción en general (Lynch y Horton 2002).

Salaverría 2003; Vilamor 2001), como anglosajonas (Horton 1990; Anderson *et al.* 1998; Kilian 1999; Sammons 1999; Hale y Scanlon 1999; McAlpine 2001; Bly 2002; Garrand 2001; Hammerich y Harrison 2001; Osborne 2002; Price y Price 2002; Lynch y Horton 2002; McGovern *et al.* 2002; Veloso 2005; Mill 2005; Jeney 2007). Igualmente, nuestra síntesis incluye los capítulos dedicados exclusivamente a la redacción web en los manuales de usabilidad (Nielsen 2000; Nielsen y Tahir 2002; Nielsen y Loranger 2006; Krug 2005; Brinck *et al.* 2002), o las publicaciones sobre la localización de software y de contenido web (Esselink 2000; Yunker 2003; 4; Lingo Systems 2004: 60-66).

### 5.9.1. Consideraciones generales

La totalidad de publicaciones consultadas recogen ciertas características básicas del texto en géneros web: ha de ser breve, conciso, informal, claro, accesible, específico y de calidad<sup>209</sup>. Estas características generales se corresponden en gran medida con las recogidas por Duque (2000: 21-22) para la redacción científico-técnica: claridad, concisión, precisión y objetividad<sup>210</sup>.

- La **brevedad** del texto web se asocia con las peculiaridades de lectura en pantalla, ya que el usuario suele escanearlo hasta que encuentra una información que le interesa (Nielsen y Loranger 2006). El texto breve reduce el nivel de ruido (Krug 2005), con lo que se disminuye la carga cognitiva y se resalta el contenido específico de cada página. En general, se apunta que el texto debe ser hasta el 50% menos que el impreso, evitando redundancias innecesarias (Jeney 2007). La brevedad se relaciona con las restricciones impuestas por el visionado en pantalla,

---

<sup>209</sup> Vilamor (2001: 101-110) resume las características del texto web en: diferenciado, informal, directo, breve, conciso, interactivo, persuasivo, claro, preciso, original, variado, atractivo, rítmico, con colorido y universal.

<sup>210</sup> El investigador realiza una síntesis de las investigaciones previas. Asimismo, el autor une estas características a otras de menor frecuencia como la transparencia, variedad, organización del contenido, legibilidad, accesibilidad, audiencia, corrección gramatical, uso de figuras estilísticas y elementos léxico gramaticales (Duque 2000: 21-22).

ya que el texto ha de ajustarse a estos límites. Con el fin de reducir el texto, Price y Price (2002: 88) recomiendan no incluir contenido que sea parte del conocimiento compartido de la comunidad de discurso, excluir palabras innecesarias o palabras forma como *really* o *just*, así como usar pronombres demostrativos cuando sea posible. Como ejemplo óptimo en la eficacia de la brevedad textual, se presenta el motor de búsqueda Google (McAlpine 2001).

- La **concisión** está en el mismo plano que la brevedad, y constituye uno de los retos de la redacción web (Mill 2005). Esta característica se asocia con la reducción de los futuros costes de la localización del texto (Lingo Systems 2004).
- La **claridad** ha de ser una característica en todo texto web (Bly 2002).
- El **carácter internacional** y global de la audiencia web se recoge en varios manuales. En el contexto anglosajón se recomienda escribir para una audiencia internacional que no domina perfectamente el inglés (Kilian 1999; McAlpine 2001; Lynch y Horton 2002). Dado este carácter internacional de la comunicación por Internet, se recomienda no incluir metáforas o juegos de palabras específicos que requieran de una situación de recepción sociocultural determinada.

### 5.9.2. Nivel léxico-semántico

En el plano léxico-semántico las recomendaciones se centran en los siguientes aspectos:

- Utilizar **términos simples** que sean comprensibles y familiares para los receptores, sin que sean sofisticados o de baja frecuencia de uso (Nielsen y Loranger 2006; McAlpine 2001; Price y Price 2002; etc.). En el contexto estadounidense se recomienda el uso de palabras de origen anglosajonas frente a las de origen latino, tal como *get-obtain*, *terminate-end* (Price y Price 2002; Bly 2002). Se recomienda el uso de verbos de un solo lexema frente a lexías compuestas por verbos de soporte como «decidir» frente a «tomar una decisión», «hacer una prueba» frente a «probar»

o «usar» frente a «hacer uso de». Asimismo se recomienda la concreción en el uso preposicional, con formas más breves y simples como *although* frente a *despite the fact*, o *but* en lugar de *notwithstanding* (Mill 2005).

- Ser **coherente** y consistente en el uso terminológico (Yunker 2003; Jeney 2007; Bly 2002; Price y Price 2002), evitando redundancias y el uso excesivo de sinónimos para un mismo concepto.
- Evitar el uso de **siglas o abreviaturas** y en caso de que usen, deletrearlas (Mill 2005)<sup>211</sup>. Éstas son apropiados en apartados técnicos de los sitios web o partes que se dirijan a especialistas (Garrand 2001).
- Evitar el uso de **tecnicismos**, incluyendo la mayoría de los términos informáticos como html, 56K o *plug-in* a no ser que se dirija expresamente a expertos o usuarios semiespecializados en informática (Garrand 2001).
- No emplear **jergas**<sup>212</sup> (Nielsen y Loranger 2006: 262), **clichés**, o **generalizaciones** (Jeney 2007).
- Evitar **usos ambiguos** (Yunker 2003; Kilian 1999), ya que el lenguaje ambiguo aumenta la carga cognitiva en el procesamiento textual (Morkes y Nielsen 1999). Prestar igualmente atención a las palabras polisémicas (McAlpine 2001: 48). No introducir expresiones idiomáticas o americanismos (McAlpine 2001: 45; Yunker 2003).
- Evitar el **sarcasmo**, los **juegos de palabras** con un uso apropiado del **humor** (Mill 2005: 26), ya que incrementan los problemas de comprensión y con difíciles de traducir a otros idiomas (McAlpine 2002: 41). Estos aspectos se han de evitar en los menús de navegación puesto que requieren una claridad absoluta (Price y Price 2002).
- Prestar atención a la **variación dialectal** en el caso de escribir para una audiencia internacional (Kilian 1999). Como ejemplo, se presenta la diferencia entre billón y su diferencia entre el inglés americano y británico (McAlpine 2001: 56).

---

<sup>211</sup> Es posible usar las siglas y abreviaturas de mayor frecuencia como DVD, PC, etc. (Nielsen y Tahir 2002).

<sup>212</sup> Mill (2005) recomienda el uso de programas informáticos de lenguaje controlado como Bullfighter TM, que controla estos aspectos. [www.deloitte.com](http://www.deloitte.com).

- Incluir URL o **direcciones de sitios** web que sean simples, lógicos y que sea posible pronunciar, leer, recordar, predecir y escribirlas con facilidad (Price y Price 2002: 165). No deben incluir símbolos tipográficos que confundirían al usuario; en inglés el nombre de la empresa Barns & Nobles puede suponer la dirección [www.barnsandnobles.com](http://www.barnsandnobles.com) o [www.barns&nobles.com](http://www.barns&nobles.com).

### 5.9.3. Nivel morfosintáctico

- La **sintaxis** ha de ser simple, evitando la profusión de oraciones subordinadas (Horton 1990). Una sintaxis compleja supone una mayor carga cognitiva en el procesamiento del texto web (Price y Price 2002).
- Las **oraciones** han de ser simples, cortas y completas, con una idea por oración. En el caso del inglés se suele apuntar una longitud máxima de 15 a 25 palabras<sup>213</sup> (Mill 2005; Yunker 2003).
- Se recomienda el uso de la **voz activa** (McAlpine 2001: 47; Jeney 2007; etc.), puesto que los usuarios encuentran el texto más accesible y resulta más fácil de procesar (Bly 2002).

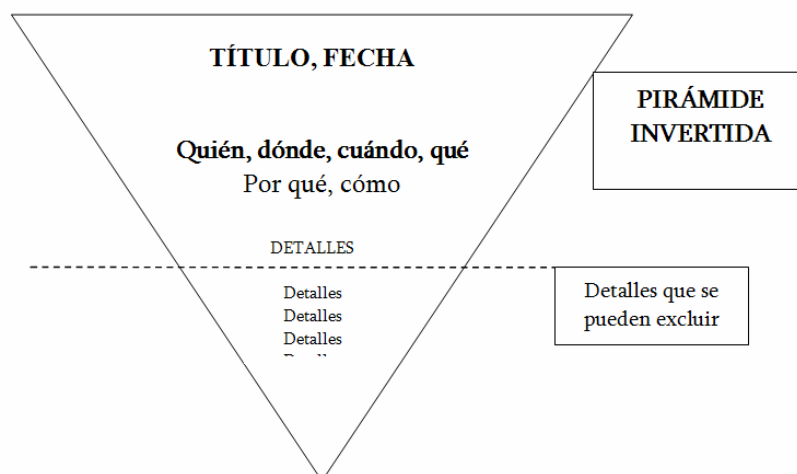


Figura 33. Estilo periodístico textual en forma de pirámide invertida. Adaptado de Jeney (2006: 148).

---

<sup>213</sup> El español técnico se caracteriza por una longitud oracional mayor (López Rodríguez 2001).

- Los usuarios web suelen escanear el texto en las primeras fases de aproximación (Nielsen 2000; Nielsen y Loranger 2006). El texto se ajusta a estas peculiaridades de lectura a través con el uso de **oraciones cortas** y párrafos que comiencen con la idea principal. Esto se consigue igualmente con la estructura periodística de párrafos en forma de **pirámide invertida**, en la que el párrafo comienza con la conclusión, sigue con la información importante y termina con los antecedentes (McAlpine 2001: 44; Nielsen y Loranger 2006: 269). Según Mill, «the effectiveness of the inverted pyramid style of writing on the Web is because the reader [...] can quickly consume the key information at the beginning» (2005: 11-12).
- Emplear **formas verbales frente a sus nominalizaciones** cuando sea necesario (Price y Price 2002). Evitar los verbos frasales en inglés (McAlpine 2001: 48), ya que son de difícil comprensión para una audiencia intencional.
- Usar principalmente **formas nominales y adjetivales** descriptivas, a la vez que se evita la profusión de adjetivos o adverbios (Garrand 2001). Emplear **tiempos verbales simples** cuando se posible (McAlpine 2001: 48).
- Emplear conectores simples que articulen el texto entre párrafos (Mill 2005).
- Se recomienda ejercer cuidado en el uso de los **pronombres anafóricos, catafóricos, y deícticos**, puesto que es imposible asumir qué ha leído previamente el usuario (Price y Price 2002). Es necesario de igual modo controlar cualquier referencia **endofórica, exofórica, deíctico, pronombres demostrativos**, etc., sin una referencia clara, ya que así se contribuye a producir textos exentos de ambigüedad.
- Ser específico y claro con las **fechas y números** por las diferencias entre lenguas y países (McAlpine 2001: 56).

#### 5.9.4. Nivel estilístico

- Aparece un **tono conversacional**, similar al oral (Mill 2002: 26), en el que el texto se dirige directamente al usuario y se incorpora en él (Bly 2002: 143; Nielsen y Tahir

2006: 262). La prescripción de usar este estilo directo se recoge en la totalidad de publicaciones; «The Internet work best as a series of two way conversation» (Price y Price 2002: 31), «write as if you are talking to an individual, not a collective group of anonymous web surfers» (Garrand 2001: 32), «use a friendly, lively, individual tone. Remember, you are conversing with one person at a time», (McAlpine 2001: 37). Las publicaciones anglosajonas aconsejan dirigirse directamente al usuario en un tono informal y directo<sup>214</sup>.

- Incluir una **actitud positiva** (Mill 2005). Ésta se produce mediante la exclusión de dobles negativas en inglés, usando preferentemente expresiones positivas como «poco» frente a «no mucho» o «no el mismo» frente a «diferente». En el plano verbal no combinar expresiones de negación con verbos con carga semántica negativa como *avoid, deny, exclude, lack, prevent, prohibit*, o con preposiciones como *against* o *without* (Price y Price 2002). El uso excesivo de expresiones negativas supone un mayor esfuerzo cognitivo de comprensión (Sammons 1999).
- Usar un **estilo informativo publicitario**, ya que es el más productivo en Internet. Los experimentos en usabilidad de Jacob Nielsen (1999) concluyeron con que el estilo informativo es 25% más efectivo que el publicitario, ya que facilita su lectura a los usuarios, que a su vez recuerdan mejor el contenido. El estilo web más efectivo ha de alejarse de las exageraciones publicitarias y de marketing, «write your copy like you are sharing information, rather than selling» (Bly 2002: 12).
- No emplear **lenguaje sexista** (Price y Price 2002: 14; McGovern *et al.* 2002).
- El texto no debe de ser excesivamente simple o estereotipado en ciertas circunstancias (Yunker 2003; Nielsen y Loranger 2006). La guía de estilo web de Kilian (1999) se hace eco de la inexistencia de guías de estilo durante el auge de Internet en los noventa, de ahí que proliferaran ciertas expresiones estereotipadas

---

<sup>214</sup> En este aspecto aparecen contradicciones entre las diferentes guías de estilo web consultadas sobre como ajustar el estilo a esta peculiaridad (Bly 2002).

a evitar como, «site under construction», «welcome to my site», «click here!» o «check this out!»<sup>215</sup> (Kilian 1999: 11).

- Prestar atención al uso de una **puntuación adecuada** (Yunker 2003), ya que ésta puede causar problemas en su localización y en su posible traducción automática. Además, ciertos elementos de puntuación como el punto y coma o los dos puntos pueden no ser captados por los usuarios en la lectura en pantalla (Price y Price 2002), por lo que cuando sea posible se recomienda la división en dos oraciones.
- Incluir la información repetitiva o numérica en **tablas y gráficos**, dado que mejoran su el proceso de visualización e identificación de información relevante (Price y Price 2002; Bly 2002: 73-78).
- La información ha de dividirse en **párrafos de menor extensión** que los impresos (Kilian 1999). Cada párrafo ha de contener una idea o ser la unidad de propósito. Los párrafos cortos facilitan el escaneado, de ahí que la extensión general recomendada sea de 3 a 5 oraciones (Mill 2005; Price y Price 2002: 91), con alrededor de 100 palabras máximo (Kilian 1999). El concepto más repetido por cuanto a esta división textual es el término inglés «chunking»<sup>216</sup>.
- No obstante, se prescribe emplear **listas** cuando un párrafo contenga cuatro o más elementos (McAlpine 2001: 5), si bien se recomienda no abusar de listados en un sitio web (Nielsen y Loranger 2006: 280). Las listas se han de alinear a la izquierda, pudiendo realizarse en negrita el término clave inicial y sin comenzar con un determinante o demostrativo inicial. No se ha de incluir más de tres niveles dentro de un listado (Price y Price 2002).
- Emplear **títulos y encabezamientos** concisos y descriptivos, un aspecto primordial para captar la atención del usuario (Mill 2005: 3). Los títulos han de contener un máximo de sesenta caracteres (Bly 2002). Éstos deben resumir de manera concisa el contenido de la página en el caso de los títulos, o la sección en caso de un

---

<sup>215</sup> Todas estas expresiones se han calcado y extendido en español, «sitio en construcción» o «haz click aquí».

<sup>216</sup> Jeney (2007: 44) señala que esta división y organización textual se puede realizar de las siguiente formas: alfabéticamente, numéricamente, cronológicamente, de lo general a lo específico, según la prioridad, por la relación causa-efecto, problema-solución, geográficamente, por tamaños, por precios, por tamaño, etc.



encabezamiento, y deben estar seguidos de doble espacio (Kilian 1999). Las investigaciones con seguimiento de movimientos oculares muestran que los usuarios sólo leen las primeras palabras en los resultados de motores de búsqueda. Como se mencionó anteriormente, en estos subtextos también se recomienda evitar el humor, juegos de palabras o sarcasmo.

### 5.9.5. Nivel tipográfico

Los aspectos tipográficos en el texto web han sido el objetivo de gran cantidad de publicaciones<sup>217</sup> e investigaciones en el ámbito de la usabilidad. Las recomendaciones sobre tipografía han de considerarse sugerencias, ya que es imposible controlar exactamente la presentación en pantalla dadas las preferencias del usuario, el navegador, las opciones preseleccionadas por el usuario, los tipos de letra presentes en el ordenador, etc. (Lynch y Horton 2002; Kilian 1999: 4). Los elementos tipográficos desempeñan un papel de mayor importancia en la labor del redactor original que en la del localizador web, puesto que se suele respetar la tipografía existente durante el proceso de traducción. De hecho, los textos web se localizan exclusivamente con herramientas informáticas que respetan automáticamente la tipografía original. Podemos resumir las recomendaciones en este nivel en:

- Emplear un **conjunto limitado de tipos de letra** (Nielsen y Loranger 2006: 235), sin que incluya tipos exclusivos o no extendidos (McAlpine 2001: 61; Lingo Systems 2004: 62).
- Emplear **fuentes diseñadas para la lectura en pantalla** como Verdana, Georgia, New York y Trebuchet, ya que estas fuentes son *sans-serif* frente a las *serif* (Kilian 1999). Los estudios de usabilidad señalan la mejora en la lectura en los textos web en *sans-serif* (Nielsen 2000).

---

<sup>217</sup> Para una bibliografía sobre la totalidad de los aspectos del diseño web se puede consultar la lista bibliográfica de la University of Minesota en la página web:  
<http://www.d.umn.edu/is/support/Training/Online/webdesign/books.html>

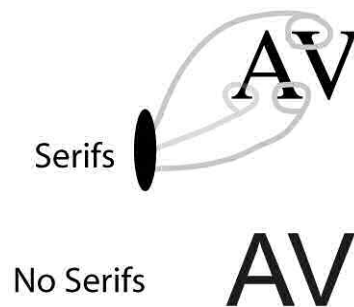


Figura 34. Diferencia entre la fuente *sans-serif* frente a la *serif*.

- El **tamaño de letra** óptimo es 12, siendo posible el 10 (Nielsen y Loranger 2006). No obstante, los usuarios pueden cambiar el tamaño dependiendo de sus preferencias o la resolución en pantalla.
- Usar preferentemente la **negrita** para dar énfasis, ya que el texto en itálicas dificulta su lectura (Nielsen y Loranger 2006), y por otro lado el subrayado en Internet se ha convencionalizado para indicar la presencia de un hiperenlace (McAlpine 2001: 61; Nilsen 2000). No se recomienda usar texto en mayúsculas para dar énfasis, ya que ralentiza la lectura en un 10% (Nielsen y Loranger 2002: 238). Del mismo modo, el uso exclusivo de mayúsculas se asocia con el grito y sería inapropiado en sitios web profesionales (ibid).
- No usar más de cuatro **colores** en el diseño (Nielsen y Loranger 2006: 235). El negro sobre blanco sigue siendo el más legible, junto al azul sobre blanco. Si se mezcla texto en color, prestar atención a las convenciones tipográficas: color morado para los enlaces visitados y azul para los no visitados (McGovern *et al.* 2002).
- Los **fondos** deben de ser simples sin gráficos complicados, ya que aparte de dificultar la legibilidad (Nielsen y Loranger 2006: 240), pueden transmitir una falta de profesionalidad en los sitios web comerciales.
- Se requiere la presencia de gran cantidad de **espacio en blanco** alrededor del texto para mejorar su legibilidad (Kilian 1999).
- Es primordial evitar el **texto en movimiento y animaciones**. En la actualidad se ha convencionalizado su uso como recurso publicitario y, por consiguiente, se suele

ignorar por los usuarios (Price y Price 249). Este efecto se denomina *banner blindness* (Janoschka 2003), el efecto por el que los usuarios ignoran cualquier texto en un recuadro con formato y tipografía diferenciada al asociarlo con textos publicitarios.

- No distribuir el texto en **columnas** que comprendan la totalidad de la pantalla (Kilian 1999: 31).
- No se recomienda usar imágenes, espacios o puntos para organizar el texto o como decoración, ya que presenta problemas de accesibilidad para personas con discapacidad (Price y Price 2002).

### 5.9.6. Nivel estructural

- Los **enlaces** son de esencial importancia con lo que es necesario elegirlos cuidadosamente (McAlpine 2001: 33). Dado que se encuentran resaltados por medio de un color diferenciado y por encontrarse subrayados, muchos usuarios los escanean antes de leer la página. Esta tipografía que indica la presencia de un enlace constituye una de las convenciones básicas del texto web, de ahí que tengan que ser implícitos y nunca han de estar precedidos de una expresión redundante como *click here!* o «haz clic» o *Follow this link to...* (Price y Price 2002; Mill 2005: 44). En caso de que se incluyan en un párrafo, es necesario ponerlo al principio o al fin del mismo, puesto que el usuario tiende a ignorar el texto posterior (Price y Price 2002: 124). Los enlaces han de colocarse preferentemente en texto, ya que si se colocan en imágenes, estos no son accesibles para personas con discapacidad visual.
- El texto web necesita ser **dividido** en caso de superar los largos de pantalla, aunque es primordial que cada página sea clara y tenga sentido por sí misma (McAlpine 2001: 13). La información más importante de cada página ha de presentarse en las dos primeras oraciones (Nielsen y Loranger 2006) y cada página, como unidad de contenido, ha de poder resumirse en una sola oración (McAlpine 2001: 26). No

obstante, las divisiones del texto no han de ser arbitrarias (Lynch y Horton 2002). Se recomienda no dividir innecesariamente textos coherentes y lineales (Nielsen 1999). Si embargo, una excepción serían las noticias largas que pueden dividirse en dos o tres páginas (Price y Price 2002: 147).

### **5.9.7. Nivel icónico-visual**

La naturaleza del texto web es caracteriza por su carácter verbo-icónico (Tercedor 2005). Dentro de esta retórica visual, el uso metafórico de imágenes se convierte en una característica intrínseca de la comunicación por Internet (Posteguillo 2003: 63). Las guías de estilo web recomiendan restringir el uso de imágenes, ya que el uso innecesario de imágenes proliferó desde los albores de Internet. Tal como apunta Jeney (2006), se desaconseja inclusión excesiva de imágenes y animaciones en estos textos ya que:

Web writers are frequently tempted to replace textual information with images, but it important to remember that image are usually static and nonhierarchical in nature, whereas the human mind prefers structure and hierarchy [...] Only use the images so that they fit the purpose of the website (Jeney 2007: 28-29).

A continuación presentamos una síntesis de las recomendaciones en esta área:

- Los estudios de usabilidad web del instituto Poynter.org<sup>218</sup>, realizados con seguimiento de movimientos oculares, confirman que los usuarios se dirigen primero al texto en lugar de a las imágenes (Nielsen y Loranger 2006).
- Los elementos de la retórica visual son tan importantes como el texto, pero no deben ser sustitutos del mismo (Jeney 2007: 29). Se recomienda usar imágenes sólo cuando se ajusta en propósito o función del sitio (ibid: 28).
- A pesar de que las imágenes e iconos son importantes para los usuarios, el texto constituye el núcleo de todo sitio web. En palabras de Crawford Kilian: «The

---

<sup>218</sup> <http://eyetrack.poynter.org/>

graphic elements of a web site are especially important to them [users], but text remain the core of most websites» (Kilian 1999: 15).

- El texto en imágenes no se puede buscar, una de las características interactivas del nuevo medio (Nielsen y Loranger 2006: 247). En caso de que se usen, es indispensable incluir el texto alternativo por medio del atributo «alt», que resume su contenido y función. El texto en imágenes requiere una inversión en localización más importante, por lo que en las publicaciones en este campo se desaconseja incrustar texto en imágenes (Lingo Systems 2004: 60).
- El texto en imágenes presenta problemas desde el ámbito de la usabilidad, ya que los lectores de páginas web no pueden acceder al mismo. De igual modo, las imágenes no se pueden ampliar o reducir para personas con problemas de agudeza visual, una posible opción con el texto web (Nielsen y Loranger 2006: 247).
- Las imágenes se recomiendan por tanto para los botones y textos con tamaños de letra pequeños (Nielsen y Loranger 2006: 247).
- Evitar el uso de animaciones Flash siempre que sea posible comunicar el mismo mensaje con texto e imágenes estáticas (Bly 2002). En caso de usar el formato Flash, se recomienda (1) evitar las introducciones innecesarias, (2) proporcionar una navegación e interactividad lógicas, (3) no abusar de animaciones, (4) reducir al mínimo el uso de elementos sonoros, y (5) ajustar el tamaño para que no presenten problemas a los usuarios sin banda ancha (Yunker 2003).
- Las guías recomiendan incluir el atributo «Alt» con una descripción textual de toda imagen en el sitio. Ésta ha de proporcionar una descripción clara y sucinta de las mismas, incluyendo cualquier imagen o icono que sirva en la navegación (Garrand 2001). McAlpine (2001: 101-113) recoge en su guía un compendio de recomendaciones para la descripción de imágenes:
  - Conocer la función de cada imagen como paso previo a su descripción.
  - Dejar el atributo vacío para imágenes decorativas o que se usen para organizar la distribución del texto, « alt="" ». En caso de que no incluir el atributo «Alt» en una imagen que sirva como decoración, el lector de páginas web para

personas con discapacidad visual repetirá la palabra «imagen» cada vez que se encuentre una imagen sin esta etiqueta.

- Si la imagen es una línea para terminar un párrafo o lista, escribir «end of section», «end of page», etc.
- Terminar el texto en alt con un punto, puesto que en caso contrario el lector unirá este texto con el siguiente sin pausa, Ej. alt="Logotipo de la empresa."
- Si la imagen es parte del contenido del texto, como una gráfica con texto incluido, describirlo en su totalidad.
- Incluir texto de calidad puesto que los buscadores e indexadores de sitios web emplean este texto para describir e indexar la página.

### 5.9.8. Edición del texto

En general, en las publicaciones en este ámbito existe un consenso amplio sobre la relajación en la aplicación de las normas que rigen todo texto escrito (Posteguillo 2003). No obstante, el proceso de edición de todo texto es primordial y se recomienda:

- Una revisión textual completa de calidad (Bly 2002: 56).
- En general, se encuentran más errores en los textos web que en los impresos, y éstos se asocian con la falta de profesionalidad de la empresa o la institución emisora del texto: «in a commercial context, writing errors are seen as evidence of weak of sloppy thinking. This in turn reflects upon the company or institution that owns the site» (Jeney 2007: 46).
- Se aconseja no confiar en exceso en el corrector automático de los procesadores de texto y revisar el texto en su totalidad (Nielsen y Loranger 2006; Jeney 2007).
- Dentro de los errores básicos del texto web Jeney señala (2007):
  - Confusión en el uso de pronombres
  - Confusión en el uso de expresiones idiomáticas

- Uso erróneo de terminología especializada
  - Problemas de puntuación
  - Coloquialismos
  - Estructura oracional confusa
  - Confusión entre palabras específicas a los usuarios de la lengua inglesa:  
*accept-except, conscience-conscious*
- 
- Los errores que señala la investigación sobre textos localizados de Noelia Corte (2002) están en consonancia con los presentados anteriormente:
    - Faltas de ortografía
    - Errores gramaticales
    - Terminología errónea
    - Imágenes incongruentes
    - Texto en idioma original en cuadros o mensajes de error

## 6. LA LINGÜÍSTICA DE CORPUS EN LA INVESTIGACIÓN EMPÍRICA EN TRADUCCIÓN

Corpus-based approaches will enable new approaches to internationalization and thus significantly improve the speed, efficiency and accuracy of computer-assisted translation and localization. Using corpora as part of an internationalization strategy implies that they can be manipulated or engineered in some way to be more effective tools for translation and localization (Shreve 2006: 315).

El último paso en el establecimiento del marco teórico pasa por analizar desde una perspectiva teórica la aplicación de técnicas de Lingüística de corpus al marco metodológico que describiremos en el capítulo 7. Desde su introducción en Traductología (Baker 1993), esta metodología ha originado una gran cantidad de investigaciones empíricas dentro su rama descriptiva (Toury 1995; Holmes 1988). En esta encrucijada entre los Estudios Descriptivos de la Traducción y la Lingüística de Corpus se sitúa la metodología que hemos adoptado en la presente tesis. Dentro de la nueva corriente traductológica basada en corpus denominada *Corpus-based Translation Studies*, defendemos siguiendo a Olohan (2004: 3)<sup>219</sup> la pertinencia de no considerarla como nuevo paradigma (Laviosa 2002: 5)<sup>220</sup>, sino más bien como metodología.

---

<sup>219</sup> El autor indica que es más fructífero y beneficioso su consideración como metodología que como una nueva disciplina que a caballo entre ambas ramas de la Lingüística y la Traductología (Olohan 2004: 3).

<sup>220</sup> Sara Laviosa aboga por que se considere un nuevo paradigma y lo denominado «Corpus-Based Translation Studies» (Laviosa 2002: 5).



Ya desde las primeras investigaciones, la Lingüística de corpus se introdujo en Traductología para investigar el producto y el proceso de la traducción (Holmes 1988: 67; Tymozco 1998). En este sentido, pretendemos establecer un marco teórico y metodológico para investigar específicamente el producto y el proceso de la localización, que defendemos será una de las aportaciones de la presente tesis. Hasta la fecha el uso de corpus no se ha integrado plenamente en la tecnología y estrategias de internacionalización de la mayoría de empresas del ramo (Shreve 2006), puesto que se suele asumir que esta metodología requiere gran cantidad de tiempo y recursos. En palabras de los investigadores Biau y Pym (2006: 18)<sup>221</sup> el uso de corpus en localización es «a not cost-beneficial tools when applied to professional translation». No obstante, hemos de defender los beneficios que aportan la perspectiva cuantitativa y descriptiva de las investigaciones traductológicas de corpus (Laviosa 2002; Nord 2003), con su subsiguiente efecto en la práctica y didáctica de los distintos tipos y modalidades de traducción (Hurtado Albir 2001: 505). De ahí que pretendamos demostrar o corroborar que su uso en localización proporcionaría claros beneficios en estos dos ámbitos. En lo referente a la práctica profesional, las mejoras en el proceso de localización se producirían por que las soluciones a problemas de traducción pasarían a no basarse exclusivamente en posibles ejemplos limitados, la intuición del traductor, o el uso exclusivo de una opción única y descontextualizada presentada por memorias de traducción (Laviosa 2002: 9).

## **6.1. La Lingüística de Corpus en Traductología**

La metodología de la Lingüística de Corpus existe en Lingüística desde hace más de tres décadas (Kennedy 1998: 1), si bien su introducción en Traductología no se

---

<sup>221</sup> Los autores erróneamente señalan que una simple búsqueda en Internet puede resolver los problemas de traducción que se presenten, «They address problems that are more easily solved with a quick web search» (Biau y Pym 2006: 18). En este sentido, defendemos que una simple búsqueda o un corpus paralelo muestra exclusivamente una de las posibles soluciones, y no necesariamente la más correcta o convencional (Olohan 2004).

produjo hasta los noventa (Baker 1993). Su evolución ha ejercido una gran influencia en cambios metodológicos y teóricos en Traductología (Laviosa 2002), así como en varias disciplinas afines como la lexicografía, la lingüística educativa, la lingüística computacional, la lingüística contrastiva, la lingüística forense, la lingüística crítica, los estudios literarios, la terminología, la adquisición de una segunda lengua, la estilística, la lingüística histórica, la psicolingüística, sociolingüística o la traducción asistida por ordenador (McEnery y Wilson 1996). En un principio, este impacto se debió a su afán empiricista por investigar el funcionamiento real de la lengua mediante análisis de grandes cantidades de textos producidos de manera natural por hablantes nativos (López Rodríguez 2001: 169), rechazando implícitamente los dualismos *language-parole* de Saussure o *competence-performance* y lengua externalizada-internalizada de Chomsky (Stubbs 1996: 23). Sus principios se basan igualmente en una visión de la lengua como fenómeno social<sup>222</sup> y heterogéneo (Laviosa 2002). El impacto de su introducción en Traductología en conjunción con avances en el campo de la informática es innegable (Tymozco 1998). En la actualidad podemos afirmar que constituye un área consolidada de investigación empírica en el seno de esta disciplina:

Within a wider empirical paradigm, corpus based studies of translations have given rise to a coherent and distinct body of research in both branches of the discipline, pure and applied (Laviosa 2002:1).

Para situar el presente apartado, hemos de señalar que nuestro objetivo no consistirá una revisión exhaustiva del papel que desempeña la Lingüística de corpus en la Lingüística en general<sup>223</sup> (Biber *et al.* 1998; McEnery y Wilson 1996

---

<sup>222</sup> Dentro del contexto del estudio de la lengua dentro de su contexto social iniciado por Firth (1957).

<sup>223</sup> Biber (Biber *et al.* 1998: 4) resume las características de los análisis basados en corpus a través de los siguientes puntos: (1) es empírica, ya que analiza las estructuras reales de uso en textos naturales, (2) utiliza como base de sus análisis colecciones de textos amplias y representativas, llamadas «corpus», (3) emplea ampliamente los ordenadores para el análisis, usando técnicas tanto automáticas como interactivas, (4) se basa en análisis tanto cuantitativos como cualitativos. (5) en su conjunto, estas características producen una amplitud y una fiabilidad en el análisis que no se podría obtener de otra manera.

2001; Kennedy 1998; Meyer 2002), o una revisión de su introducción en Traductología (Kenny 2001; Laviosa 2002; Olohan 2004), sino más bien poner de relieve los puntos de confluencia en los que la Lingüística de Corpus en su orientación no prescriptiva se ha introducido como metodología dentro de las bases teóricas más amplias de los estudios descriptivos en traducción. En este sentido, recogemos la síntesis que ofrece Olohan de los rasgos que ha introducido esta metodología en la Traductología<sup>224</sup> (2004: 16):

- El interés en el estudio descriptivo de las traducciones tal y como existen.
- Un interés por la lengua tal y como aparece en el producto de la traducción, en contraposición al estudio orientado hacia el sistema que se produce en la Lingüística contrastiva. En este sentido, y siguiendo a Kenny (1998), el carácter «contrastivo» de nuestra investigación se centra en la peculiaridad de la lengua de traducción cuyos rasgos se explican debido a «the nature and pressures of the translation process» (Baker 1997: 177), en lugar del sentido de la Lingüística contrastiva, que justifica estos cambios dentro del contexto de los sistemas lingüísticos comparados.
- Un interés por descubrir qué es probable y típico en la traducción, y de igual modo, interpretar aquellos rasgos que sean atípicos.
- Una descripción que combina el análisis de corpus cualitativo y cuantitativo, que a su vez se puede centrar en el léxico, la sintaxis y los rasgos discursivos.
- El uso de esta metodología en diferentes tipos de traducción, como la traducción en contextos socioculturales, modos o tipos diferentes, etc. En nuestro caso, aplicamos esta metodología a una modalidad específica de traducción, la localización web.

---

<sup>224</sup> No obstante, el investigador señala que la unidireccionalidad casi exclusiva desde la Lingüística de Corpus que se refleja por ejemplo, en que en esta disciplina aún se considere al TO y al TM un único texto: «a bilingual parallel corpus is a corpus that contains the same text samples in each of the two languages, in the sense that the samples are translations of each other» (Oakes and McEnery 2000: 1, cf. Olohan 2004: 13).

La introducción de la Lingüística de corpus en Traductología se manifiesta en una gran cantidad de investigaciones empíricas desde diversos ámbitos geográficos y combinaciones lingüísticas<sup>225</sup> (Baker 2004). De hecho, esta introducción se ha llegado a identificar como uno de los principales avances en la misma (MalmKjaer 2003, cf. Olohan 2004: 23). No obstante, su introducción no ha estado libre de críticas. Ciertas investigaciones han señalado las posibles deficiencias de esta metodología. Por ejemplo, la investigadora Tymoczko (1998: 658) defiende el carácter positivo de sus aportaciones pero advierte a los investigadores ante un celo demasiado científico que dependa exclusivamente de datos cuantitativos. En este mismo sentido, Ian Mason, (2001: 71) señala la necesidad de incorporar un marco más amplio que dé cabida a al contexto y co-texto, el género, o los aspectos discursivos en investigaciones que realicen generalizaciones mediante datos cuantitativos, tal como las líneas de concordancia. Nuestra investigación asume y defiende la opinión del investigador, de ahí que en nuestro análisis partamos desde el marco teórico que proporcionan los estudios sobre el género textual.

A continuación, ya que el concepto de corpus se sitúa en el núcleo de nuestra metodología, pasamos a definir este concepto desde una perspectiva multidisciplinar.

## **6.2. El concepto de corpus**

El centro de nuestro estudio lo constituye la recopilación y análisis posterior de un corpus comparable monolingüe (Laviosa 1997: 292) de sitios web desarrollados en

---

<sup>225</sup> Como ejemplo, en España han proliferado la creación de corpus en el seno de ciertas universidades con objetivos traductológicos: TURICOR en la universidad de Málaga, COVALT de la Universidad Jaume I de Castellón (Corpus Valencia de la Literatura Traducida), el GITRAD de géneros jurídicos en la misma institución, el P-Actes contrastivo de la Universidad de León, etc. Consúltese Borja (2008) para un resumen de los distintos corpus desarrollados en España.

España. Dada su importancia, un primer paso ineludible en nuestra labor sería definir teóricamente el concepto de corpus en general.

La concepción de corpus en Traductología proviene de la Lingüística de Corpus, de la que parte una de las definiciones más extendidas; la propuesta por John Sinclair (1991):

A corpus is a collection of naturally-occurring language text, chosen to characterize a state or variety of a language (Sinclair 1991: 171).

Esta definición remarca la necesidad esencial de que los textos sean naturales<sup>226</sup> (Stubbs 1996: 4), principio que entronca con los principios *firthianos* del estudio de la lengua en su uso (Firth 1935, 1957). Esta concepción de naturalidad se encuentra asimismo en la definición propuesta por Biber (1998), que pone de relieve la importancia de que la colección textual sea extensa y se recopile según unos criterios explicitables claramente establecidos: «a large and principled collection of natural texts» (Biber 1998: 4).

Con el advenimiento de la informática a finales de los 80, se introdujo como requisito el formato electrónico para analizar los corpus de forma automática o semiautomática (Baker 1995: 226). El investigador Leech (1992: 106) introduce el concepto de corpus informatizado de la siguiente manera: «On the face of it, a computer corpus is an unexciting phenomenon; a helluva lot of text, stored on a computer», reseñando así cómo su tamaño aumentó exponencialmente<sup>227</sup>. En la actualidad, el formato digital es una necesidad reflejada en las propias definiciones de corpus (Johansson 1998; McEnery y Wilson 1996: 24; Baker 1995: 226)<sup>228</sup>. El

---

<sup>226</sup> Para Stubbs el matiz fundamental es que no haya intervenido el lingüista en su creación: «an instance of language in use, either spoken or written: a piece of language behavior which has occurred naturally, without the intervention of the linguist» (Stubbs 1996: 4).

<sup>227</sup> El British National Corpus cuenta con unos 100 millones de palabras de inglés escrito y oral. El corpus CREA de la Real Academia Española cuenta en la actualidad con 320.759.416 de palabras.

<sup>228</sup> McEnery and Wilson, (1996: 24) indican en su definición que un corpus debe ser *machine-readable text*.

investigador Johansson define el concepto de corpus informatizado como «a body of texts put together in a principled way and prepared for computer processing» (Johansson 1998: 3).

Otras definiciones ahondan en las diferentes características esenciales de los corpus, como que deben de ser construidos según una serie de criterios específicos (Atkins *et al.* 1992: 1; Pérez Hernández 2002), que los textos recopilados deben ser seleccionados según unos criterios lingüísticos explicitables que garanticen su representatividad de la variedad de lengua seleccionada (McEnery y Wilson 1996: 24; Biber 1993; EAGLES; 1996), o que deben tener un tamaño limitado, formato electrónico y constituir una referencia estándar para la variedad de lengua que se pretenda representar (McEnery y Wilson 1996).

Finalmente, quizás la definición más extendida y estandarizada de corpus es la presentada por EAGLES (Expert Advisory Group on Language Engineering Standards), que lo define como:

A collection of pieces of language that are selected and ordered according to explicit linguistic criteria in order to be used as a sample of the language (EAGLES 1996: 4).

En esta definición el concepto de texto se deja a un lado a favor de extractos textuales o «pieces of language», para dar cabida a los corpus compuestos de extractos o subtextos. En el caso de la presente investigación, un subcorpus de páginas de contacto o de imágenes que contengan texto constituiría un subcorpus compuesto por extractos textuales.

### **6.3. Tipología de corpus**

Una vez definido el concepto de corpus, el siguiente paso lógico consiste en revisar sus diferentes tipos con el fin de situar nuestro conjunto representativo de textos

en el marco más amplio de la Lingüística de Corpus y la Traductología. La primera matización a realizar gira en torno a la diversidad de corpus existentes para el estudio del proceso y producto de la traducción, su didáctica, o para los objetivos de la Lingüística contrastiva. Esta multiplicidad ha supuesto una falta de coherencia terminológica entre estas diversas disciplinas afines, como la Lingüística contrastiva, la Lingüística de corpus o la Traductología (McEnery y Xiao 2008). La inconsistencia se debe a la falta de una tipología multidisciplinar ampliamente aceptada, y queda patente principalmente en referencia al concepto de «corpus comparable» que usamos en la presente investigación (Laviosa 1997, 2002; Kenny 1998; Olohan 2004). Esta falta de concreción nos conduce a revisar las propuestas de tipología de corpus para situar nuestra investigación en este contexto.

Varias publicaciones presentan tipologías de corpus dirigidas a estandarizar la descripción de los corpus, como EAGLES (1996)<sup>229</sup> o Teubert (1996) en Lingüística de corpus, o Laviosa (2002) y Corpas (2001) en Traductología. Como ejemplo de la variedad de perspectivas, en las investigaciones sobre géneros textuales del grupo GENTT (García Izquierdo y Nebot 2003) se usa la propuesta de la traductóloga Gloria Corpas, (2001: 3 y ss.), que parte de una propuesta en Lingüística de Torruella y Llisterri (1999), combinándola con las propuestas del lingüista Johansson (1998) y la de la traductóloga Mona Baker (1995)<sup>230</sup>.

---

<sup>229</sup> La tipología de EAGLES (1996) distingue entre: corpus de referencia (BNC o CREA), corpus monitor, corpus oral, corpus de fragmentos textuales, corpus especiales / especializados y corpora diseñados con fines especiales, corpus bilingüe o multilingüe.

<sup>230</sup> Esta propuesta (Corpas 2001: 3 y ss) clasifica los corpus según los siguientes criterios: (1) Según el porcentaje y la distribución de los diferentes tipos de texto: corpus grande o extenso, corpus equilibrado, corpus piramidal, corpus monitor, corpus paralelo y corpus comparable. (2) Según la especificidad de los documentos contenidos en el corpus: corpus general, corpus especializado, corpus canónico, corpus periódico o cronológico y corpus diacrónico. (3) Según la cantidad de texto que se reúne en cada uno de los documentos que componen el corpus: corpus textual, corpus de referencia y corpus léxico. (4) Según la codificación y anotación: corpus no anotado, corpus anotado. (5) Según la documentación que acompaña al corpus: Corpus documentado y corpus no documentado.

### 6.3.1. Tipología de corpus de Sara Laviosa (2002)

Consideramos que la tipología corpus propuesta por Sara Laviosa (2002: 34-38) es la más concisa para nuestros fines, tanto por sus objetivos normalizadores, como por pretender alcanzar la consistencia descriptiva en el objeto de estudio entre investigadores de diferentes disciplinas.

La investigadora señala que su afán no es exhaustivo, dado que su objetivo consiste en proporcionar una base común por la que describir un corpus en relación con el resto. La propuesta se estructura jerárquicamente en cuatro niveles posibles: el primer nivel presenta seis parámetros contrastivos esenciales en la descripción de todo corpus; el tipo de texto - completo, extracto o corpus monitor; la dimensión temporal - corpus sincrónicos o diacrónicos; el propósito de corpus - general o terminológico; la combinación lingüística - corpus monolingües, bilingües o multilingües; las lenguas incluidas en el corpus; o el modo del texto, escrito, oral o ambos. Este primer nivel se complementa con tres niveles consecutivos de mayor concreción y de parámetros más específicos. La siguiente tabla presenta una adaptación de la presente tipología adaptada a los objetivos de la presente tesis en la que se han incluido los corpus «desechables» (Varantola 2003), aquellos recopilados específicamente para una tarea de traducción y de especial importancia en la didáctica de la misma, que incluimos por su relativo uso en las tareas de localización (Shreve 2006). Igualmente hemos introducido los corpus «multimedia», que suponen una realidad relativamente reciente dentro de esta disciplina (Laviosa 2002: 7). Éstos intentan resolver el problema de la representación de elementos no textuales de aquellos corpus que combinan el discurso oral o una combinación de texto y fotos<sup>231</sup>. Hasta la fecha, los corpus multimedia se centran principalmente en estudios monolingües y descriptivos. Nuestro corpus constituiría, en este sentido, un corpus multimedia al incorporar

---

<sup>231</sup> Como ejemplo, en el *Children's Writing Project* de McEnery y Wilson, (1997) se mezclan entrevistas con los niños y fotografías.



imágenes, imágenes con texto, funcionalidades básicas y animaciones Flash en formato .swf.

Dimensión	Nivel I	Nivel II	Nivel III	Nivel IV
Tipos de Textos	<b>Full text</b> Sample Mixed Monitor			
Temporal	<b>Synchronic</b> Diachronic			
Propósito	<b>General</b> Terminological Desechable <sup>232</sup> o preciso (Varantola 2003)			
Combinación lingüística	<b>Monolingual</b>  Bilingual  Multilingual	<u>Monolingual</u> : Single <b>Comparable</b> <u>Bilingual</u> : Parallel Comparable <u>Multilingual</u> : Parallel Comparable	Bilingual parallel: Mono -directional Bi-directional  Multilingual: Mono-source-language Bi-source-language Multi-source Language	
Lenguas	Language(s) of corpus <b>Español</b> , inglés, francés, etc.		<b>Translational</b> Non-translational	Translational: <b>Mono-source-language</b> Bi-source-language Milti-source Language
Modo	<b>Written</b> Spoken Mixed Multimedia			

**Tabla 3. Tipología de corpus adaptada de Sara Laviosa (2002: 34-38), con el objeto de estudio de la presente tesis.**

Dentro de esta tipología, nuestro corpus constituiría un corpus comparable monolingüe español sincrónico, multimedia y especializado con textos completos y una sección original y otra traducida. De entre los distintos parámetros pasamos a profundizar en los tres tipos básicos de interés en la presente investigación descritos por Baker (1995): los corpus monolingües, los comparables y los paralelos.

<sup>232</sup> Este tipo de corpus se ha denominado también *precision corpora* (Varantola 1997) que se definen como corpora especializados creados y dirigidos a una tarea de traducción específica. Shreve (2006) también recoge que este tipo de corpus se ha denominado *virtual corpus*.

### 6.3.2. Corpus Monolingües

Un **corpus monolingüe** es aquel que contiene una sola lengua (Kenny 2001: 58) y a su vez se puede subdividir entre los corpus monolingües traductológicos o naturales u originales (Laviosa 2002: 35). Nuestro corpus es monolingüe, si bien en nuestro caso podríamos denominar *monolocale*<sup>233</sup>, debido a la necesidad de enfocar nuestra investigación desde el punto de vista de la localización. En este campo, la lengua meta se convierte en la *locale* meta (Pym 2004a), con lo que definir nuestro corpus como monolingüe quizás no fuera suficiente. La importancia del concepto de *locale* en la recopilación del corpus es esencial y lo revisaremos posteriormente (véase 7.2.2).

En el desarrollo del estudio de la lengua de la traducción<sup>234</sup> se ha debatido la terminología a emplear para denominar la sección no traducida. Baker (1993) comenzó denominándolo «original», pero debido a las posibles implicaciones por cuanto a la originalidad o creatividad de la lengua de la traducción, la investigadora pasó a denominarlo *non-translational, non-translated language* o no traducido. Olohan (2004) señala que a pesar del sesgo que suponen hacia la traducción, son términos que entrañan una menor confusión en inglés. Otros investigadores denominan la sección sin traducir del corpus *spontaneously sourced texts* (Mason 2001:68) o *spontaneously produced texts* (Zanettin 2000: 106). En la presente tesis emplearemos los términos «originales» y «traducido o localizado».

En Traducción los corpus monolingües originales o no traducidos de textos especializados se han usado sobre todo en la enseñanza de la traducción (Bowker 1998; Pearson 1999), como ayuda en la cuantificación de la calidad de las traducciones (Bowker 1999; Bowker 2001) y para la extracción de terminología

---

<sup>233</sup> Definimos nuestra locale según la W3C en es-Es, es decir, español de España.

<sup>234</sup> Según Baker (1993, 1995, 1996, 1997, 1998).

(Pearson 1999; Faber, López y Tercedor 2001). Debido a la necesidad de comparar los datos obtenidos (Kenny 2001: 125)<sup>235</sup>, se usan además varios corpus monolingües de control como el *British National Corpus* o *Cobuild Bank of English* en inglés o el CREA en español<sup>236</sup>. La sección original en nuestro corpus surgió con la intención adicional de crear un corpus de «control» claramente delimitado para estudios futuros evaluativos, didácticos o empíricos sobre el proceso de traducción en localización.

Si nos centramos específicamente en la localización, los posibles usos o aplicaciones de los corpus monolingües se resumirían en (1) los trabajos destinados a la estandarización de la terminología, (2) en las iniciativas de lengua controlada o (3) como parte de las estaciones de trabajo para la creación asistida por ordenador de textos técnicos o especializados (Shreve 2006). Nuestra aportación en este campo se centra en desarrollar una metodología con un corpus monolingüe en el estudio de las convenciones de género esenciales en este tipo de traducción.

En el contexto de la WWW, nos vemos abocado a matizar el concepto de texto monolingüe. En los textos web es común encontrar dos o más idiomas dentro de una misma página web, en parte por el carácter multinacional y multicultural del medio. Por ejemplo, es convencional la presencia de enlaces hacia otras versiones localizadas directamente en la lengua de destino. A pesar de que definimos nuestro corpus como monolingüe, en el proceso de análisis de los listados globales de palabras obtenidos con la herramienta Wordlist<sup>237</sup> de WordSmith Tools, queda patente que existe un número considerable de palabras en varios idiomas, siendo mayoritarios el inglés, francés o catalán. Como ilustración de este punto, en el siguiente gráfico muestra esta presencia multilingüe en un formulario del sitio web [www.fcasas.com](http://www.fcasas.com):

---

<sup>235</sup> Al igual que en Lingüística de corpus, como Stubbs (1996: 69-70) nos comenta «the need for comparative data is well known in corpus-based studies».

<sup>236</sup> Disponible en <http://corpus.rae.es/creanet.html>.

<sup>237</sup> Componente del programa de software WordSmith tools (véase 7.7).

Titular Name of cardholder	<input type="text"/>
Dirección Address	<input type="text"/>
Población Town	<input type="text"/>
C.Postal ZIP code	<input type="text"/>
País Country	<input type="text"/>
<b>CREDIT CARD</b> TARGETA DE CRÉDITO	
<input type="radio"/> AMERICAN-EXPRESS <input type="radio"/> MASTEI	
Nº tarjeta Card number	<input type="text"/> - <input type="text"/> - <input type="text"/> - <input type="text"/>

Figura 35. Inclusión de inglés en formulario bilingüe. [www.fcasas.es/form-2.html](http://www.fcasas.es/form-2.html)

La posibilidad conjugar varios idiomas supone una de las características textuales básicas de los sitios web corporativos, y por su importancia la analizaremos en mayor detalle en el capítulo dedicado a los aspectos extratextuales (8.1.2.1)

### 6.3.3. Corpus comparables

**Corpus comparable** en Traductología se identifica con una colección estructurada de textos digitales en una lengua junto con traducciones hacia esta misma lengua. Aunque esta metodología fue empleada con anterioridad (Gellerstain 1986), se considera a Mona Baker como la primera gran investigadora que promovió el uso de corpus comparables con la creación del *Translational English Corpus* (TEC) en la Universidad de Manchester (Baker 1995). La investigadora los define como:

A structured electronic collection of texts originally written in a particular language, alongside texts translated into that same language<sup>238</sup> (Baker 1995: 234).

Esta colección de textos se ha de recopilar de acuerdo a unos criterios que aseguren su comparabilidad, por lo que deben incluir un mismo campo, variedad de lengua,

<sup>238</sup> Esta idea de Baker se ha cristalizado en la creación del *English Comparable Corpus* (ECC) que consiste en dos colecciones de textos, una es el *Translational English Corpus* (TEC), de textos traducidos al inglés desde diferentes idiomas y accesible en <http://ubatuba/ccl.umist.ac.uk/tec/>, y otro corpus denominado *Non-Translational English Corpus* (NON-TEC) compuesto por textos originales en inglés de similares características producidos durante el mismo periodo de tiempo.

temporalidad y una extensión similar, «both corpora should cover a similar domain, variety of language and time span, and be of comparable length» (Ibid).

La investigadora señala que los corpus comparables contienen el mayor potencial para el estudio de la traducción por sí misma, vista como cualitativamente diferente de la producción textual normal (Baker 1995: 240). En la actualidad existen además corpus comparables multilingües (Bowker 2002) y corpus comparables monolingües con múltiples lenguas origen (Laviosa 1997, 2002). Definimos nuestro corpus como un corpus comparable monolingüe o *monocale*, ya que ambas secciones del corpus contienen textos en la variedad regional español de España (*es-ES*). Dentro del corpus, una sección contiene textos originalmente producidos en español, y la otra sección consiste de textos localizados hacia esta misma región o *locale*.

No debemos confundir la noción de corpus comparable en Traductología con la más extendida en Lingüística contrastiva y de corpus<sup>239</sup> (Peters *et al.* 1996: 69; Teubert 1996: 250). En estas disciplinas este término se identifica con una colección de textos en dos o más lenguas que sin ser traducciones, pertenecen a un mismo tema, campo, género o motivación situacional y comunicativa<sup>240</sup>. Esta confusión queda patente en ciertas publicaciones en las que se ha introducido el concepto de comparable que se corresponde con su acepción en Lingüística (Shreve 2006)<sup>241</sup>. En este caso, la acepción dentro de la tipología adoptada sería un corpus comparable bilingüe, opuesto al corpus comparable monolingüe (Baker 1995).

---

<sup>239</sup> Peter *et al.* (1996: 154) definen un corpus comprable en Lingüística como «sets of translation equivalent texts in which generally one text is the source text and the other (s) are translations».

<sup>240</sup> Un ejemplo sería el *Network of European Textual Reference Corpora* (NERC).

<sup>241</sup> En el área de la localización Shreve (2006) describe los corpus comparables como «documents in a foreign language that are not translations of the source-language corpus but are in the same domain». Esta definición se corresponde con los corpus bilingües comparables en Traductología.

A pesar de que ciertos investigadores señalan la escasez de investigaciones basadas en corpus comparables (Olohan 2004)<sup>242</sup>, en la actualidad están en aumento el número de investigaciones traductológicas que adoptan esta metodología (Mauranen 2008; McEnery y Xiao 2008). Éstas investigan las características de los textos traducidos comparados con los textos originalmente producidos en la lengua idioma meta. Destacan en especial los estudios de Baker (1995, 1996, 2001) sobre los controvertidos «universales de la traducción»<sup>243</sup> en textos traducidos al inglés, una corriente adoptada por otros investigadores que han estudiado las peculiaridades textuales de las traducciones en otros idiomas (Toury 1980; Vanderauwera 1985; Puurtinen 1995; Laviosa 1996, 1997, 1998; Kenny 2001). En estas investigaciones basadas en corpus comparables se ha demostrado los textos traducidos suelen presentar estructuras y frecuencias atípicas en comparación con textos producidos originalmente en la misma LM (Gellerstain 1986; Laviosa-Baithwate 1996, 2002)<sup>244</sup>. Las peculiaridades de los textos traducidos se justifican por la «lengua de la traducción», que refleja rasgos que se observan exclusivamente en los textos traducidos o que aparecen en una frecuencia significativamente superior o inferior en los mismos (Baker 1995: 235). En otras investigaciones este concepto se ha denominado «third code» (Frawley, 1984), «the third language» (Duff 1981) o «hybrid language» (Trosborg 2000). Estas diferencias se justifican en la naturaleza del texto traducido en general y la naturaleza del proceso de traducción en particular (Baker 1995: 236), ya que la traducción se concibe a como «a communicative event which is shaped by its own goals, pressures and context of production» (Baker 1996: 175). De ahí se concluye que el producto, el texto

---

<sup>242</sup> El investigador identifica la escasez de estudios sistemáticos de corpus comparables con el resultado de las mismas más centrado en la descripción de la metodología y de los descubrimientos sobre los datos lingüísticos (Olohan 2004: 41).

<sup>243</sup> Entre los «universales» que se han propuesto se encuentra la simplificación, la explicitación, la normalización, el «leveling out», la desambiguación y la estandarización (Baker 1995, 1996; Laviosa 1998a, 1998b, Olohan y Baker 2000; Olohan 2001). Los universales de la traducción se definen como «features that typically occur in translated text rather than original utterances and which are not the result of interference from specific linguistic systems» (Baker 1996: 176).

<sup>244</sup> En Laviosa (2002) y Olohan (2004) se encuentran revisiones en profundidad de las diferentes investigaciones y sus resultados.

traducido, mostrará ciertos rasgos lingüísticos recurrentes y significativos debido a ciertas restricciones que operan durante el proceso:

Translated language reflects the constraints which operate in the context of production and reception: these constraints are social, cultural, ideological, and of course, of cognitive nature (Baker 1999: 285).

Estas restricciones se deben a las diferencias en el contexto de producción que se manifiestan en estructuras textuales distintas:

A translation, like any kind of text production, develops in response to the pressures of its own immediate context and draws on a distinct repertoire of textual patterns (Baker 1996: 176).

Las investigaciones con corpus comparables se conciben por tanto como una metodología para el estudio del producto y del proceso de traducción (Baker 1993; 1995, 1996, 1999; Olohan 2004). Entre otros, Gideon Toury (1995) defiende esta noción del estudio de la traducción sin la inclusión del TO. El investigador señala que el estudio conjunto del contexto de la traducción y del producto puede mostrar parte de las decisiones que se realizan durante el proceso y los factores que restringen la actuación de los traductores (ibid: 37). Estos aspectos son identificados por el autor como los «observables»: las propias unidades de habla traducidas junto con sus elementos constituyentes (ibid: 36). De igual modo, Bell (1991: 39) sugiere el estudio del proceso a través del producto de la traducción, ya que éste puede revelar rasgos que sugieran la existencia de elementos particulares y de relaciones sistemáticas en el proceso. Olohan sintetiza esta posición de la siguiente manera:

The compilation and use of corpora of translations would allow us to analyze features of translation products which can provide evidence of translation processes, both conscious and subconscious (Olohan, 2002: 6).

A pesar de lo expuesto en el párrafo anterior, el uso de corpus comparables en Traductología no ha estado exento de críticas. Olohan (2004) sintetiza las mismas

en aquellas dirigidas hacia (1) la primacía del producto de la traducción frente al proceso, (2) el papel que desempeña el TO en el análisis de la traducción o (3) los usos o aplicaciones potenciales de los resultados de tales investigaciones. De entre los detractores, Stewart defiende la mayor idoneidad de los corpus paralelos en la investigación del proceso de traducción (Stewart 2000: 210). No obstante, nuestra posición está en consonancia con la de Maeve Olohan (2004: 39), que defiende cómo mediante las investigaciones con corpus comparables puede observarse el resultado de la actividad traductora y, consecuentemente, se prioriza el proceso y los factores que influyen en éste, siendo el texto origen uno de los factores que más influye.

Hemos de apuntar que la propia Mona Baker reflexiona en una publicación reciente sobre las posibles deficiencias de esta metodología (Baker 2004: 183-184). Sus conclusiones se centran en (1) la necesidad de entender los datos lingüísticos obtenidos como un simple punto de partida hacia una explicación situada de los fenómenos, dando cabida a los aspectos sociales y cognitivos, (2) la imposibilidad de identificar exclusivamente los datos cuantitativos obtenidos con su interpretación, y (3) los problemas de subjetividad del investigador por su método de interpretación y por el sesgo se produce al centrarse exclusivamente en ciertas características que selecciona.

#### 6.3.4. Corpus paralelos

Un **corpus paralelo** lo conforman textos originales en una lengua con las traducciones de los mismos hacia una o varias lenguas (Laviosa 2002: 37; McEnery 1996: 58). En general, ésta es la descripción más extendida en Lingüística<sup>245</sup>, aun cuando varios investigadores en Traductología toman el concepto proveniente de análisis contrastivo que equivaldría al corpus comparable (véase Laviosa 2002: 37).

---

<sup>245</sup> Es preciso no confundir el concepto de corpus paralelo de lingüística de corpus usado en esta tesis con el concepto de de textos paralelos en traducción desde una perspectiva didáctica.



Con el objetivo de clarificar la confusión Johansson (1998) defiende que este tipo de corpus se pase a denominar «corpus de traducciones» o *translation corpus*.

Las críticas a su uso se basan en que los textos traducidos no representan la gama de posibilidades lingüísticas de la lengua de destino y pueden reflejar el estilo particular de traductores individuales (Teubert 1996: 247), por lo que no representa la lengua de destino de forma natural y genuina (McEnery y Xiao 2008). No obstante, si el corpus es suficientemente extenso, este puede mostrar ejemplos de estrategias de traducción en casos concretos o puede emplearse para el análisis de propuestas de traducción (Zanettin 1998).

### **6.3.5. Corpus comparable frente a corpus paralelo en la investigación traductológica**

En las fases previas de la investigación, tras un análisis de la idoneidad de ambos tipos de corpus al objetivo de nuestro estudio, nos decantamos por el uso de corpus comparable frente al paralelo, cuyo uso está más extendido en estudios descriptivos de género en nuestro país (ej. Gamero 2001; Fernández Sánchez 2004; Vázquez del Árbol 2006; Sánchez Nieto 2006). Como indicamos anteriormente, no incorporamos la metodología de corpus paralelo ya que éstos normalmente constan de una sola versión traducida por texto origen, y ésta responde exclusivamente a un proceso de traducción único (Malmkjær 1998). En este sentido, los traductores pueden cometer errores y cada traducción supone una de las posibles traducciones en un caso determinado (Johansson 1998: 137). Varios investigadores han señalado que las investigaciones con corpus paralelos pueden esconder ciertos rasgos de interés del fenómeno traductor, como pueden ser las traducciones diferentes de un mismo original (Laviosa 2002: 28; López y Tercedor 2008).

Los resultados de nuestra investigación piloto realizados con un corpus comparable (véase 1.3.1), se ajustan a los de otras investigaciones pioneras en el uso de esta metodología (Laviosa 1997; Kenny 1998, 2001; Olohan y Baker 2000; etc.); por un lado en la aparición de estructuras inexistentes en el corpus de páginas en español, como «cláusula de privacidad», o «términos de uso» o «home»<sup>246</sup>, y por otro, en la frecuencia atípica de varias colocaciones como «Contacta con Nosotros» o «Sobre nosotros»<sup>247</sup>. En nuestro caso, los resultados obtenidos se interpretan además por los efectos y presiones del entorno del proceso de traducción durante la localización, que según nuestra hipótesis, incide de manera diferenciada a la traducción literaria (Kenny 2001; Laviosa 1997) o periodística (Baker 1995)<sup>248</sup>. Estos resultados motivaron la ampliación de la investigación y la decisión de adoptar ésta para investigar el proceso y el producto de la localización web. El uso de un amplio corpus comparable nos permitirá obtener traducciones múltiples a un mismo segmento textual recurrente o una forma lingüística convencional en el género de estudio, por lo que se posibilita obtener el grado de convencionalización de cada solución posible y compararlo con el corpus de control original.

No podemos terminar este epígrafe sin hacernos eco de la posible combinación de ambas metodologías por la que abogan varios investigadores: el uso de comparable y paralelo para el estudio del proceso y producto de la traducción. Las razones que se esgrimen serían la influencia de la producción diferenciada y la necesidad de ir más allá de soluciones únicas realizadas por un traductor individual (Johansson 1998: 6; Tercedor 1999; López Rodríguez 2001; Kenny 2003; McEnery y Xiao 2008).

---

<sup>246</sup> Reflejo de la colocación conveconalizada en los géneros web de páginas corporativas norteamericanas «Privacy Policy», «Terms of use» y «home» respectivamente.

<sup>247</sup> «Contact us» y «About us» respectivamente en inglés.

<sup>248</sup> Un ejemplo de la influencia de este contexto de traducción diferenciado sería la frecuente aparición de faltas de ortografía raramente presentes en textos en otros medios (véase capítulo 11).



**PARTE II:**

**MARCO  
METODOLÓGICO**



## 7. METODOLOGÍA DE LA INVESTIGACIÓN EMPÍRICA

Una vez establecido el marco teórico en el que se fundamenta la presente tesis, pasamos a describir el marco metodológico conforme al ciclo del proceso heurístico en investigaciones empíricas traductológicas (Hurtado Albir 2001: 190). En esta fase describiremos el diseño de la investigación empírica y el proceso de recogida de datos, reflejado en un corpus comparable monolingüe<sup>249</sup> del género objeto de estudio.

### 7.1. Metodología de análisis

Realizaremos el análisis del corpus de modo integrador entre la teoría y la práctica como propone Göpferich (1995: 74), a su vez basándonos en la superestructura como *tertium comparationis* siguiendo la metodología de Gamero (1998, 2001). Siguiendo una progresión lógica, el análisis textual procederá en un *continuum* metodológico de arriba-abajo como propone Baker (1992). El mismo comenzará con los aspectos extratextuales definitorios del género (véase 5.5), obligatoriamente similares entre ambas secciones del corpus y que no admiten variación (Gamero 2001). A continuación el análisis procederá hacia los aspectos intratextuales que se enmarcan y delimitan a través de los extratextuales. Esta progresión de lo extra a lo intratextual se observa en la gran mayoría de los análisis de géneros previos<sup>250</sup> (Gamero 1998, 2000; Göpferich 1995; Fernández Sánchez 2004; Bolaños 2002a, 2002b; Sánchez Nieto 2006).

---

<sup>249</sup> En la concepción de corpus comparable nos basamos en Baker (1995) y Laviosa (2002).

<sup>250</sup> Esta progresión desde lo extratextual hacia el nivel intratextual se recoge en los análisis textuales aplicados a la traducción (Baker 1992; Bravo 1999; García Izquierdo 2000).

Las investigaciones traductológicas sobre géneros se hacen eco de la inexistencia de un modelo de análisis contrastivo consensuado entre la comunidad científica (Gamero 1998, 2000; Göpferich 1995; Fernández Sánchez 2004). Ante este estado embrionario de los análisis traductológicos basados en géneros, nos es indispensable examinar los diferentes estudios previos para recoger sus categorías de análisis, problemas y experiencias<sup>251</sup> (Pöckl 1998: 18, cf. Fernández Sánchez 2004). No obstante, hemos de mencionar que recientemente han surgido en nuestro país propuestas desde la Lingüística (Ciapuscio y Kugel 2002), ya aplicadas a análisis contrastivos de géneros desde una perspectiva traductológica (Sánchez Nieto 2006).

En nuestro caso, los sitios web corporativos han sido objeto de investigaciones previas desde la Lingüística (Fernández 2003; Alamán 2002, 2003; e. o.), el IFE (Rennau 2004) o la Traductología (Bolaños 2002a, 2002b; Bolaños *et al.* 2005). La mayoría de estas investigaciones toma como punto de partida la corriente que analiza los géneros digitales desde las Ciencias de la Información (Crownston y Williams 1997; Shepherd y Watters 1998).

### **7.1.1. Análisis extratextual e intratextual**

Como hemos mencionado, nuestra metodología de análisis parte del nivel extratextual y extralingüístico, analizando contrastivamente la función textual, el contexto sociocultural, y el contexto comunicativo, para pasar al nivel intratextual, de mayor concreción y basado en la macro y microestructura. En el análisis convergerán métodos cuantitativos<sup>252</sup> y cualitativos de forma diferenciada, con

---

<sup>251</sup> En este sentido, Silvia Gamero (2000: 141) recoge las categorías que se han analizado con anterioridad a su estudio.

<sup>252</sup> Según Hurtado (2001: 175), los métodos cualitativos se centran en la cualidad y su objetivo es la comprensión, la descripción o el descubrimiento, y se emplean métodos inductivos. Los métodos cuantitativos es la cantidad y su objetivo es siempre la predicción, el control, la descripción, la confirmación y la comprobación de hipótesis y se utilizan análisis deductivos mediante métodos estadísticos.

métodos cualitativos en la consideración de los criterios extratextuales, y mayormente cuantitativos en el análisis de los criterios intratextuales. Ambos procedimientos constituyen un *continuum* metodológico (Hurtado Albir 2001: 175), que no polos antagónicos, con lo que combinaremos ambos métodos dependiendo del objetivo de cada fase o la técnicas empleadas (Orozco 2000). Esta combinación de criterios cualitativos y cuantitativos en los análisis de corpus se ha señalado como necesaria (Ian Mason 2001: 71), sobre todo en los análisis que se centren en el discurso o el género (ibid: 78).

Nuestro análisis intratextual partirá con la obtención de la superestructura prototípica del género siguiendo a Gamero (1998, 2001), sobre la que se fundamentará el estudio descriptivo que, a su vez, nos proporcionará un *tertium comparationis* sobre el que basar el análisis contrastivo entre textos originales y traducidos. Debido a la extensión y la enorme posibilidad de elementos de análisis, en una etapa temprana de la investigación nos fue necesario seleccionar los elementos intratextuales del género en los que centrar el análisis contrastivo del género (Paltridge 1997: 3). Nuestra posición está en consonancia con la defendida por investigadores que manifiestan la imposibilidad de analizar exhaustivamente el género, sobre todo por la extensión de los textos tratados (Hasan 1989; Paltridge 1997). Desde su perspectiva, Halliday apoya igualmente la noción del análisis lingüístico selectivo (1985b):

The guiding principle is to select and develop what is needed for the particular purpose in hand. There are many different purposes for analyzing a text, and the scope and direction of the analysis will vary accordingly. Often we may want to scrutinize only one or two features, but to follow them to a considerable depth (Halliday 1985b: 54).

Por lo tanto, en cada etapa realizaremos una selección de los criterios objeto de análisis a través de métodos cuantitativos y deductivos. Tras el primer paso, el estudio contrastivo de las superestructuras prototípicas, estudiaremos el nivel de convencionalización de las formas lingüísticas asociadas al mismo. En este sentido,



los menús de navegación presentan los mayores niveles de convencionalización (Nielsen y Tahir 2002) y, por consiguiente, los hemos seleccionado para un análisis contrastivo de este aspecto. A continuación, y con el fin de realizar un estudio con un mayor nivel de concreción, analizaremos contrastivamente dos bloques que seleccionamos por su importancia dentro del género de estudio: el bloque **contacto** (B), el de mayor frecuencia en el conjunto, y el bloque **legal** (F). Como se verá más adelante, éste último presenta las mayores diferencias superestructurales entre ambas secciones del corpus, y esta diferencia nos permitirá analizar la posible relación existente entre los diferentes grados de convencionalización de bloques textuales en distintos contextos socioculturales y su impacto en el texto traducido.

En segundo lugar se analizarán contrastivamente los elementos textuales integrantes del segundo nivel estructural presente en los sitios web, los elementos internos de estructuración de toda página web. Estos elementos son dados por las normas de base tecnológicas (véase 8.1.3.1) y que permiten la imbricación de toda página y sitio web en el contexto más amplio de la WWW. Estos segmentos estructurales secundarios poseen una menor importancia en el intercambio comunicativo directo entre el emisor y el usuario<sup>253</sup>, puesto que se destinan a un proceso comunicativo entre el sitio web y la hiperweb global. Los segmentos que estudiaremos son los títulos, las descripciones en los encabezamientos de páginas web, los segmentos que facilitan la accesibilidad y los segmentos basados en la interactividad (véase 9.2).

### **7.1.2. Metodología de análisis de la superestructura**

La metodología de análisis intratextual comenzará con la obtención y contraste de la superestructura prototípica de ambos pares textuales. Con este objetivo se descompondrá en sus distintos bloques, secciones, unidades y subunidades

---

<sup>253</sup> Con la excepción de los segmentos que facilitan la accesibilidad tales como los segmentos en <alt>.

significativas para permitir y delimitar su estudio contrastivo (véase 5.3.1). El procedimiento establecido para identificar los bloques comunicativos y distinguirlos de las secciones es de carácter cuantitativo y se basa en investigaciones previas sobre textos lineales impresos, por lo que ha sido necesaria su adaptación al entorno hipertextual. En éstas, los procedimientos de identificación se fundamentan en tres rasgos esenciales de carácter tanto verbal como no verbal, a partir de los cuales se establece cuantitativamente la superestructura prototípica como base contrastiva (Göpferich 1995; Gamero 1998, 2001; Fernández Sánchez 2004; etc.). En resumen consisten en:

- La presencia de títulos, subtítulos, encabezamientos de sección y párrafos (Gamero 1998: 262) o texto en negrita, que son representativos de un bloque o sección y a su vez pueden estar recogidos en un índice de contenidos.
- Los aspectos tipográficos, puesto que los bloques y secciones pueden aparecer como entidades propias dentro del texto separadas por espacios dobles, distinto formato, etc. (Göpferich 1995: 389-390).
- En último lugar, la identificación por medio de la competencia textual genérica e intuición del investigador que realiza el análisis de un bloque o sección independiente que no pertenece a ningún apartado lógico de orden superior (Gamero 2001: 109).

La superestructura adquiere en los hipertextos una importancia mayor que en los textos impresos (Price y Price 2002: 68), y se encuentra representada explícitamente en los menús de navegación y en mapas web<sup>254</sup>. Los menús de navegación poseen una finalidad metacomunicativa esencial en todo texto web (Alejo 2005), puesto que han de ser el referente superestructural para todo usuario

---

<sup>254</sup> Otros investigadores identifican la superestructura o estructura subyacente en los sitios web con la organización subyacente en el servidor organizada en árboles estructurales y en carpetas (Mata Pastor 2005). Sin embargo, ambas estructuras, la superestructura del hipertexto y su estructura subyacente en carpeta, pueden no corresponderse necesariamente. De hecho, en nuestra investigación hemos constatado que la prescripción de organizar la estructura subyacente según la estructura lógica del sitio suele no cumplirse en los sitios web corporativos españoles.

que desconozca los límites reales del hipertexto (Conklin 1987). Concederemos una importancia especial a estos menús que, a través de este valor metacomunicativo, denotan el contenido de sus respectivas páginas o bloques textuales.

El procedimiento de identificación se basa recopilar manualmente las unidades léxicas presentes en menús de navegación que contienen enlaces hacia los distintos bloques y secciones. A continuación analizamos exhaustivamente las páginas hacia las que se dirigen estos enlaces, procediendo a asignar cada unidad léxica a un bloque o sección determinado. Siguiendo la investigación de Bouffard y Craignon (2006), comprobamos que los menús de navegación se componen principalmente de unidades léxicas, aunque pueden asimismo incluir combinaciones léxicas. La pertinencia de este análisis manual se debe a que las herramientas de software empleadas no recogerían los elementos textuales de navegación codificados en imágenes o en archivos en formato Flash. Adicionalmente, nos sería imposible asignar cada unidad léxica a un bloque determinado sin haberlo analizado manualmente.

Nuestro análisis se cimienta en una de las peculiaridades básicas de la hipertextualidad: la necesidad de dividir los contenidos en páginas web debido a la presentación en pantalla (Brinck *et al.* 2002: 284). Esto implica que el contenido de cada página representa una entidad comunicativa diferenciada y unitaria. El hecho de crear una página web independiente es indicativo de la presencia de un bloque o sección independiente que lógicamente posee una función específica (Askehave y Nielsen 2005). Argumentamos que esta particularidad da validez a nuestro análisis y, consecuentemente, emplearemos los porcentajes de frecuencia de bloques y secciones como páginas independientes en la superestructura para analizar sus grados de convencionalización. No obstante, hemos de matizar que un bloque comunicativo puede incrustarse en el seno de otra página web, por lo que

la inexistencia de una página independiente no apunta necesariamente a que tal o cual bloque no esté presente. Por ejemplo, una página web puede mostrar directamente la información de contacto: en las páginas web españolas esto sucede en un 13,96% de casos (véase 9.1.2.2). Consecuentemente, los porcentajes que presentaremos son indicativos de la existencia de páginas web independientes para cada bloque o sección y que relacionamos con:

- (1) La otorgación tácita de un mayor rango por parte del emisor.
- (2) Un mayor nivel de convencionalización de esta parte del texto.
- (3) Una mayor relevancia de la función parcial de un bloque o sección determinados.

En tal sentido, los porcentajes son relativos y específicos para nuestra investigación sobre textos web y responden a la metodología que hemos desarrollado, por lo que no se corresponden con los porcentajes absolutos de frecuencia presentes en otras investigaciones sobre géneros impresos<sup>255</sup>.



Figura 36. Bloque «contacto» incluido en la página de inicio.

<sup>255</sup> En el caso de comparar los porcentajes obtenidos con investigaciones de géneros impresos emitidos por empresas como los folletos bancarios (Alberola 2005).

A pesar de las diferencias entre nuestra metodología y los métodos de identificación utilizados en investigaciones similares sobre textos impresos, nuestro procedimiento basado en elementos verbales y no verbales se corresponde con los utilizados anteriormente (Göpferich 1995a; Gamero 1998; Fernández Sánchez 2004; Vázquez del Árbol 2006). Justificamos estas similitudes a través de los siguientes puntos:

- a) En primer lugar, cada elemento en un menú de navegación se asemeja al título<sup>256</sup> de una sección o párrafo en los textos impresos. A su vez, y por economía en pantalla, los menús de navegación suelen aparecer jerarquizados y reflejan el rango secundario correspondiente a ciertas secciones: las secciones de menor rango dentro de un bloque comunicativo suelen visualizarse exclusivamente cuando se activa el enlace de rango superior. Este carácter jerárquico de bloque y secciones se identifica igualmente en los mapas web gráficos representado en la Figura 37.

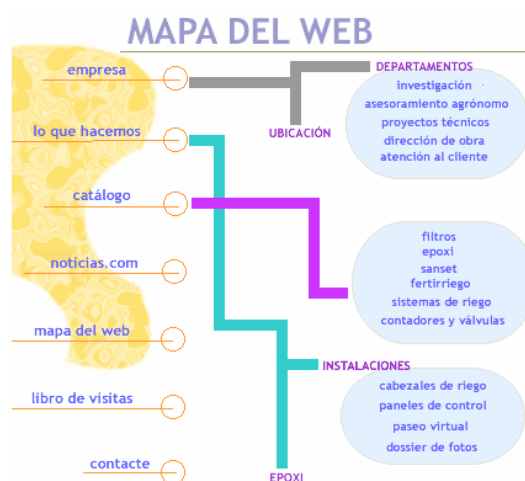


Figura 37. Superestructura representada en el mapa web. De [www.arteriego.com](http://www.arteriego.com).

<sup>256</sup> Cada página web incluye un título que aparece en la barra superior del navegador. Sin embargo, metacomunicativamente su importancia es menor y se suelen repetir los mismos o no se incluyen. Este elemento textual se tratará en detalle en el apartado (véase 9.2.1).

En nuestro caso, podríamos haber empleado el análisis de los «mapas web» que representan visualmente la superestructura textual, aunque debido a su baja presencia en los sitios originales (20,30%), su uso para este propósito no fue factible.

- b) En segundo lugar, los bloques se pueden identificar tipográficamente (Göpferich 1995): los elementos en los menús de navegación aparecen marcados por (1) un tipo y tamaño de letra diferente, (2) por su situación prominente en la periferia de la pantalla, (3) por estar rodeados de más espacio en blanco que el resto del texto (Price y Price 2002), y sobre todo, (4) funcionalmente al incluir un hiperenlace.
- c) En tercer lugar, y en caso de que un bloque o sección no pertenezca a ningún apartado lógico de orden superior, recurriremos al análisis del bloque textual y a la intuición del investigador para adjudicar una sección a un bloque determinado. De igual modo, en caso de conflictos tales como la incorporación de dos o más secciones o bloques en una misma página, procedimos a su análisis y adjudicación en función del mayor porcentaje de contenido.

El conjunto de las variables de análisis que forman la superestructura del género se estableció tras el examen de la totalidad de los sitios web en ambas secciones del corpus. La lista de variables identificadas nos permitirá dilucidar qué bloques y secciones están presentes en cada texto, y a su vez, se utilizarán para obtener los porcentajes de frecuencia de cada bloque. A través de estos porcentajes se establecerá qué bloques y secciones son recurrentes u ocasionales. El porcentaje de aparición que confiere el carácter de recurrente u ocasional a un elemento lingüístico fluctúa según la investigación. Para adquirir el carácter de recurrente varía desde el 90% (Fernández Sánchez 2004: 256; Nielsen 2004), el 70% (Hoffman 1988: 154; cf. Fernández Sánchez, 2004), o incluso al 50% mínimo para los

elementos textuales que de otra manera se consideran ocasionales (Gamero 1998: 314).

Como se mencionó anteriormente, nuestra metodología de análisis se basa en la identificación y recuento de páginas web independientes indicativas de la existencia de un bloque o sección comunicativos. Los porcentajes de investigaciones sobre textos lineales se han de adaptar a este método de análisis y por consiguiente, consideraremos como «altamente recurrentes» aquellos bloques o secciones con una frecuencia de aparición superior al 70% (Hoffman 1988). A nuestro parecer, este porcentaje se ajusta a la metodología que hemos adoptado, y queda ampliamente justificada por la frecuencia de los dos bloques de mayor frecuencia y que consideramos básicos en el prototipo genérico de los sitios web corporativos: la página web de inicio (100%) y el bloque contacto en los textos originales (86,4%). Consideremos como «recurrentes» aquellos bloques o secciones con una frecuencia entre el 70% de aparición propuesto por Hoffman (1988), y el 50% que propone Silvia Gamero (1998). El carácter «ocasional» se atribuirá a aquellos elementos con una frecuencia inferior al 50%<sup>257</sup>. A los bloques o secciones recurrentes les corresponderá una función parcial principal, y a los bloques ocasionales una función parcial subsidiaria.

## **7.2. Recopilación del corpus comparable**

Durante todo proceso de recopilación de un corpus se presentan ciertas cuestiones relativas al establecimiento de los preceptos básicos que garanticen sea representativo de la población textual seleccionada (Pérez Hernández 2002). En este sentido, existe un consenso amplio sobre la importancia de las consideraciones teóricas en la construcción de corpus, que han de ser claras y preceder a su

---

<sup>257</sup> Desde el ámbito de la usabilidad web, Jakob Nielsen distingue entre las normas o «standards», con un 80% de frecuencia o superior, las convenciones, del 50% al 80%, y la «confusión» a todo elemento con una frecuencia inferior al 50% (Nielsen 2004).

creación: «theoretical research should always precede the initial corpus design and actual compilation of text» (Biber 1993: 256).

Los estudios traductológicos de corpus se hacen eco de la gran importancia que entraña el diseño del corpus (Kenny 2001; Zanettin 2000; Varantola 2003), y dado que su proceso de construcción entraña un gran esfuerzo y tiempo, nos resulta indispensable establecer unas bases teóricas sólidas que permitan obtener el mayor rendimiento posible (Zanettin 2000). Además, en general la validez de los resultados obtenidos está en principio ligada a la validez y solidez del propio corpus:

The decisions that we take about what is to be in the corpus, and how the selection is organized, controls almost everything that happens subsequently. The results are only as good as the corpus (Sinclair 1991: 13).

Partimos así con la premisa de que la validez y fiabilidad en investigaciones basadas en corpus dependen de las bases teóricas que guíen su proceso de recopilación (Kenny 1998). De hecho, somos conscientes de que las posibles deficiencias pueden suponer la inoperabilidad del corpus para los objetivos que motivaron su concepción:

Issues in corpus design and compilation are fundamentally concerned with the validity and reliability of research based on a particular corpus, including whether that corpus can serve the purposes for which it was intended (Kenny 1998: 60).

En nuestra opinión, éste es uno de los puntos de que adolecen varias de las investigaciones sobre textos web hasta la fecha, en las que se usa el concepto de corpus sin definir qué entienden por el mismo o analizar y explicitar qué criterios han llevado a su recopilación. Desde un punto de vista empírico, estas deficiencias suponen la imposibilidad de replicar los resultados presentados.



En este sentido, defendemos que nuestro proceso de recopilación cuenta con unos principios teóricos que dan solidez a nuestra investigación. Así, en nuestro marco teórico establecimos varias de las cuestiones fundamentales, tales como la delimitación del género que pretendemos representar (capítulo 5) o qué constituye la unidad textual que lo representa (capítulo 3). La recopilación de nuestro corpus se llevó a cabo según los criterios más aceptados en los estudios de traducción (Pérez Hernández 2002), y procedimos de manera cíclica, con la revisión de nuestros principios de recopilación durante la fase de análisis cuando lo estimamos oportuno (Biber 1993: 256).

Regardless of the initial design the compilation of a representative corpus should proceed in a cyclical fashion (Biber 1993: 256).

Por ejemplo, tras los errores en pruebas preliminares de descarga observamos que las páginas de presentación en formato Flash impedían el análisis sintáctico de la herramienta Htrack, con lo que nos fue indispensable prescindir de las mismas (véase 7.7). Así, aun cuando defendemos que estas presentaciones forman parte del texto global, debimos eliminarlas para garantizar la descarga sincrónica de los sitios seleccionados. Otro problema surgió por la presencia de sitios web dinámicos con servidores activos que impedían el análisis del programa Htrack, por lo que procedimos a descargarlos a partir del mapa web. Como describimos previamente, los mapas web reflejan la estructura jerárquica hipertextual<sup>258</sup> y constituyen un punto de entrada al sitio web óptimo para nuestros fines.

### **7.2.1. Criterios en la recopilación del corpus**

Tal y como hemos apuntado, los criterios que han de guiar toda recopilación de corpus han de ser sólidos y explicitables. Actualmente, las publicaciones sobre

---

<sup>258</sup> Como fue el caso de [www.lexmark.es](http://www.lexmark.es) en el corpus. Este sitio web se descargó a partir del mapa web incluido en el mismo.

criterios de recopilación es extensa, y podemos sintetizar que los conceptos más repetidos en el seno de la comunidad científica son: la **representatividad**, la **estandarización** y **tipología de los corpus** (Pérez Hernández 2002). En nuestro caso específico, otro criterio de primordial importancia será la **comparabilidad** (Baroni y Bernardini 2003; Laviosa 1997).

En consonancia con la mayoría de investigaciones, consideramos que la **representatividad** constituye el criterio de mayor relevancia (Biber 1993, 1998; Kennedy 1998; Laviosa 2002; Olohan 2004; Zanettin 2000: 105; Halverson 1998; etc.). En palabras de Biber (1993):

A corpus is put together in a principled way so as to be representative of a larger textual population, in order to make it possible to generalize findings concerning that population<sup>259</sup> (Biber 1993: 243).

La representatividad se asegura al incluir toda posible variación en la población textual, «the full range of variability in a population» (Biber 1993: 243). Desde nuestra perspectiva traductológica, y con el fin de unir esfuerzos con el sector de la localización, la **población textual** a la que hacemos referencia la representan los sitios web corporativos en la *locale es-ES*<sup>260</sup>, o español de España, definida por el World Wide Web Consortium (W3C), el consorcio que regula Internet. Este código surge de la combinación de la norma ISO 639, que representa los códigos de lenguas<sup>261</sup>, con la norma ISO 3166, que representa los códigos de países<sup>262</sup>. Según las recomendaciones de este consorcio, las páginas web deberían incluir este código en su encabezamiento marcado con la etiqueta html <HEAD><\HEAD>, bajo el atributo normalizado<sup>263</sup> <LANG=>. No obstante, en nuestra investigación hemos constatado que la inclusión de este atributo es un *desideratum* que no se refleja ampliamente en las páginas web: el 30,56% de páginas localizadas del corpus

---

<sup>259</sup> Énfasis propio.

<sup>260</sup> Los códigos de países se pueden obtener en <http://www.w3c.org/International/O-HTML-tags.html>

<sup>261</sup> ISO 639 *Codes for the representation of names of languages*.

<sup>262</sup> ISO 3166 *Codes for the representation of names of countries*.

<sup>263</sup> Este atributo se encuentra normalizado por la directiva RFC 3066 del consorcio W3C.

incluyen este atributo, y de entre ellas, sólo el 25,69% indican explícitamente el español (<lang="es"> o <lang="es-es">)<sup>264</sup>. Otro indicador de pertenencia a la *locale* es-ES sería el dominio «.es»<sup>265</sup> aunque, metodológicamente, no podríamos establecerlo como representativo de la población textual objeto de estudio dado que muchas empresas españolas se decantan por los dominios *.com*, o *.net* como primario.

La decisión metodológica de usar el concepto de *locale* y no lengua está en consonancia con investigaciones previas en localización. Como comenta Pym (2004), su uso se debe a la gran amplitud del concepto de lengua, «due to many countries, communities, dialects, etc.» (Ibid: 22). Concebimos el concepto de *locale* como un «prototipo» (Rosch 1978, 1977), sin unos límites claros de categorización que establezcan la separación. En la acotación de una *locale*, Anthony Pym señala que son el mercado y la resistencia a la distribución la que hacen que aparezca:

Locales do not exist until they show themselves by resisting some process of distribution. (Pym 2004a: 22).

Según el autor, si un texto se puede distribuir sin localización, no existirá una *locale* nueva<sup>266</sup>. Consecuentemente, según el investigador la diferencia fundamental entre *locale* y un dialecto sería la dependencia de un mercado en la constitución. En nuestra opinión, hemos de añadir que la resistencia en el contexto web se manifiesta en estudios empíricos que han demostrado que los usuarios prefieren realizar transacciones en sitios web en su lengua (Singh y Pereira 2005: 5). Así, aun cuando los usuarios internacionales puedan usar sitios web en inglés, desde un punto de vista económico la localización produce importantes beneficios a las empresas (Yunker 2003: 12-13). Esta resistencia puede depender igualmente

---

<sup>264</sup> La mención explícita de la traducción puede observarse en el sitio web de Fedex <html lang="es-TRAD">. Esta explicitación del carácter de traducción queda recogido en el 4,21% de sitios web localizados.

<sup>265</sup> Los códigos de país de dominios se asigna por la [Internet Assigned Numbers Authority \(IANA\)](http://www.iana.org).

<sup>266</sup> Como podría ser el caso de las lenguas indígenas en Norteamérica.

del campo de especialidad: DePalma (2006) observó cómo en la comunicación web se suele preferir el inglés dependiendo de la especialización de la audiencia o del precio del producto. Esta preferencia se encuentra asimismo en ciertos sectores como la ciencia, la banca internacional o el software, paralelamente al uso del inglés como lengua de especialidad internacional (Posteguillo 2003).

Dentro de los estudios sobre corpus comparables, la **comparabilidad** de las dos secciones es primordial y se equipara en importancia a la representatividad (Baroni y Bernardini 2003: 2; Laviosa 1997, 2002; Baker 2004; Olohan 2004). Podemos definir este concepto como la obligatoria similitud entre los textos de ambas secciones en tantos aspectos como sea posible, con el objetivo de que las diferencias lingüísticas sean atribuibles al estatus de traducción en lugar de variables externas o no controladas (Laviosa 2002). La comparabilidad ha de estar presente en el mayor grado posible (Olohan 2004: 42), siempre teniendo en cuenta las limitaciones existentes o lo que sea práctico o posible (ibid). Al igual que la representatividad, el grado de comparabilidad no se concibe como un valor absoluto sino como el mayor que garantice un grado aceptable (Bernardini y Baroni 2003). En el establecimiento de la comparabilidad los investigadores apuntan:

Original a translations address the same audience, deal with virtually the same topics, conform to the same editorial policies, it would appear to guarantee an acceptable level of comparability (Bernardini y Baroni 2003: 2).

Dentro de la comparabilidad, se plantea además la necesidad de que los textos sean sincrónicos, es decir, que se hayan traducido en un marco temporal similar<sup>267</sup> (ibid). En nuestro caso, la descarga del corpus se produjo de manera sincrónica en dos días en un mismo año 2006.

---

<sup>267</sup> Sobre este aspecto se ha criticado la validez del English Comparable Corpus (ECC) debido a la presencia de traducciones diacrónicas o a factores socioculturales en la traducción literaria (Bernardini y Baroni 2003).

El establecimiento de la comparabilidad en nuestra investigación se produce por la recopilación de textos pertenecientes a un mismo género textual, el sitio web corporativo tradicional, con unos rasgos de caracterización extratextual similares. Esta similitud en los parámetros de caracterización que cumplen los textos de ambas secciones asegura el máximo grado de comparabilidad para nuestra investigación. Los sitios web originales y localizados poseen una misma funcionalidad, una misma situación comunicativa con una audiencia y perfil del emisor similares, así como un contexto sociocultural de recepción similar. No obstante, existen divergencias en los rasgos de caracterización del género entre ambas secciones, tal como el distinto contexto de producción, entorno sociocultural de emisión, o los posibles rasgos convencionales en que difieran, si bien éstos son precisamente el objeto de estudio de la presente tesis.

Las investigaciones en Lingüística de corpus indican la preferencia por incluir **textos completos**<sup>268</sup>, puesto que hay características lingüísticas que raramente se distribuyen de manera igualitaria a lo largo del texto y pueden tener baja o alta representación en extractos (Kenny 20001; Baker 1995: 240, Biber 1994: 389; Sinclair 1991: 19). Como ya establecimos en el capítulo 3, el texto web o hipertexto lo constituye el sitio web completo. Si analizamos varias investigaciones previas sobre textos web, éstas incluyen exclusivamente páginas web aisladas (Fernández 2003; Ruiz 2003; Biber 2004; Rennau 2004; Santini 2005a), o por motivos metodológicos se excluyen páginas de menos de 200 palabras (Biber 2004)<sup>269</sup>. En nuestra opinión, desde el punto de vista del proceso de traducción y siguiendo nuestra definición de texto en localización (véase 3.4.1), las páginas web no constituirían textos completos sino subtextos. Además, una página web depende de las relaciones de coherencia y cohesión que proporciona el sitio web para

---

<sup>268</sup> Existen además corpus de muestras textuales, como el

<sup>269</sup> Aun cuando uno de las principales características de las páginas web bien escritas es la brevedad y concisión (Krug 2005, McGovern *et al.* 2002). Como ejemplo, la media de palabras por página web el componente original de nuestro es de 258,8.

completar su textualidad<sup>270</sup>. En cualquier caso, hemos de reseñar que la inclusión de textos completos puede distorsionar las estadísticas en análisis léxicos (Kenny 2001). En parte, ésta fue una de las variables que debimos controlar debido a la repetición léxica en el texto «interfaz» como menús de navegación o los atributos html que describen las páginas web (véase 9.2.3).

La **estandarización** o convencionalización (Pérez Hernández 2002) constituye otro criterio de importancia, de ahí que excluyéramos sitios web en construcción o que no cumplieran ciertas convenciones básicas, como la presencia de un menú de navegación<sup>271</sup>. La decisión de excluir los textos no convencionalizados quedaría respaldada por estudios previos sobre tipologías textuales (Göpferich 1995; Werlich 1975). El concepto de estandarización está relacionado con el de **calidad**, sobre todo por el acercamiento a las expectativas del receptor (Nobs 2005: 26-29), y puesto que los emisores son empresas de reconocido prestigio, hemos de asumir preliminarmente que las traducciones han sido realizadas por expertos. En este sentido, la «experticidad» del traductor afecta tanto al proceso como al producto (PACTE 2000, 2005, 2007). En principio, hemos de presuponer que los textos recopilados son representativos del género objeto de estudio y reflejan las distintas restricciones dadas por sus contextos de producción (Baker 1999), entre las que se encuentra la localización como actividad económica o los problemas relacionado con los plazos de tiempo o los recursos disponibles (Sager 1989: 91).

La **tipología** de corpus es el último criterio presentado por Pérez Hernández (2002). Dentro de la tipología en Traductología adoptada (Laviosa 2002), el corpus recopilado se corresponde con un corpus comparable monolingüe español

---

<sup>270</sup> En el contexto de la situación comunicativa web, una página que indicara pertenecer a [www.telefonica.es](http://www.telefonica.es), aun cuando imitara el entorno gráfico y tipológico de la misma, no se aceptaría como un texto creíble sin las relaciones por hipervínculos con el resto del sitio web de la empresa.

<sup>271</sup> Cabe destacar que el diseño de estas páginas web se basa en las recomendaciones del consorcio W3C para la creación de las mismas, y todas incluyen los elementos básicos de navegación y organización indispensables y comunes a todas las páginas web accesibles.

multimedia sincrónico y especializado de textos completos con una sección original y otra traducida.

### **7.2.2. Comparabilidad y representatividad en el proceso de recopilación**

El proceso de recopilación del corpus se produjo en dos fases sucesivas y complementarias. La recopilación comenzó con un corpus representativo de sitios web corporativos españoles sobre el que se realizó un primer estudio descriptivo (Jiménez 2008c, 2006a). Siete meses más tarde se produjo la segunda fase, que culminó con la recopilación el corpus de sitios web localizados, que junto al primero conformaría nuestro corpus monolingüe comparable.

El objetivo de la primera fase se dirigió a asegurar la recopilación de un corpus representativo de la totalidad de sitios web corporativos en la *locale* es-ES. Siguiendo la investigación sobre registros web de Biber (2004), usamos el directorio de sitios web corporativos de [www.google.com](http://www.google.com) para el proceso de selección textual. Dentro de directorio, usamos la categoría [World> Español> Regional> Europa> España> Economía y Negocios](#)<sup>272</sup>. A continuación, y puesto que se pretende realizar un estudio del género web como forma textual convencionalizada, elegimos el primer sitio web de cada subdirectorío siempre que fuera una creación original en español y no el producto de un proceso de localización (véase 7.2.3). En la recopilación se excluyeron varios subdirectoríos, como «Compras», ya los sitios web corporativos cuyo objetivo sea el comercio electrónico representan un género diferente (véase 5.1.2.1). También excluimos todo subdirectorío que no incluyera el género web de estudio, como «Asociaciones», «Noticias y publicaciones», «Organizaciones» o «Autoridades».

---

<sup>272</sup> El directorio de Google.com se basa en el proyecto *Openproject* en el que los sitios web se catalogan por medio de editores voluntarios y aplicando las funcionalidades de Google.com para un refinamiento de los mismos.

Procedimos a seleccionar el primer sitio dentro de cada subdirectorio siempre que respondiera a una serie de criterios preestablecidos (véase 7.2.3). En caso de que éstos no se cumplieran, seleccionamos en el mismo subdirectorio el siguiente sitio web que se ajustara a estos criterios. En su conjunto, este procedimiento garantizó que diéramos cabida a la mayor variedad textual posible dentro de la población textual seleccionada, y defendemos que el resultado constituye una muestra representativa de los sitios web corporativos más usados en España de todas las áreas económicas.

El directorio de Google se basa en el directorio *Open Source Directory*<sup>273</sup> de Netscape y, a su vez, Google ordena los sitios web siguiendo su tecnología *Pagerank*<sup>274</sup>, que depende del uso y relevancia de cada sitio. Argumentamos que este procedimiento de selección en el que priman los sitios web más usados en todas las áreas económicas garantiza la representatividad. Los sitios web menos usados suelen tener unas características no convencionalizadas y pueden presentar un lenguaje más idiolectal, de ahí que no los incluyéramos (Göpferich 1995)<sup>275</sup>. Este corpus original se constituye en el corpus textual de referencia hacia el que los sitios web localizados se han de parecer según los objetivos del sector de localización (LISA 2003). Esta representación de sitios web corporativos a todos los niveles asegurará que las convenciones de género que observemos se han establecido ampliamente en este contexto sociocultural específico.

El proceso de selección y recopilación de la sección localizada se produjo tras la recopilación y análisis descriptivo del corpus de sitios web corporativos españoles (Jiménez 2008c, 2006). La primera decisión metodológica fue seleccionar un sólo

---

<sup>273</sup> <http://dmoz.org/>

<sup>274</sup> Para una descripción de la tecnología *Pagerank*: [www.http://www.google.es/intl/es/why\\_use.html](http://www.google.es/intl/es/why_use.html)

<sup>275</sup> Göpferich (1995) no se introduce los textos no-convencionales puesto que representan un uso del lenguaje idiolectal y por lo tanto no se les puede abstractizar de manera suficiente para que quepan en una tipología textual. Werlich (1975: 55) también los excluye.



contexto sociocultural de procedencia de los sitios web, el estadounidense<sup>276</sup>. Esta decisión fue motivada por el afán de reducir al mínimo la introducción de variables externas debidas a diferencias en los contextos socioculturales de los TO. Seleccionamos este contexto sociocultural por la mayor presencia de empresas estadounidenses entre las mayores empresas del mundo (Forbes 2007), un aspecto nos garantizaría una mayor probabilidad de encontrar localizaciones dirigidas exclusivamente a la *locale* es-ES. Nuestro análisis en una fase previa demostró que las medianas y pequeñas estadounidenses se decantan por localizaciones en español «internacional» o «neutro»<sup>277</sup> y, consecuentemente, no se correspondería con una variedad claramente definida (García Izquierdo 2006; Castro 1996).

Para el proceso de selección de los sitios localizados utilizamos la lista *Forbes en Internet* ([www.forbes.com/lists/](http://www.forbes.com/lists/))<sup>278</sup>, en la que examinamos los sitios web de las 650 mayores empresas estadounidenses públicas y privadas. El objetivo de esta fase preliminar consistió en identificar localizaciones exclusivas para España. En ciertos casos sólo identificamos un sitio web localizado de una filial o de un producto, y éste se recopiló en representación de la empresa en su conjunto. En total, identificamos un sitio web dirigido específicamente a España en el 20,9% (n. 136) de empresas, aunque sólo el 17,53% (n. 114) son el resultado de producto de localización-traducción dirigido a España. El resto son producciones locales en España, 3,38% (n.22) y, consecuentemente, no son el resultado de un proceso de localización. Estos sitios web responderían a un modelo de localización descentralizado (Yunker 2003), en el que sitio web para cada región se desarrolla y

---

<sup>276</sup> El origen de Internet en este ámbito geográfico supuso que en sus inicios el inglés se convirtiera en la *lingua franca* en este medio (LISA, 2004; Tercedor, 2005).

<sup>277</sup> García Izquierdo (2006) se decanta por denominar esta variedad «español neutro» (2006: 152), aunque apunta que esta variedad se ha denominado español «internacional», «castellano general», «español común» o «español estándar».

<sup>278</sup> Otras investigaciones sobre sitios web corporativos se decantan por la lista Fortune 500 (Liu *et al.* 1997; Robbins y Stylianou 2003), pero la misma se consideró reducida para nuestros propósitos dirigidos a conseguir un corpus localizado de similares dimensiones al no traducido recopilado anteriormente (Baker 1995).

controla in situ en cada región, si bien se mantiene una identidad visual corporativa común.

Dado que esta sección del corpus se recopiló en una segunda fase, el objetivo de este proceso de selección fue contar con un corpus comparable que asegurara un número de páginas lo más similar posible entre ambas secciones (Baker 1995: 234)<sup>279</sup>, de ahí que nos limitáramos a examinar 650 sitios. Para llegar a este punto realizamos diversas pruebas piloto de descarga hasta que decidimos que la composición de la sección traducida del corpus era comparable a la original. Necesitamos realizar este proceso ya que, como apuntamos anteriormente, la lectura hipertextual se realiza sin un conocimiento de los límites reales del hipertexto (*informational shortsightedness*), que influye en su procesamiento cognitivo (Conklin 1987), por lo que durante la fase de selección del corpus fue imposible conocer a priori el número de páginas de cada sitio web seleccionado.

Como resumen, asumimos que ambos corpus son representativos de las poblaciones textuales que pretendimos representar y comportan el mayor grado de comparabilidad posible dentro del marco metodológico empleado, los objetivos de la presente tesis y los medios a nuestro alcance (Olohan 2004). En cualquier caso, no debemos olvidar que los sitios web de las empresas localizadas se corresponden exclusivamente a un posible subgénero, el «sitio web corporativos de grandes empresas», que se puede contraponer los subgéneros «sitios web de las medianas empresas» o «sitio web de pequeñas empresas (PYMES)» y que sí que se encuentran representados en el subcorpus original. En nuestro análisis posterior tendremos en cuenta este aspecto que, como mencionamos en el párrafo anterior, fue motivado metodológicamente para garantizar que pudiéramos contar con un número suficiente de localizaciones de sitios web corporativos para España desde un mismo contexto sociocultural. Este estudio sería impracticable en caso de que

---

<sup>279</sup> La autora señala que ambos componente del corpus deben ser de extensión similar «both corpora should be [...] of comparable length» (Baker 1995: 234).

pretendiéramos incluir localizaciones hacia España en los sitios de medianas y pequeñas empresas estadounidenses.

### **7.2.3. Criterios para asegurar que se trata de un texto original o localizado**

Dado que nuestro estudio se sitúa en la marabunta de la web, nos fue indispensable elaborar unos criterios que garantizaran la posible adjudicación de cada sitio web a (1) una producción original, o (2) al producto de un proceso de traducción-localización. La concepción de texto traducido-localizado que asumimos no se corresponde exclusivamente con la concepción de Toury (1995) de «assumed translations». El investigador las define como aquellas unidades de habla que se presentan o se consideran como tal en la cultura meta:

All utterances which are presented or regarded as such within the target culture, on no matter what grounds (Toury 1995: 32).

Nuestro objetivo no se centra en delimitar qué constituye una traducción en el sentido más amplio en Traductología como referencia al objeto de estudio de esta disciplina (Toury 1995; Halverson 1998; 1999), sino que desde unos criterios objetivables y explicitables, necesitamos delimitar qué constituye una localización y qué texto no ha de ser considerado como tal.

Para este fin desarrollamos unos criterios empíricos en una investigación previa (Jiménez 2005a) para determinar la naturaleza de sitio web como texto original o natural<sup>280</sup> en la *locale* es-ES (Halverson 1998), o como localización hacia este mismo contexto sociocultural o *locale*. Debemos destacar que su concepción se basa en la teoría de prototipos de Eleanor Rosch (1977, 1978), por lo que los criterios de asignación a una u otra categoría no son rígidos siguiendo el concepto

---

<sup>280</sup> Halverson (1998) indica que deben ser «naturally-occurring language».

de traducción desde los principios cognitivos más amplios propuesto por Halverson (1998, 1999).

En ciertos sitios web pudimos identificar subtextos o páginas producidas originalmente en español junto a otros productos de un proceso de localización, con lo que en ciertos casos la adscripción a la categoría de traducción se realizó siguiendo esta concepción de prototipo, ya que existen casos fronterizos en aquellos sitios web con distinto porcentaje de subtextos producidos localmente en España<sup>281</sup>. Estos criterios pueden ser aplicados a la evaluación de fuentes de documentación en Internet en el contexto profesional o en la didáctica de la traducción (Palomares 2000; Pearson 1999).

Los criterios que se detallan a continuación se usaron para determinar si el texto no constituye una creación original en la *locale* de estudio, es-ES, y se observaron en orden descendente:

- La empresa multinacional tiene su sede principal en un país en el que el español no es idioma oficial y usa el inglés como *lingua franca*.
- La posibilidad de elegir idioma o sitio global en España desde la página principal en español, tanto en menú desplegable como en menú gráfico con banderas<sup>282</sup>.
- La ausencia de caracteres exclusivos del español como ¡, ¿, ñ.
- La presencia de comentarios<sup>283</sup> y variables en inglés u otro idioma en el archivo fuente formato .txt en la página en español. Los comentarios de los desarrolladores y las variables de los *scripts* son de gran utilidad para

---

<sup>281</sup> Como el caso de los sitios localizados que no obstante presentan un bloque comunicativo legal redactado en España (véase 10.3)

<sup>282</sup> La W3C desaconseja el uso de banderas para la elección del lenguaje debido a que, como el español, se hablan la misma lengua en diferentes países.

<sup>283</sup> Ejemplo de línea de comentarios en una página traducida el español de Samsung [n <http://www.samsung.es/cgi-bin/sesa/main/samsunges.jsp>]  
</TABLE></DIV><!-- end Main Promo --><!-- end Main Promo -->

identificar el carácter de original o traducción de una página web. Por ejemplo, la página española de la multinacional Adecco<sup>284</sup> incluye en sus *javascript* variables en español, con lo que asumimos que la página se desarrolló originalmente en esta lengua.

```
cerrado=true;
function despe(){
if(cerrado){
bmen.frenadoTo(null,67,4,5,'cerrado=false');
}
else bmen.frenadoTo(null,48,4,5,'cerrado=true;')
```

Estas variables en español también pueden aparecer en la dirección o URL de la página web, sobre todo en las páginas dinámicas que se crean con servidores activos. Como ejemplo:

```
http://www.hasbroiberia.com/WEB/Crea_Frames.asp?pagina2=FormulariosContacto/formcontacto.as
p&pagina=Contactos&opcion=7&htmAsp=asp
```

En esta dirección observamos como se incluye el español en la estructura del servidor activo con «**creaframe**», «**página**» u «**opción**».

- La presencia de subtítulos en inglés u otra lengua en una página web en español. A menudo uno o varios marcos se quedan sin traducir.
- La existencia de una página web original en una lengua, junto con una página en español que tenga exactamente la misma composición. En el caso más común, el inglés, la página original tiene el dominio «.com» y la traducción al español «.es».
- Y por último, tomamos como creación original los sitios de multinacionales elaborados mediante un proceso de globalización web descentralizado (véase 3.3.2) y, por lo tanto, si presentan un lenguaje publicitario claramente destinado a la cultura de España. Como ejemplo podemos poner a [www.adecco.es](http://www.adecco.es), cuya página web en español, aun resultando muy parecida a la página principal en inglés [www.adecco.com](http://www.adecco.com), contiene una tipología textual que apunta a que su adscripción a una creación original en

---

<sup>284</sup> <http://www.adecco.es/>

esta *locale*. No obstante, debido a los diferentes niveles de localización, es posible que un sitio localizado a su nivel máximo según Signh y Pereira (véase 2.7), pueda confundirse con una creación original en español. En ese caso es necesario observar el código fuente y encontrar los comentarios de los desarrolladores `<!>`, o analizar la composición de los URL para observar si el sitio web es el producto de un proceso de localización.

### 7.3. Descarga del corpus

La selección textual del corpus original se realizó entre los días 6 y 7 de marzo del 2006. La descarga final del corpus se realizó sincrónicamente con la herramienta informática Htrack<sup>285</sup> en un solo día, el 4 de mayo del 2006<sup>286</sup>. La descarga del corpus traducido se produjo de manera sincrónica el 21 de noviembre del 2006.

En ambas descargas se usaron tres ordenadores con una conexión a Internet por cable de 100T, dividiendo los sitios web en tres grupos. Los estudios de Lingüística de corpus suelen ser mayormente sincrónicos para una caracterización en un momento temporal determinado (Kenny 2001: 106). Además, el carácter sincrónico se considera una variable esencial en el establecimiento de la comparabilidad en este tipo de corpus (Baroni y Bernardini 2003; Baker 2004). Nuestra descarga en un solo día consiguió almacenar una representación sincrónica de cada población textual objeto de estudio.

Al recopilar sitios web completos se descargaron páginas web en diferentes extensiones como: .html, .htm, .xml, .dhtml, xhtml, .jsp (*Java server pages*), .aspx (*active server pages*), etc., que el programa Htrack almacena exclusivamente en formato .html. Adicionalmente, se descargaron y formaron parte del corpus todas

---

<sup>285</sup> El programa se puede obtener a través de [www.htrack.com](http://www.htrack.com).

<sup>286</sup> Se realizaron 10 intentos previos en los que se estudiaron los parámetros a seguir para conseguir la descarga sincrónica de un mayor número de sitios web.

las imágenes, tablas, tipografía, etc., puesto que son parte integrante del texto (Göpferich 1995). Su presencia añade una dimensión adicional «multimedia» a nuestro corpus (Laviosa 2002).

Nos fue inevitable excluir los archivos en formato .pdf debido a limitaciones de almacenamiento y descarga. Estos archivos constituyen géneros *extant* (véase 5.5.6.1), géneros que existen paralelamente de forma impresa y cuya presencia en la web se destina a imprimir documentación distribuida por Internet (Nielsen 2003). Por el mismo motivo decidimos excluir las presentaciones en PowerPoint, .ppt, o documentos en formato Word, .doc o .rtf. Finalmente, y también por razones de almacenamiento, excluimos los archivos multimedia como los vídeos, los archivos de sonido y los archivos comprimidos<sup>287</sup>.

#### **7.4. Composición final del corpus comparable**

El **corpus original** final se descargó el 4 de mayo del 2006. Se compone de 172 sitios web y 19.102 páginas web, con una media de 111,50 páginas por sitio web. A su vez, cada página contiene una media de 453,34 palabras, aunque en estas cifras no se incluyen las palabras codificadas en imágenes o archivos Flash al ser imposible contabilizarlas con nuestras herramientas informáticas. A su vez, la sección original se compone de 56.961 archivos, 331,16 por sitio web. El total de bytes es de 2.604.756.488 o 2,42587 Gigas. La media sería 15.143.933.06 bytes por sitio web, equivalente a 14,44 Megas. Como señalamos anteriormente, nuestra muestra representativa consistía en 248 sitios web seleccionados entre todas las áreas económicas. En la descarga definitiva nos fue posible descargar el 29,8% de sitios web (n. 74).

---

<sup>287</sup> Archivos de vídeo y sonido excluidos: .wma, .wmv, .mpeg, .rm, .avi, .asf, .mov, .wav, .vob. Archivos de compresión: .rar, .tar, .zip, .tgz

Por su parte, el **corpus localizado** se descargó el 21 de noviembre del 2006. En un principio se seleccionaron 107 sitios web localizados para España entre las mayores empresas estadounidenses. En el proceso de recopilación sólo pudimos descargar 95, con lo que se descartaron el 11,21% de sitios preseleccionados (n. 12). Esta sección localizada consta de 21.322 páginas web, con una media de 224,93 páginas por sitio. A su vez, se compone de 54.526 archivos o 576,95 por sitio. Su tamaño total es 1.170.133.080 bytes o 1,089 Gigas. La media por sitio web es 11,7423 Megas o 12.317.190,31 bytes por sitio. La composición global del corpus queda recogida en la siguiente tabla, en la que hemos incluido las estadísticas sobre aquellos elementos textuales de interés en el proceso de traducción.

<b>Tipo</b>	<b>Corpus original</b> 172 webs españoles		<b>Corpus localizado</b> 95 webs estadounidenses	
	<b>Total</b>	<b>Media</b>	<b>Total</b>	<b>Media</b>
<b>Páginas web</b>	19.102	111,5 Páginas/sitio	21.322	224,93 Páginas/sitio
<b><u>Palabras en Cuerpo de la página</u></b>	4.945.103	258,87 Palabras/Pág.	8.871.512	416,07 Palabras/Pág.
<b>Palabras total</b>	8.659.856	453,34 Palabras/Pág.	12.562.894	589,50 Palabras/Pág.
<b>Palabras en &lt;Meta&gt;</b>	932.255	48,80 Palabras	742.112	34,80 Palabras/Pág.
<b>Palabras en &lt;alt&gt;</b>	290.031	15,18 Palabras/Pág.	395.275	18,53 Palabras/Pág.
<b>Palabras en &lt;input&gt; y &lt;select&gt;</b>	1.559.949	81,66 Palabras/Pág.	897.937	42,11 Palabras/Pág.
<b>Palabras en Scripts</b>	254.069	13,30 Palabras/Pág.	580.676	27,23 Palabras/Pág.
<b>Palabras en &lt;OnMouse&gt;</b>	678.449	35,51 Palabras/Pág.	1.075.382	50,43 Palabras/Pág.
<b>Número de enlaces HTML</b>	726.443	38,02 Enlaces/Pág.	1.061.227	49,77 Enlaces/Pág.
<b>Número de Imágenes</b>	676.297	35,40 imágenes/Pág.	1001915	46,98 Imágenes/Pág.

**Tabla 4. Composición final del corpus y análisis comparativo. Estadísticas elaboradas con Webbudget.**

Si procedemos a analizar contrastivamente estas estadísticas, la primera diferencia significativa es la media de páginas web por sitio: el corpus español cuenta con 111,50 páginas y una media de 453,34 palabras por página, mientras que la media de páginas web por sitio web localizado es 224,93 y una media de 589,50 palabras



por página. Los sitios web localizados constan de una cantidad mayor de páginas web, casi al doble, y además la media de palabras por página es 136,16 superior.

Por otro lado, el corpus original muestra una mayor presencia de archivos en formato Macromedia Flash<sup>288</sup>, 4,29 por sitio web (n. 739), frente 3,5 archivos por sitio en el corpus localizado (n. 339). Este hecho puede influir en la mayor media de palabras por página, ya que las palabras inmersas en estos archivos no forman parte de las estadísticas finales. Tal como describimos en el apartado 7.5.2, los expertos en usabilidad desaconsejan la inclusión de presentaciones Flash, especialmente en menús de navegación (Nielsen 2000; Nielsen y Tahir 2002; Nielsen y Loranger 2006; Brinck *et al.* 2003)<sup>289</sup>, con lo que podríamos concluir que los sitios web localizados se ajustan mejor a estas pautas de desarrollo web. De entre las posibles razones que podrían explicar este hecho, nos inclinamos a justificarlo en la mayor complicación e inversión económica que requiere la localización de elementos en formato Flash, de ahí que se reduzca su presencia en los sitios a localizar (Yunker 2003; Jeney 2007).

Otro punto de interés es el mayor tamaño global en bytes del corpus original. En investigaciones previas el tamaño global de un sitio web se incluye como variable contrastiva asociada a la dimensión «forma»<sup>290</sup> del género digital (Kennedy y Shepherd 2005; Robbin y Stylianou 2003). A pesar de que ambos contienen una cantidad similar de páginas web, el tamaño global en disco difiere considerablemente, 2,425 Gigas en el original y 1,089 Gigas en localizado. En este sentido, las guías de desarrollo recomiendan reducir al máximo el tamaño de los sitios web para que los usuarios puedan descargarlos en el menor tiempo posible

---

<sup>288</sup> Con extensión <.swf>.

<sup>289</sup> Nielsen y Loranger (2006: 249) señalan que el texto en movimiento en general se asocia actualmente con publicidad y suele ser ignorado por los usuarios de la web.

<sup>290</sup> Como se señaló en el apartado 5.5.6.1, las investigaciones sobre género digitales previas se fundamentan en la triple vertiente de forma, contenido y funcionalidad.

(Price y Price 2002)<sup>291</sup>. Los sitios web localizados presentan un menor tamaño, ajustándose mejor estas recomendaciones relativas a la reducción y optimización del tiempo de descarga<sup>292</sup> siempre manteniendo un alto nivel de calidad (Lynch y Horton 2002; Yunker 2003). El mayor tamaño del corpus original lo asociamos a una mayor presencia de elementos en Flash y de imágenes no optimizadas para la web.

Estos resultados contrastan con los obtenidos en la investigación previa sobre un subgénero de sitios web corporativos de Rennau (2004: 307). La investigadora afirma que los sitios españoles analizados no abusan de imágenes o presentaciones en movimiento que ralenticen su descarga en contraste con los de otros países. Nuestros resultados no apuntarían específicamente a un posible abuso de estos elementos en los sitios web españoles, sino más bien a su posible optimización de acuerdo con los modelos que emplean las empresas que localizan sus sitios y que tratan con una multitud de versiones. Nuestro análisis sólo muestra que el tamaño medio de los sitios españoles es relativamente superior a los localizados (50%), y hecho que podría apuntar a un menor grado de profesionalización en su desarrollo.

## 7.5. Limitaciones de la recopilación

En las pruebas preliminares de descarga del corpus se hicieron patentes ciertas limitaciones debidas a la descarga sincrónica y a las particularidades del texto web. Podemos agruparlas en dos grupos: (a) las limitaciones debidas al volumen de textos recopilados, y (b) aquellas que surgen de las características del software empleado en combinación con ciertas características de los textos web.

---

<sup>291</sup> Esta variable en diseño web se denomina velocidad o *speed*, definida como «speed feature relates to the amount of time that visitors to the site have to wait» (Robbin *et al.* 2003: 207).

<sup>292</sup> Como ejemplo, se señala un intervalo de diez segundos como el máximo para la descarga de una página web (Nielsen y Loranger 2006).

1. El volumen de textos recopilados implica la imposibilidad de descargar ciertos componentes del texto, como los archivos .pdf, los de vídeo, sonido, o los archivos comprimidos. Partimos de la base de que las investigaciones de corpus han de contar con conjuntos textuales de gran tamaño (Sinclair 1991), de ahí que nuestro objetivo de descargar nuestro corpus sincrónicamente con los medios disponibles tuvo como contrapartida la eliminación de estos elementos multimedia.
2. En segundo lugar, la herramienta informática seleccionada para la descarga se basa en el análisis sintáctico del lenguaje de marcaje html u otro código para detectar enlaces internos<sup>293</sup> y almacenar una copia de las páginas hacia los que se dirigen. Los problemas en nuestro estudio surgieron por: los *scripts* que presentan páginas web dinámicas (véase 7.5.1) y, consecuentemente, distintas dependiendo de la sesión<sup>294</sup>, al uso de marcos o *frames* en las páginas (véase 7.5.1), y a la existencia de Macromedia Flash en un gran número de sitios web (véase 7.5.2). En estos tres casos, el análisis sintáctico de las páginas web no detecta los enlaces, de ahí que el programa no pueda almacenar el sitio web completo. Estas limitaciones supusieron que en el corpus original descartamos el 29,8% de sitios web (n. 74 de un total de 248 sitios inicialmente seleccionados), y 16,52% en la sección localizada (n. 19 de sitios localizados preseleccionados). De igual modo, fue imposible acceder y descargar las *intranets* o zonas reservadas que requieran autorización o la suscripción al sitio web.

### 7.5.1. Los marcos y la arquitectura servidor-usuario

En la actualidad, gran cantidad de sitios web se generan dinámicamente a través de un servidor activo que produce una página web **virtual** (véase 5.5.6.1), es decir,

---

<sup>293</sup> Los enlaces en hipertextos pueden ser internos, hacia el mismo sitio web, o externos, hacia otros sitios web (ver Janoschka 2003: 178).

<sup>294</sup> En este caso suelen ser *Active Server Webpages*.

páginas potencialmente distintas dependiendo del momento temporal o las especificaciones o preferencias del usuario. Las páginas se generan en servidores web activos, pudiendo reconocerse por su extensión: .asp. .do. .cgi o .aspx (Mata 2005). Estos formatos surgen con la aparición de componentes dinámicos producidos con lenguajes como JavaScript, Java Applets, VBScript, ASO, CGI, ActiveX, Pearl ShockWave, VRML, Flash, etc. Representan además una realización del género web **virtual** (Shepherd y Watters 1998), es decir, textos que cuentan con una realización única, como serían los resultados de una búsqueda.

En ciertos casos, dependiendo del lenguaje de programación y de las funcionalidades de Httrack, se hace imposible descargar un sitio web puesto que no existe un único sitio web en un momento dado como tal. Por ejemplo, [www.iberia.es](http://www.iberia.es) se basa en un programa Java controlado por el servidor Web<sup>295</sup> que dota a las páginas web de la extensión «.do» y apunta a la existencia de un programa Java que funciona en un servidor activo. Este aspecto imposibilita almacenar estáticamente un sitio en un corpus. Ahora bien, se podrían estudiar las páginas web **virtuales** una por una, pero como mencionamos anteriormente, no consideramos que constituyan un texto unitario.

Entre las grandes empresas se están imponiendo estos lenguajes de programación junto con los programas de Global Content Management (LISA 2006). Este aspecto dificultará aún más las investigaciones futuras sobre la textualidad web de sitios corporativos completos. En el corpus original fue imposible descargar el 8,6% (n. 20) de sitios preseleccionados. Por su parte, en el corpus traducido fue imposible descargar el 10,52% (n. 12). Una posible solución a este problema consistiría en descargar el sitio web a partir del mapa web, un método que utilizamos en cuatro sitios web del corpus localizado<sup>296</sup>.

---

<sup>295</sup> Normalmente IBM WebSphere.

<sup>296</sup> [http://www.wyeth.es/mapa\\_sitio.htm](http://www.wyeth.es/mapa_sitio.htm), [www.lexmark.es](http://www.lexmark.es), <http://www.starbucks.es/es-es/>, y <http://www.stryker.es/st-index>

La utilización de marcos o *frames* se basa en incluir varias páginas web independientes dentro de una misma pantalla. Cada marco en la página web funciona como una página independiente, con lo que se rompe la premisa de que cada página sea una unidad de información, navegación y almacenamiento. Nielsen indica la necesidad de prescindir de los marcos debido a los problemas que producen en la categorización, navegación o búsqueda<sup>297</sup> (Nielsen 2000: 85-87). Este hecho dificulta el análisis que realiza la herramienta Httrack, puesto que la dirección o URL de la página web global es distinta a la de las páginas de cada marco y, por lo tanto, el programa considera que pertenecen a otros sitios web. Los marcos presentaron problemas sólo en el corpus original, en el 4,43% (n. 11) de casos, algunos de ellos combinados con Flash. Este hecho no supuso problema alguno en la sección localizada del corpus.

### **7.5.2. Macromedia Flash**

Durante años, el uso de Macromedia Flash ha ido aumentando de manera exponencial, ya que proporciona una experiencia multimedia, animada y de un alto grado de interactividad. Estas presentaciones en Flash han llegado a investigarse independientemente desde una perspectiva del género (Paolillo *et al.* 2006). El uso de Flash en su formato FLA (*Flash Movie Authoring File*), convertido para distribuirse por Internet en formato SWF (*Small Web Format/ShockWave Flash*), puede presentar los siguientes problemas en general: (a) el equipo informático del usuario puede no contar con este componente instalado, (b) su inclusión puede molestar, distraer o confundir al usuario durante la sesión (Brinck *et al.* 2002), y (c) una conexión lenta puede llevar al usuario a abandonar el sitio

---

<sup>297</sup> Véase también Brinck *et al.* (2002: 361-364) para un tratamiento de las ventajas y desventajas del uso de marcos en páginas web.

web<sup>298</sup>. Nielsen (2000) indica que se debe restringir el uso de Flash debido a que los usuarios acceden a los sitios web con el propósito claro de obtener información, y estas presentaciones y animaciones le distraen de este propósito básico.

Si atendemos a su traducción, el formato Macromedia Flash se aleja de la simplicidad del lenguaje html o xml, y su proceso de traducción se asemeja más a la localización de software. En el manual de localización de Macromedia se señala la necesidad de separar el texto del código en archivos de recursos, siendo necesaria su compilación para su uso (Macromedia 2004). Para la traducción de estos archivos es necesario recurrir a paquetes de software que se especializan en su traducción, tales como Avral Tramigo<sup>299</sup>, o contar con los archivos originales sin ensamblar.

El uso de Flash en el corpus original se suele centrar en anuncios web, en presentaciones iniciales en páginas web o *splash screens*, y como método de navegación dentro de las páginas web. Estos dos últimos casos se desaconsejan por los expertos en usabilidad y, de hecho, esta práctica casi ha desaparecido<sup>300</sup> (Nielsen y Loranger 2006: 90). En nuestra recopilación, estos dos casos presentaron problemas puesto que el proceso de análisis sintáctico del programa Htrack no detecta los enlaces internos en el seno de animaciones Flash y, consecuentemente, se hace imposible la descarga. Un ejemplo sería el sitio web preseleccionado [www.lechepascual.es](http://www.lechepascual.es) que usa Flash en su menú de navegación. Para intentar solucionar el problema eliminamos las presentaciones Flash de inicio<sup>301</sup>, incluyendo las que sirven como selectores del idioma. En el corpus original el

---

<sup>298</sup> Los manuales de desarrollo recomiendan incluir la opción de «saltar la introducción» o «pasar a información» e incluir una barra que indique el estado de la descarga de la presentación.

<sup>299</sup> Para información sobre el paquete de software: <http://www.avral.com/tramigo/>.

<sup>300</sup> Yunker (2003) recomienda en caso de su uso: (a) no incluir introducciones Flash innecesarias, (b) incluir una interactividad lógica y estructurada, (c) no abusar de las animaciones, (d) no emplear sonidos innecesarios (e) reducir el tamaño para que la presentación no suponga un problema para los usuarios con conexión telefónica.

<sup>301</sup> A pesar de que según nuestra concepción del texto web, estas páginas forman parte integrante del texto global.

10,46% (n. 18) de sitios web no se pudieron descargar. En el corpus traducido sólo el 3,5% de los sitios web seleccionados no se pudieron descargar, una cifra menor que reflejaría una mayor concienciación de los emisores de sitios web localizados hacia las recomendaciones desde el campo de la usabilidad (Nielsen y Loranger 2006).

El uso de Flash presenta además problemas por cuanto a la accesibilidad para personas con discapacidades, ya que (1) no incluye atributos básicos como <alt>, (2) el usuario no puede cambiar el tamaño de letra desde su navegador en caso de que sea necesario, o (3) no puede ser usado por lectores de páginas para personas con discapacidades visuales (Tercedor y Jiménez 2007).

Finalmente, en caso de usar este formato se recomienda presentar dos versiones del sitio web, una en formato html y otra en Flash (Nielsen 2000). Nuestro estudio demuestra que esta prescripción no encuentra su reflejo en la práctica real en España, puesto que sólo identificamos un sitio web con ambas versiones y la posibilidad de seleccionarlas: [www.culture02.com](http://www.culture02.com)<sup>302</sup>. Esta doble opción no se identificó en el corpus localizado.

### **7.5.3. Sitios web eliminados por distorsionar las estadísticas**

En el proceso de recopilación de corpus de sitios web es difícil conocer a priori la extensión del hipertexto seleccionado, en parte por la falta de apreciación de los límites de los hipertextos en su procesamiento (Conklin 1987). Ya indicamos que concebimos el proceso de diseño y compilación de nuestro corpus como un proceso cíclico (Biber 1993: 256), de ahí que decidiéramos eliminar dos sitios web durante el proceso de análisis. Se trata de dos sitios web cuyas características distorsionarían las estadísticas: [www.fotolab.es](http://www.fotolab.es) y [www.masson.es](http://www.masson.es). El primero

---

<sup>302</sup> El porcentaje global sería un 0,58% de los sitios web españoles.

cuenta con 22.103 páginas web totales y el segundo con 4.082. A pesar de ser sitios web corporativos, el objetivo final de estos sitios web es principalmente la venta *online*, lo que implica la inclusión del catálogo de venta en línea. Ya indicamos anteriormente que los sitios web de compra en línea constituyen un género diferenciado, aunque en ciertos casos la línea de separación entre géneros web en constante evolución es difusa, un hecho que no se observó en un primer momento. Como ejemplo de la posible distorsión que causaría si los incluyéramos en nuestro análisis, la media de páginas por sitio web en nuestro corpus ascendería a 310,5 en lugar de la media actual, 161,99 páginas.

### **7.6. El análisis textual y la elaboración de subcorpus**

Los distintos niveles de procesamiento a los que sometimos nuestros datos, en conjunción con las peculiaridades de corpus web y problemas dados por sus distintos formatos, nos llevaron a elaborar distintos subcorpus comparables en varias fases consecutivas. Los objetivos tras la elaboración de subcorpus se dirigieron a analizar específicamente los distintos bloques comunicativos o segmentos textuales significativos y, en este sentido, la base teórica que proporciona el género textual nos fue inestimable a la hora de recopilar estos subcorpus. Otra de las razones tras la elaboración de subcorpus se debe al uso de ciertas herramientas informáticas que impusieron unas restricciones, de ahí que debiéramos convertir selectivamente el corpus del formato html a formato texto plano (.txt). La Tabla 5 resume las distintas fases de análisis y los subcorpus recopilados.

Durante la primera fase de análisis de textos globales mediante un análisis manual, y con el objetivo de analizar la superestructura global, utilizamos los textos del corpus en formato html, tal y como se recopilaron por la herramienta informática Htrack.



Corpus		Análisis
Fase I	Corpus comparable global en formato HTML	Estadísticas globales con Webbudget Análisis superestructural Análisis contrastivo de la superestructura Tenor digital Función textual global
Fase II	Corpus comparable global en formato .txt y codificación ASCII	Análisis léxico y sintáctico global
Fase III	Subcorpus comparable del bloque comunicativo de contacto (B)	Análisis contrastivo de unidades y subunidades significativas Análisis contrastivo de formas lingüísticas convencionales
	Subcorpus comparable del bloque comunicativo legal (F)	Análisis contrastivo de unidades y subunidades significativas Análisis contrastivo de formas lingüísticas convencionales
	Subcorpus comparable de menús de navegación	Análisis de formas lingüísticas convencionales
	Subcorpus comparable de segmentos textuales en <i>Alt</i>	Análisis léxico contrastivo
	Subcorpus comparable de segmentos textuales en <i>Title</i>	Análisis léxico contrastivo

**Tabla 5. Correlación de las distintas fases de análisis y el uso de corpus o elaboración de subcorpus comparables.**

En una segunda fase de análisis, que consistió en el análisis léxico y sintáctico por medio de la herramienta Wordsmith Tools<sup>303</sup>, se presentaron ciertos problemas debidos a las distintas codificaciones presentes en el corpus, Unicode, ISO-8859-1, y Windows-1252, y a la redundancia léxica que suponen los elementos metatextuales incorporados en los encabezamientos de las páginas web (véase 9.2.3). Estos problemas exigieron la elaboración de un corpus paralelo global en formato texto plano (.txt) a través de la herramienta Detagger, que permitió además la eliminación de los segmentos textuales redundantes en las cabeceras de las páginas web y su conversión a formato ASCII. Este segundo corpus nos permitió realizar un análisis léxico de textos globales. El control de estos problemas es primordial para evitar que los resultados que obtengamos queden sesgados. En esta fase hemos de apuntar que la extracción automática de texto en imágenes y

<sup>303</sup> <http://www.lexically.net/wordsmith/>.

archivos en formato .swf de Flash es imposible con las herramientas de que disponemos, aunque defendemos que dado el tamaño del corpus esta posible deficiencia queda compensada.

En una tercera fase seleccionamos dos bloques comunicativos por su alto grado de convencionalización y su carácter representativo: el bloque comunicativo contacto (B) (véase 10.2) y el bloque comunicativo legal (F) (véase 10.3.). Para llevar a cabo su estudio contrastivo recopilamos manualmente dos subcorpus comparables del bloque contacto y del bloque legal a partir del corpus global. Finalmente, disgregamos dos subcorpus comparables de segmentos textuales de la estructura interna de cada página: los segmentos en los atributos <alt> y <title>. Éstos se extrajeron mediante una combinación de varias herramientas informáticas que diseñamos para tal fin.

## 7.7. Herramientas de análisis

Como último paso en la descripción de nuestra metodología resumimos brevemente las características de las herramientas informáticas que hemos usado. Su selección se produjo tras un análisis de la pertinencia de diversas herramientas para el objetivo de nuestro estudio, que incluyó gran cantidad de pruebas piloto. En todo caso, sus capacidades y limitaciones imponen unas posibilidades y restricciones que hemos de tener en cuenta:

1. *Httrack*<sup>304</sup>. Herramienta informática para la descarga de sitios web completos de acuerdo con unos parámetros definidos por el investigador. Entre los parámetros principales destaca la posibilidad de reestructurar el sitio web o mantener su estructura en el servidor, la selección de qué archivos descargar según su formato, o la posibilidad de reescribir el código para redireccionar la

---

<sup>304</sup> Programa disponible en <http://www.httrack.com/>.

totalidad de los enlaces internos. Por medio de este programa se descargaron de forma sincrónica en un solo día la totalidad de textos de cada sección del corpus. Este programa se basa en el análisis sintáctico de las diferentes páginas web y la descarga de la totalidad de los enlaces internos que encuentra. El programa almacena una copia de cada sitio web respetando, o no, su estructura original.

2. *Wordsmith tools*<sup>305</sup>. Herramienta desarrollada por Mike Scott (1996) para el análisis de elementos intratextuales en el corpus, con la inclusión de un generador de concordancias, analizador de colocaciones o elaboración de estadísticas de frecuencias. El programa incluye a su vez tres herramientas esenciales en el análisis intratextual del corpus (Pérez Hernández 2002):

- *Concordancer*, una herramienta analizadora de concordancias muy completa, que además de las características usuales de los este tipo de herramientas, permite computar colocaciones, análisis de patrones léxicos o *patterns* y agrupaciones de palabras, *cluster*.
- *Wordlist*, un programa que permite obtener listados de palabras incluidas en un corpus según diferentes criterios como alfabéticamente o por frecuencia, junto con estadísticas detalladas sobre la composición del corpus. Las listas de palabras se pueden basar según unidades léxicas simples o en grupos de dos o más palabras. Esta herramienta permite indexar el corpus, y a su vez este índice puede usarse para obtener otros tipos de información, como el índice de información muta o *MI-Score*.
- *KeyWords*, un programa que extrae las «palabras clave» de un corpus, palabras cuya frecuencia es estadísticamente relevante en comparación con la frecuencia de aparición en un corpus comparándolo con un corpus de

---

<sup>305</sup> Programa disponible en: <http://www.lexically.net/wordsmith/>

referencia. Permite además estudiar como las palabras clave se distribuyen por el texto y los enlaces que existen entre una palabra clave y el resto.

3. *Detagger*<sup>306</sup>. Herramienta informática que convierte todo tipo de archivo en formato .html o .htm a formato texto con la posibilidad de excluir selectivamente etiquetas .html o diversos tipos de código. El programa es de gran utilidad debido a la multiplicidad de codificaciones y formatos y lenguajes de programación presentes en el conjunto de los sitios web del corpus. La exclusión del contenido incluido en la cabecera de las páginas web <Head> </Head>, sobre todo los valores de los atributos <Meta> (véase 9.2.3), es fundamental puesto que su inclusión afectaría de manera negativa a los análisis léxicos realizados con Wordsmith Tools.
4. *Webbudget*<sup>307</sup>. Es un paquete de software destinado a la gestión de localización de sitios web. El programa puede realizar estadísticas del corpus separando los distintos elementos textuales y gráficos. El paquete proporciona además un listado individualizado con las estadísticas de cada página web del corpus dentro de la estructura del sitio web.
5. *InfoRapid Search and Replace*<sup>308</sup>. Programa desarrollado por Ingo Straub que permite la búsqueda de palabras o cadenas de manera automática recogiendo en los resultados el contexto de cada unidad encontrada.
6. *Avral Tramigo*<sup>309</sup>. Paquete de software dirigido a la extracción y localización de los elementos textuales en los archivos multimedia Flash.

---

<sup>306</sup> Programa disponible en: <http://www.jafsoft.com/detagger/>

<sup>307</sup> Programa disponible en: <http://www.webbudget.com/>

<sup>308</sup> Programa disponible en: <http://www.inforapid.de/html/searchreplace.htm>

<sup>309</sup> Programa disponible en: <http://www.avral.com/tramigo/>

7. *ExpPrint*<sup>310</sup>. Herramienta que proporciona un listado completo estructurado de la totalidad de carpetas y archivos de todo tipo dentro de un corpus.

---

<sup>310</sup> Programa disponible en: <http://www.jddesign.f2s.com/expprint.htm>

**PARTE III:**

**RESULTADOS**

**Y**

**DISCUSIÓN**



## 8. ANÁLISIS EXTRATEXTUAL DEL GÉNERO *SITIO WEB CORPORATIVO*

### 8.1. Criterios de análisis extratextual

Nuestro análisis extratextual del género *sitio web corporativo* se desarrollará en consonancia con las investigaciones traductológicas previas sobre géneros textuales. El análisis que propondremos se centra en los factores (1) pragmáticos, (2) socioculturales, y (3) los comunicativos (Göpferich 1995; Gamero 1998, 2001; Fernández Sánchez 2004). Estos criterios se corresponden con los rasgos extratextuales en el modelo de caracterización de género presentado por Silvia Gamero (1998, 2001), que se fundamenta en la propuesta previa de Göpferich (1995). Ésta a su vez, desde una perspectiva funcionalista, adopta los tres criterios de análisis de Kuntz (1979)<sup>311</sup>, la función comunicativa, la relación emisor-receptor, y la situación espacio-temporal, que coinciden con los expuestos anteriormente. A estos tres criterios la investigadora alemana Göpferich añade una categoría adicional que denomina «otros» (1995: 201-211). En nuestro caso incluiremos (4) el marco socioprofesional, ya que a nuestro parecer influye de manera determinada en la producción y localización de este tipo de textos.

Nuestro análisis de factores extratextuales combina los enfoques inductivos y deductivos siguiendo a Göpferich (1995: 75). El enfoque inductivo se fundamenta en las observaciones propias sobre el género objeto de estudio y recurrimos a la competencia textual de quien realiza el análisis del género (ibid: 65). El enfoque deductivo se basa tanto de clasificaciones y estudios previos, como en el análisis de

---

<sup>311</sup> La investigación de Kuntz se centra en las obras de referencia y manuales de la industria del automóvil (Kuntz 1979: 58-59, cf. Göpferich 1995).



estos criterios comunicativos, socioculturales y pragmáticos. Si nos detenemos en los estudios previos, el análisis extratextual se realiza mediante métodos cualitativos que se sitúan en la encrucijada entre la deducción y la inducción. En nuestro caso, incorporamos métodos cuantitativos cuando su introducción sea pertinente al análisis en cuestión. Éste será el caso del análisis del tono discursivo (véase 8.1.2.3).

### **8.1.1. Factores pragmáticos**

Defendemos que el foco contextual o función dominante en los sitios web corporativos tradicionales es **expositivo**<sup>312</sup>, principalmente en la descripción, con un foco contextual secundario exhortativo, tanto en exhortación con alternativa como sin alternativa (Hatim y Mason 1990). La exhortación con alternativa quedaría reflejada en los textos publicitarios dentro del sitio web, mientras que la exhortación sin alternativa se asocia a los términos legales o contratos de registro para acceder a zonas o restringidas. Consecuentemente, nuestra posición está en consonancia con varias investigaciones previas que defienden la multifuncionalidad de los sitios web corporativos, reflejada en sus diferentes objetivos que se solapan o corresponden con los diversos motivos que llevan a un usuario a dirigirse a los mismos (Nielsen y Tahir 2002: 3; Jeney 2007).

El carácter primario expositivo-descriptivo de este género web <sup>313</sup> que defendemos queda respaldado por un gran número de investigaciones (Askehave y Nielsen 2005; Robbin y Stylianou 2003; Nielsen y Tahir 2002; Sullivan 1999; Roberts 1998). En éstas se recoge que el objetivo básico de los sitios web corporativos sería (1) la creación de un canal de comunicación interactivo por el que diseminar

---

<sup>312</sup> En la terminología de Hatim y Mason (1990, 1997).

<sup>313</sup> En una investigación longitudinal reciente sobre los sitios web corporativos estadounidenses se concluye que una de las principales características de su evolución es el aumento de los contenidos informativos «Our longitudinal study of the S&P 500 websites revealed that corporate websites are incorporating more information and higher levels of interactivity and service than ever before.» (Heinze y Hu 2006: 323).

información entre la empresa y un conjunto amplio de receptores (Sullivan 1999; Manyard y Tian 2004; Robbin y Stylianou 2003; Nielsen y Tahir 2002), y (2) como un vehículo para la creación y mantenimiento de la imagen corporativa (Askehave y Nielsen 2004; Maynard y Tian 2004; Liu *et al.* 1997; Bolaños 2002), sin que la función primaria sea exclusivamente la promoción de productos<sup>314</sup> (Sullivan 1999). En este sentido, los manuales de estilo web desaconsejan el uso excesivo de un lenguaje asociado con la publicidad, puesto que los usuarios web reaccionan negativamente hacia este tipo de discurso publicitario<sup>315</sup> (Jeney 2007: 119; McAlpine 2001: 37; Bly 2002: 53; Nielsen y Loranger 2006: 265; Price y Price 2002; Yunker 2003).

Nuestra posición encuentra su contrapartida en investigaciones traductológicas que han defendido la naturaleza dominante exhortativa-apelativa de este género (Bolaños 2002, 2003; Bolaños *et al.* 2005). En las mismas se defiende el predominio de un foco contextual primario exhortativo dirigido a influir en el modo de actuar del receptor:

Este género multifuncional posee un foco contextual dominante, el exhortativo o instructivo, ya que se pretende influir en el modo de actuar o de pensar del receptor; se trata de lo que denominan instrucción con alternativa, típica de los textos publicitarios (Bolaños 2003: 5).

Este predominio queda recogido en otras investigaciones lingüísticas sobre sitios corporativos (Rennau 2004; Borja *et al.* 1999: 448). En este sentido, creemos pertinente diferenciar entre dos subgéneros corporativos presentados anteriormente: el sitio web corporativo tradicional y el de comercio en línea (véase 5.1.2.1). El primero posee un foco contextual dominante expositivo con un foco contextual secundario exhortativo, mientras que en el subgénero de comercio en

---

<sup>314</sup> El investigador señala sobre los sitios web corporativos «because our research suggests that marketing is not an important home page function» (Sullivan 1999: 209).

<sup>315</sup> Bly (2002: 53) señala como ejemplo la diferencia de estilo entre un anuncio en prensa «Buy this face cream today!» y el estilo de información publicitaria apropiado para transmitir el mismo mensaje por Internet «If you want to prevent flakiness, apply this particular cream twice a day after washing!».

línea estos focos contextuales se invierten, con un foco contextual dominante exhortativo y un expositivo secundario. Las conclusiones presentadas por Alicia Bolaños están motivadas, en nuestra opinión, por dos aspectos: por un lado no se ha asentado convenientemente el subgénero objeto de estudio en un marco más amplio que separe los sitios web corporativos de los sitios web centrado en el comercio en línea (véase 5.1.2.1), y por otro dado, que se estudia exclusivamente las páginas de inicio, uno de los múltiples bloques textuales que componen los sitios web corporativos. Estas páginas principales constituyen un resumen atractivo de los contenidos globales y animan a los usuarios a visitar sus distintas áreas (Askehave y Nielsen 2005), de ahí el posible predominio del componente exhortativo. Nuestro análisis se basa en sitios web completos, un aspecto que nos confiere una visión global del conjunto de subfunciones constitutivas de una función global primaria (Nord 1997a). Tal como hemos expuesto, ésta se corresponde principalmente con una función primaria expositiva.

Si situáramos este género en el contexto de la clasificación de géneros técnicos realizada por Gamero (2001: 69), el mismo se encontraría entre los géneros expositivos con un foco contextual secundario exhortativo, como los artículos comerciales o folletos informativos publicitarios. Este aspecto queda patente además en las indicaciones de las guías de estilo web que aconsejan un estilo informativo publicitario frente al estilo publicitario directo (Bly 2002; Price y Price 2002. etc.). La descripción pragmática de los sitios web se asemeja así, en parte, a la presentada por Silvia Gamero (2001) para el género impreso «folletos informativos publicitarios»:

Son de escasa longitud [...] se dirigen a potenciales consumidores y usuarios, o bien a los distribuidores, y en los que las empresas describen las características de los productos de la forma más clara posible (función expositiva) para ganar compradores (función exhortativa) (Gamero 2001: 76).

La investigadora describe los textos técnicos expositivos con foco secundario exhortativo dirigidos al público general como textos en los que prima la exposición de conceptos, aunque con una finalidad última de tipo publicitario, que denomina «información publicitaria» (ibid: 75). Más específicamente, los sitios web se dirigen a unos receptores semiespecializados, con lo que pertenecen a un subgrupo que siguiendo a Bolaños (2003) podríamos denominar «comunicación semiespecializada». Así, aunque se dirigen a usuarios generales y no a especialistas<sup>316</sup>, sí que se les presuponen ciertos conocimientos técnicos mínimos. Nuestra defensa del carácter secundario de la función exhortativa se fundamenta en que los usuarios se dirigen activamente hacia el sitio web principalmente en busca de información, un aspecto que recogen acertadamente los investigadores Robbin y Stylianou:

Globally accessible web sites enable corporations to communicate [...] to anonymous recipients who are **actively** seeking information<sup>317</sup> (Robbin y Stylianou 2003: 205).

Por consiguiente, la función de estos sitios web no consiste exclusivamente en modificar la conducta de los visitantes, ya que gran cantidad de bloques comunicativos en el género, como los productos, informaciones legales, noticias, páginas de contacto, descripción de la empresa, aun tendiendo un foco contextual secundario exhortativo, son expositivas-descriptivas en su naturaleza. La guía de estilo de Bly (2002) recoge esta dualidad expositiva y exhortativa del texto en sitios web corporativos:

Even a purely informative business document has an agenda or persuasion when it is posted on a business website (Bly 2002: 59).

Coincidimos con el autor en que no existirían textos puramente expositivos en este género. Asimismo, no podemos olvidar el solapamiento de dos niveles

---

<sup>316</sup> En el análisis del género se observan bloques dentro del género dirigidos a especialistas, como inversores, desarrolladores, etc.

<sup>317</sup> Énfasis propio.

comunicativos concurrentes en todo sitio web, la proceso comunicativo emisor-receptor y la interacción máquina-usuario (véase 5.5.6), que a su vez se realizan a través de una combinación textual e icónica-visual (Bolaños *et al.* 2005; McAlpine 2001: 34). Todo sitio web ha de «invitar» al usuario a interactuar y acceder a sus diferentes áreas, una intención que se realiza por medio del texto «interfaz» (Price y Price 2002). En estos elementos marcados por un hipere enlace queda reflejada esta dualidad pragmática, puesto que asumimos la naturaleza descriptiva del contenido de la página o nodo al que apuntan (Neumüller 2004), y a su vez, ésta es exhortativa puesto que convidan al usuario a que interactúe. Estos enlaces presentan una tipología textual en la que predominan unidades léxicas nominales (Craignon y Bouffard 2006), y su carácter exhortativo forma parte del conocimiento genérico convencional de los usuarios web, que asumen la presencia de enlaces en los menús de navegación o texto subrayado.

Puesto que contamos con un amplio corpus representativo de este género, pasamos a analizar esta dimensión pragmática sobre los análisis léxicos del corpus textual de sitios web.

#### **8.1.1.1. Análisis de los factores pragmáticos mediante formas verbales**

Nuestra hipótesis sobre el carácter primario expositivo de este género podría validarse mediante un análisis de las formas verbales recogidas en el corpus. Introducimos un método cuantitativo mediante el recuento y análisis de las formas verbales en la lista global de palabras del corpus realizada por medio de la herramienta Wordlist de Wordsmith Tools. No obstante, hemos de tener presente que la exhortación no se formula exclusivamente por medio de formas verbales monoléxicas, sino que la misma se manifiesta igualmente en aspectos visuales o en perífrasis verbales. Nuestro análisis es limitado puesto que se reduce al texto

extraíble de los sitios web, con lo que queda excluido todo texto en imágenes o presentaciones Flash. Esta posible deficiencia queda en parte compensada por la inclusión de las descripciones textuales de imágenes presentes en las etiquetas <alt> (véase 9.2.2) Igualmente, siguiendo investigaciones previas excluimos las palabras forma más comunes (Pérez Hernández 2002)<sup>318</sup>.

Lista lematizada en sección original del corpus					
Posición	Verbo	Frecuencia	Posición	Verbo	Frecuencia
1	Ser	0,6%	180	Descargar	0,05%
12	Estar	0,18%	192	Imprimir	0,04%
19	Poder	0,16%	198	Permitir	0,04%
25	Viajar	0,13%	199	Ir	0,04%
43	Tener	0,1%	233	Realizar	0,04%
50	Trabajar	0,09%	255	Pinchar	0,04%
52	Ampliar	0,09%	264	Solicitar	0,04%
53	Haber	0,09%	283	Deber	0,03%
54	Buscar	0,08%	336	Agregar	0,03%
57	Contactar	0,08%	383	Volver	0,03%
80	Comprar	0,07%	388	Acceder	0,03%
85	Mejorar	0,07%	400	Comparar	0,03%
86	Hacer	0,07%	440	Utilizar	0,02%
114	Consultar	0,06%	476	Incluir	0,02%
122	Recibir	0,06%	479	Obtener	0,02%
147	Ofrecer	0,05%	493	Completar	0,02%

**Tabla 6. Verbos más frecuentes en lista global de palabras en corpus original lematizada.**

Los resultados muestran que los verbos constituyen el 6,4% del total de lexemas presentes entre las primeras 500 formas lematizadas. Asimismo, al igual que la lista de frecuencia del Corpus de Referencia del Español Actual (CREA), el verbo más frecuente es el verbo «ser», 0,6%, seguido del verbo «estar», 0,18%, con lo que podríamos deducir que existe una clara primacía del carácter descriptivo en el conjunto de contenidos proposicionales de estos textos. Asimismo, la presencia del verbo «poder» es indicativa del carácter secundario de la función apelativa, y que apunta a un mecanismo para suavizar la formulación de la apelación frente a la apelación directa.

<sup>318</sup> Que incluimos en el apéndice 14.2

27	net para que nos encuentre. Si usted	DESEA vender sus productos por Internet
28	te seleccione el papel normal que usted	DESEA utilizar; ¡incluso de su tienda d
29	tión de mamparas! Sin embargo, si usted	DESEA tomar las medidas. aquí le facili
30	( Tracking de pedidos UPS) Si usted	DESEA que sus clientes sepan en todo mom
31	e hacerlo desde 1.000 Ptas. Si usted	DESEA que gestionemos su propia cartera
32	y series para televisión. Si usted	DESEA ponerse en contacto con XTREM SOLU
33	l teléfono *91 708 90 00*. Si usted	DESEA obtener información sobre eventos
34	de Ptas. Por último, si lo que usted	DESEA es constituir una SIM - SIMCAV. a
35	mi patrimonio mobiliario?* Si usted	DESEA convertirse en partícipe de cualqu
36	iferentes ofertas y acciones. Si usted	DESEA rectificar o cancelar sus datos en
37	iferentes ofertas y acciones. Si usted	DESEA rectificar o cancelar sus datos en
38	siguientes servicios: Nota: Si Ud.	DESEA un servicio de espicha sentado vea
39	: Servicios Completos Si Ud.	DESEA realizar cualquier evento gastronó
45	----- Si	DESEA ver otra extensa gama de novedosos
46	o vehículos. inmobiliarias etc... Si	DESEA ver como funciona y donde se ubica
47	nes de Comercio Electrónico* *  * Si	DESEA utilizar la potencia de la red Int
48	o web - Presupuestos personalizado Si	DESEA una web presencial sin muchos camb
49	o web - Presupuestos personalizado Si	DESEA una web presencial sin muchos camb
50	ntraseña? ADSL 2048 ===== Si	DESEA una velocidad de navegación superi
51	€ *Promociones* Si	DESEA un servicio de acceso a Internet d
52	de los datos que Vd. nos facilite. Si	DESEA un estudio más ajustado a la reali
53	des. · Gestión de rendimiento. Si	DESEA Ud. recibir más información sobre

Tabla 7. Extracto de líneas de concordancia de «desea» con Wordsmith Tools.

Por su parte, la lista no lematizada muestra una frecuencia del 7% de formas verbales, y una mayor frecuencia de la forma verbal «es» y «está» en tercera persona, con lo que se corrobora de nuevo que la función expositiva prima en este género.

Lista no lematizada en sección original del corpus					
Posición	Verbo	Frecuencia	Posición	Verbo	Frecuencia
2	Es	0.32%	294	Han	0.03%
34	Está	0.09%	295	Agregar	0.03%
35	Ver	0.09%	303	Realizar	0.03%
43	Buscar	0.08%	312	Compra	0.03%
57	Puede	0.07%	314	Pueden	0.03%
59	Son	0.07%	336	Recibir	0.03%
62	Ampliar	0.07%	353	Solicitar	0.03%
92	Somos	0.06%	371	Hay	0.03%
121	Ser	0.05%	380	Comparar	0.02%
147	Descargar	0.04%	393	Estás	0.02%
156	Imprimir	0.04%	396	<b>Sea</b>	0.02%
160	Tiene	0.04%	416	<b>Contacta</b>	0.02%
190	Comprar	0.04%	418	Volver	0.02%
195	<b>Contacte</b>	0.04%	430	Viajar	0.02%
211	<b>Pinche</b>	0.04%	435	Podrá	0.02%
218	Ofrece	0.04%	468	Recibirás	0.02%
229	<b>Desea</b>	0.03%	483	<b>Subscríbete</b>	0.02%
232	Permite	0.03%	486	Tener	0.02%

Tabla 8. Verbos más frecuentes en lista global de palabras en corpus original no lematizada.

La función exhortativa se asocia parcialmente con la presencia del modo imperativo del verbo, que constituye el 17.14% de las formas verbales encontradas<sup>319</sup>. Este aspecto contrasta con los resultados obtenidos por Renau (2004) sobre los sitios web de empresas azulejeras, en los que el modo imperativo se encuentra mayoritariamente en empresas anglosajonas, siendo casi inexistente en las españolas. La investigadora señala que la apelación directa por medio de imperativo no responde a las convenciones de este género en español. En este sentido, se observa además la presencia de modulación con formas verbales en segunda persona, «pueden», «desea», «pueden», «sea», «podrás» que señalan una formulación de la función apelativa atenuada.

Como resumen, este análisis apoya nuestra defensa del foco contextual primario expositivo-descriptivo en el género *sitio web corporativo tradicional*, con un foco contextual secundario exhortativo. Los objetivos fundamentales de estos textos se pueden sintetizar en:

- a) La creación de un canal de comunicación tanto para la interacción con los diferentes tipos de usuarios – clientes, distribuidores, inversores, futuros compradores, etc.-, como para la diseminación de información de manera económica hacia el conjunto de su audiencia.
- b) El mantenimiento o creación de una imagen corporativa de prestigio que sea atractiva al usuario. En este sentido, cobran una gran importancia tanto el diseño web como el bloque comunicativo relativo a la historia, experiencia o filosofía de la empresa.

---

<sup>319</sup> El corpus no se encuentra actualmente etiquetado morfológicamente por lo que se incluye la forma «compra» que puede ser igualmente una forma nominal.



- c) La creación de un conjunto publicitario en el que se describen atractivamente los productos y servicios ofrecidos por la empresa y su posible compra en línea, así como los servicios de pre y post venta.

### 8.1.2. Factores comunicativos

La autoría del texto en los sitios web corporativos se encuentra a cargo de un redactor técnico, un desarrollador web, o un empleado de la compañía responsable del sitio<sup>320</sup>, cuya identidad no suele aparecer implícitamente en el texto<sup>321</sup>. La empresa representa la emisora del texto y, convencionalmente, su nombre o logotipo aparece en la esquina superior izquierda de las páginas web<sup>322</sup>. Esta identificación de la empresa es obligatoria para el establecimiento de la credibilidad textual necesaria en Internet (Posteguillo 2003).

El contexto tecnológico de creación textual suele ser cooperativo, por lo que no suele existir un autor único (Bolaños 2003). Como resultado, se produce un estilo heterogéneo y una posible falta de homogeneidad terminológica, que como veremos, es más acusada en los textos localizados (véase 11.3.1). En su estudio sobre el lenguaje de informática e Internet, Belda (2003) identifica este problema con la ausencia de auténticos profesionales en este campo:

La falta de **homogeneidad terminológica**<sup>323</sup> en el ámbito de la informática, debido también en parte a la nula presencia de verdaderos profesionales de la palabra (traductores, terminólogos, correctores de texto) en el sector empresarial [...] que presten sus servicios en pos de la necesaria uniformidad para traducir los adelantos tecnológicos que nos llegan del exterior (Belda 2003: 17-18).

---

<sup>320</sup> Las guías de estilo web anglosajonas engloban a su audiencia a «technical writers, journalists, copy editors, content providers».

<sup>321</sup> En ciertos sitios web aparece el emblema o correo electrónico de la empresa desarrolladora web, existiendo de manera normalizada el atributo «DC.Creator», como ejemplo de [www.ylos.com](http://www.ylos.com), <META NAME="DC.Creator" CONTENT="Ylos.com Soluciones de Comercio electrónico.">, o la empresa a cargo de la localización como en el caso de [www.boeing.es](http://www.boeing.es), <meta name="creator" content="Translation Solutions"/>.

<sup>322</sup> Nielsen y Tahir (2002) indican que esta es la localización mayoritaria en sitios web estadounidenses, 84%, un 6% en la derecha y un 6% en el centro.

<sup>323</sup> Énfasis del autor.

El perfil del autor o redactor se decanta mayoritariamente hacia la formación técnica en desarrollo web, sin una formación específica en redacción técnica en este medio (Jeney 2007; McAlpine 2002). Por consiguiente, queda patente un paralelismo entre una menor importancia de la redacción técnica sobre el desarrollo técnico web<sup>324</sup> (McAlpine 2002), y la primacía otorgada a los procesos técnicos sobre los de traducción en el sector de la localización (Pym 2004a). En cualquier caso, la escasez de manuales de estilo web producidos específicamente en España atestigua la escasa relevancia que se otorga a la figura del redactor técnico web. Las únicas publicaciones específicas sobre la redacción web se dirige exclusivamente al periodista digital (Díaz Noci y Salaverría 2003; Vilamor 2001), mientras que en el ámbito anglosajón existen gran cantidad de publicaciones dirigidas a redactores técnicos web, un claro reflejo del reconocimiento socioprofesional de este perfil laboral (véase 5.9).

Tal como expusimos en nuestro marco teórico (véase 3.3.2), los receptores de los sitios web corporativos son heterogéneos, siendo difícil identificar a un receptor único medio<sup>325</sup> (Krug 2006). El receptor o usuario constituye la audiencia primaria del texto, mientras que los traductores-localizadores se incluirían en su audiencia periférica, ya que estos nunca son sus destinatarios directos sino receptores fortuitos (Muñoz Marín 1995: 45)<sup>326</sup>. El análisis léxico de nuestro corpus, principalmente en el bloque comunicativo dirigido a usuarios específicos (véase 9.1.2.7), identificamos el conjunto términos relacionados con el concepto «usuario» recogidos en este género. A continuación presentamos la lista que atestigua la gran variedad de receptores posibles: (1) usuarios, (2) clientes, (3)

---

<sup>324</sup> Los aspectos textuales suelen ser casi inexistentes en las publicaciones sobre desarrollo web, «Most manuals on creating web sites barely mention content, if at all. [...] web developers and their clients assume that writing is the last of their problems» (McAlpine 2001: 3).

<sup>325</sup> Existen no obstante autores que reivindican que es inapropiado dirigirse a una audiencia global y es necesario ajustar el texto web a un destinatario medio (Price y Price 2002: 4).

<sup>326</sup> No obstante, en los comentarios incluidos en el código html los desarrolladores pueden incluir comentarios específicamente dirigidos a los traductores-localizadores como se puede observar en el caso del sitio web de Starbucks (véase 9.2.3).

afiliados, (4) particulares, (5) médicos, (6) pacientes, (7) agentes de viajes, (8) viajeros, (9) desarrolladores o *developers*, (10) revendedores, (11) administradores de sistemas, (12) diseñadores, (13) programadores, (14) docentes, (15) estudiantes, (16) consumidores, (17) grandes empresas, (18) medianas empresas, (19) Pymes, (20) proveedores o (21) socios.

Adicionalmente, y como se mencionó anteriormente, el proceso comunicativo que se entabla es doble: por un lado se lleva a cabo entre la empresa emisora y el usuario y, por otro, entre el usuario y el propio sitio web por la interactividad que proporciona el medio. Además, se produce un proceso comunicativo entre el sitio y la **hiperweb**, el conjunto global de sitios web activos en la Red. Este proceso garantiza que el sitio web sea incluido en los diferentes motores de búsqueda como Google, y se plasma en la información metatextual incluida en las etiquetas <META> en sus atributos *Content, Description, Language*, etc. (véase 9.2).

### **8.1.2.1. El carácter multilingüe de los sitios web corporativos**

En la delimitación de la dimensión comunicativa del contexto textual, Hatim y Mason (1990) distinguen aspectos relacionados con los usuarios de la lengua, como por ejemplo el dialecto, y otros con los usos de la misma, como el registro. El dialecto se podría en cierta manera comparar con el concepto de *locale* puesto que introduce la diferencia dentro de lenguas según la situación geográfica. En esta área los autores incluyen también el sociolecto o idiolecto. Éste último se define como «el modo particular e irrepetible que cada hablante hace uso de la lengua»<sup>327</sup> (Muñoz Martín 1995), y tendría una importancia doble en los sitios web: por un lado, las empresas suelen decidir sobre la terminología que van a usar por motivos de coherencia textual necesaria (Gamero 2001), y por otro lado, hay que señalar que para llegar al mayor número posible de receptores en el contexto de la

---

<sup>327</sup>En lo que concierne a los sitios web, el idiolecto se puede referir igualmente a las características idiosincrásicas compartidas por un grupo de hablantes (Hatim y Mason 1997: 102).

globalización (Yunker 2003), el sector pretende escribir en un lenguaje lo más neutro posible<sup>328</sup>, alejado del idiolecto<sup>329</sup>.

Tras el análisis léxico del corpus recopilado, podemos concluir que el contexto comunicativo en el que se enmarcan los sitios web corporativos es multilingüe en su naturaleza. A pesar de que en la fase de procesamiento del corpus excluimos toda página redactada exclusivamente en otras lenguas, el análisis léxico global de ambas secciones del corpus recoge gran cantidad palabras en varios idiomas. Presentamos, como ejemplo, la siguiente captura perteneciente a la lista de palabras de la sección original que muestra una frecuencia similar de las palabras en español «experto», «expertos», o «experta» y las inglesas *expert*, *expertise* y *experts*.

30658	EXPERT	168
30659	EXPERTA	30
30660	EXPERTAS	2
30661	EXPERTISE	5
30662	EXPERTO	76
30663	EXPERTOS	266
30664	EXPERTS	11

Figura 38. Extracto de lista global de palabras en corpus original que demuestra la presencia del inglés.

La siguiente página web de un sitio web español en el corpus original ilustra adicionalmente este aspecto, ya que muestra bloques textuales en tres idiomas: español, francés e inglés. Esta página de inicio forma parte del sitio web de la empresa española Fabregas, e ilustra el posible multilingüismo presente en este tipo de textos. En cualquier caso, este sitio supondría un ejemplo extremo en este sentido, ya que responde a un modelo de localización ineficiente dada la

<sup>328</sup> En el caso del español, es normal que la industria localice a tres variedades del español, español peninsular, centroamericano y del cono sur.

<sup>329</sup> Los estudios de usabilidad indican la necesidad de escribir lo más claramente posible, sobre todo en los segmentos con enlaces, puesto que la necesidad de interpretar los términos idiolectales puede ralentizar la navegación (Krug 2005).

sobrecarga cognitiva que produce una ingente cantidad de texto en diversos idiomas que el usuario ha de descartar.



Figura 39. Página de inicio multilingüe del sitio web [www.fabregas.com](http://www.fabregas.com)

De entre los posibles idiomas identificados, la presencia de palabras en inglés es especialmente significativa en los textos españoles y localizados. Aparte de la existencia específica de marcos sin traducir o subtítulos en este idioma, la profusa presencia del inglés que recogen nuestros análisis léxicos queda reflejada en los siguientes elementos:

- Nombres de empresas: «CIGNA Life Insurance Company of Europe, S.A.» [Cigna]
- Nombres de productos o servicios
- Direcciones postales
- Eslóganes publicitarios: «ADM es Resourceful by Nature™.» [empresa ADM]
- Derechos de propiedad «Copyright © The Dow Chemical Company (1995-2006). All Rights Reserved.»

- Explicitación de siglas y acrónimos: «Almacenamiento de información accedido por contenido (CAS, Content Adressed Storage)» [EMC]
- Neologismos y anglicismos puros (Aguado de Cea 1994), «Plug-ins, Herramientas, info de browser, RSS Feeds\*»<sup>330</sup>[EMC]

Como resumen de este epígrafe, deducimos que este carácter multilingüe constituiría uno de los rasgos básicos de este género web. Este hecho entronca con una de las peculiaridades de la situación comunicativa en Internet, la mezcla de varios idiomas o la presencia de versiones en otras lenguas en un mismo sitio web (Yunker 2003). De hecho, en virtud de la alta especialización de las áreas dirigidas a expertos, desarrolladores, exportadores<sup>331</sup> o inversores, éstas se redactan en inglés o no se localizan, fiel reflejo de su estatus como *lingua franca* en los campos de especialidad (LISA 2003; Posteguillo 2003). En cierta manera, este aspecto entronca con el principio de la resistencia a la distribución textual como elemento determinante de la localización propuesto por Anthony Pym (2004). De hecho, se asume que los desarrolladores o los inversores no ofrecerán resistencia al uso del texto original en inglés y, por tanto, no se localizan. Esta tendencia es más acusada en los sitios web localizados, con la totalidad de las áreas dedicadas a desarrolladores o inversores redactadas en inglés o simplemente no localizadas<sup>332</sup>.

En el caso de los sitios web localizados, la presencia de áreas en inglés queda patente al incluir la mención expresa a la lengua de la página o sección. La imagen que presentamos a continuación muestra esta posible alusión al idioma de la página de destino, aunque muestra una falta clara de homogeneidad<sup>333</sup> tanto en el idioma en que se expresa como en su tipografía.

---

<sup>330</sup> Extraído del menú de navegación de [www.emc.com](http://www.emc.com).

<sup>331</sup> [www.keapapel.com](http://www.keapapel.com) en los textos originales.

<sup>332</sup> El mantenimiento en inglés de las áreas dirigidas a desarrolladores es total en los sitios web localizados. IBM, Sun, Motorola o Electronic Arts son ejemplos de la permanencia de esta sección en inglés.

<sup>333</sup> Este ejemplo ilustra el efecto del uso de *Global Management Systems* o sistemas de gestión de contenidos multilingües y la consecuente falta de homogeneidad en el texto final.



Figura 40. Idioma de la página de destino en menú de navegación. De [www1.euro.dell.com](http://www1.euro.dell.com).

### 8.1.2.2. Estadísticas de localizaciones de los sitios web españoles

Durante el proceso de recopilación del corpus recogimos la información referente a los posibles idiomas de localización de cada sitio web. En lo que respecta al corpus original, el inglés es la lengua hacia la que más se localizan los sitios corporativos españoles (46% de sitios web), seguida del francés (13,3% de sitios web), el catalán y el alemán (9,2%) o el italiano (6,3%). En la siguiente gráfica se puede observar la distribución de versiones localizadas de los sitios web españoles.

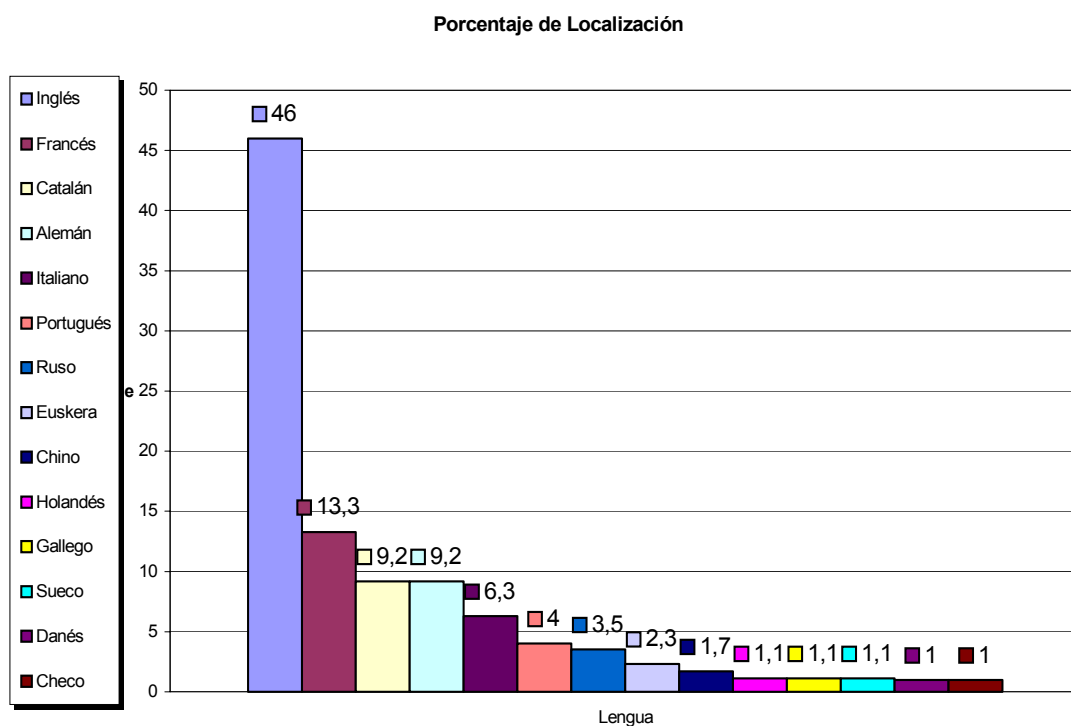
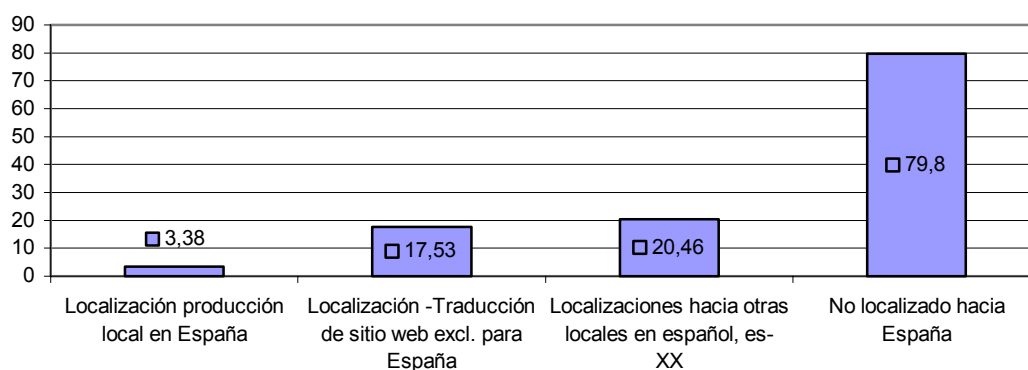


Tabla 9. Distribución de versiones localizadas de sitios web españoles.

Esta tabla muestra catorce posibles idiomas hacia los que se localizan los sitios web corporativos españoles. La alta presencia del inglés refleja su carácter de lengua internacional en Internet (Posteguillo 2003), con el resto de versiones dirigidas principalmente hacia países europeos. El único idioma hacia el que se localizan sitios españoles del corpus fuera del contexto europeo sería el chino. Es igualmente significativa la presencia de las tres lenguas de carácter co-oficial en España, el catalán, el gallego y el euskera, aunque la presencia del primero es marcadamente superior. El porcentaje de idiomas de localización obtenido es congruente con los resultados obtenidos por Rennau (2004: 309).

Por su parte, si analizamos el porcentaje de localización hacia el español de los sitios web pertenecientes a las 650 mayores empresas estadounidenses, las estadísticas serían las siguientes:



**Figura 41. Porcentaje de localizaciones en español de las mayores empresas norteamericanas.**

- a) Sitios web dirigidos a España: 20,92% (136 sitios web).
- b) Sitios web localizados hacia España: 17,53% (114 sitios web).
- c) Sitios web producidos originalmente en España: 3,38% (22 sitios web).
- d) Sitios web adicionales con localizaciones hacia otras variedades del español o *locales*, mayormente español para Estados Unidos: 20,46% (133 sitios web)



- e) Total localizaciones en español de sitios web de las mayores empresas norteamericanas, 41,38%.

Estos resultados apuntan a la importancia del español en el contexto sociocultural estadounidense, ya que el 41,38% de los sitios web de las mayores empresas estadounidenses presentarían una versión en este idioma. Este porcentaje está próximo al 46% en la dirección inversa, localización de sitios web españoles hacia el inglés.

Con el objetivo de contrastar estos resultados obtenidos, recogemos las cifras referentes a los porcentajes de localización de los sitios web de empresas multinacionales ofrecidos por LISA (2007: 27). Estos resultados representan la dirección desde el inglés estadounidense hacia otras lenguas, y muestran una mayoría de localizaciones hacia el francés (56%), el español (53%), y el alemán (50%), incluyendo además, el inglés<sup>334</sup> (35%), el japonés (29%), el chino simplificado<sup>335</sup> (27%), el italiano (21%), y un 7% representado por el resto. Esta publicación sitúa por tanto en el 53% el porcentaje de localización al español, un 11,62% superior a los resultados que hemos de nuestro estudio recopilados en base a las 650 mayores multinacionales estadounidenses.

El carácter multilingüe de los sitios corporativos aparece explícitamente mediante la presencia de un área de navegación global o selector de idiomas, un elemento que se ha convencionalizado ampliamente en este género (Adkisson 2002). Las siguientes figuras se elaboraron a partir del conjunto de unidades léxicas que recogimos en el menú de selección de la región lingüística en los sitios localizados.

---

<sup>334</sup> Aunque la publicación no lo indica implícitamente, se trata de localizaciones hacia el inglés británico.

<sup>335</sup> El estándar Unicode contiene dos versiones de conjuntos de caracteres para el chino, el tradicional y el simplificado. El estándar simplificado, GB18030-2000, se usa en Singapur y la República Popular China, que constituye el mayor mercado para este idioma, mientras que el estándar de chino tradicional se emplea en Hong Kong, Macao y Taiwán. Para más información consúltese el sitio web del W3C: <<http://people.w3.org/rishida/scripts/chinese/#simpchin>>.

La primera tabla muestra las unidades léxicas más frecuentes, y la segunda la frecuencia de las palabras más usadas en las su conjunto.

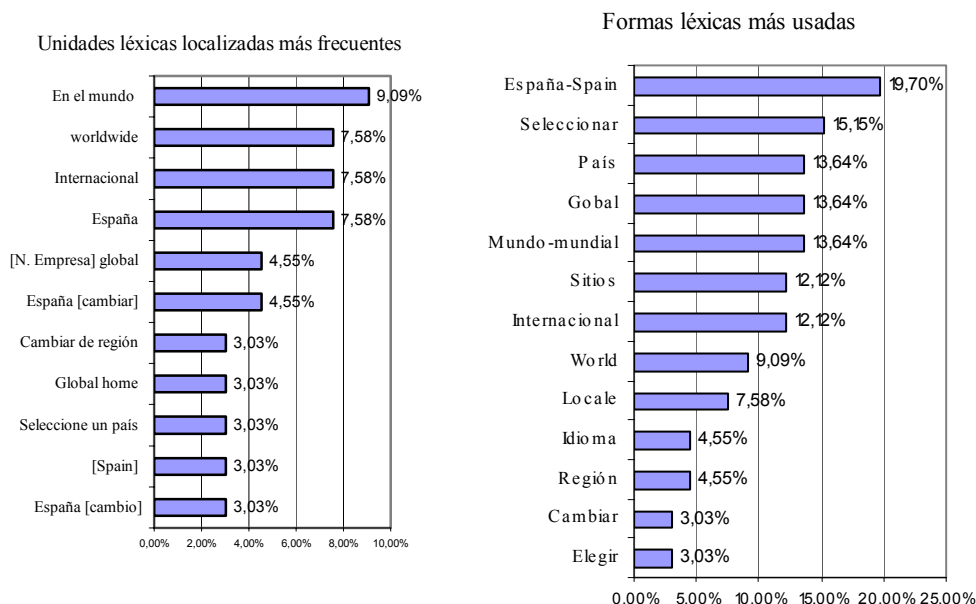


Figura 42. Unidades léxicas y formas léxicas más frecuentes en la selección de locale.

El análisis de las combinaciones léxicas y de las palabras más usadas en la selección de *locale* muestra que el nombre del país, España o *Spain*, sería la palabra más usada mientras que la unidad léxica más recurrente sería «en el mundo». En general, la referencia al país o región en sí sería mayoritaria frente a las menciones a los idiomas en particular, un hecho que entronca con el concepto de *locale* como combinación de región lingüística y cultural (Pym 2004a), que no un país en sí. Por cuanto a la forma verbal asociada, en los sitios localizados es mayoritario el verbo «seleccionar» frente a «cambiar».

### 8.1.2.3. Análisis del tenor

Como expusimos anteriormente, el análisis del tenor se sitúa dentro de la dimensión comunicativa en la propuesta de Hatim y Mason (1990) (véase 4.2.1.3.1). La importancia de analizar contrastivamente este aspecto surgió por (1)

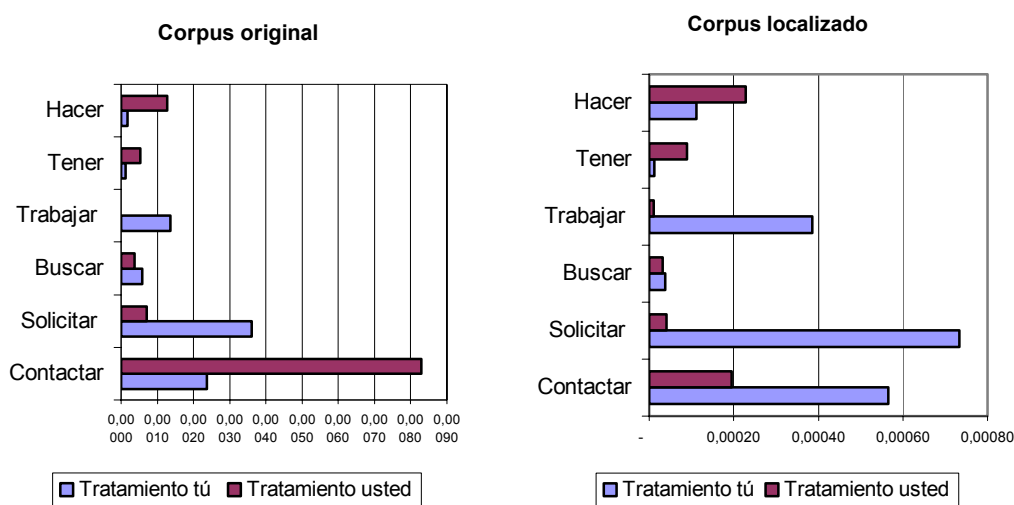
las diferencias identificadas entre textos originales y traducidos en nuestros estudios preliminares (Jiménez 2005a, 2006), (2) por las diferencias interlingüísticas en este sentido y (3) su posible impacto en los textos localizados (Tercedor 2005). Con el objetivo de analizar contrastivamente el tenor digital se realizaron tres tipos de análisis; dos de ellos por textos globales y un tercer análisis basado en un subtexto recurrente, los menús de navegación.

En primer lugar analizaremos las formas verbales en modo imperativo más recurrentes en textos globales, siempre teniendo en cuenta que el tono del discurso va más allá de propia selección de formas pronominales formales o informales, cubriendo igualmente el léxico escogido, las estructuras sintácticas usadas, o la propia estrategia de los discursos (Muñoz Martín 1995: 44). Nuestra selección de las formas verbales imperativas se debe a que los localizadores deben decidir en su traducción entre un tono formal o informal y, por consiguiente, este rasgo responde directamente a decisiones tomadas durante el proceso de localización. De ahí que defendamos quede justificado su estudio contrastivo como variable indicativa del tono discursivo.

Las formas verbales en modo imperativo más frecuentes en la lista global corresponden a los verbos: «contactar», «solicitar», «buscar», «trabajar», «tener», y «hacer». Procedimos a comparar las frecuencias de aparición de las formas verbales en modo imperativo con el tratamiento de «tú» o de «usted» de los verbos seleccionados. En la primera gráfica contrastiva se recogen los resultados de las formas verbales en modo imperativo separadas según el tratamiento de «usted» o «tú». Las barras azules corresponden a la frecuencia global del corpus original de esta forma verbal y las rojas en el corpus localizado.

El análisis contrastivo muestra una mayor frecuencia global en el corpus localizado que en el original del modo imperativo en «tú». El tratamiento informal es mayoritario en las formas localizadas del corpus localizado, aunque éste es menor

en los verbos «hacer» y «tener», aspecto similar en el corpus original<sup>336</sup>. La coincidencia más significativa entre ambos corpus se encuentra en la superioridad absoluta de las formas verbales informales del verbo «trabajar», aunque ésta es más prominente en el corpus localizado. No obstante, la mayor frecuencia relativa de las formas verbales en modo imperativo en el corpus localizado podría corresponderse con un mayor predominio de una tipología textual exhortativa en los textos localizados, con lo que sería necesario un segundo análisis.

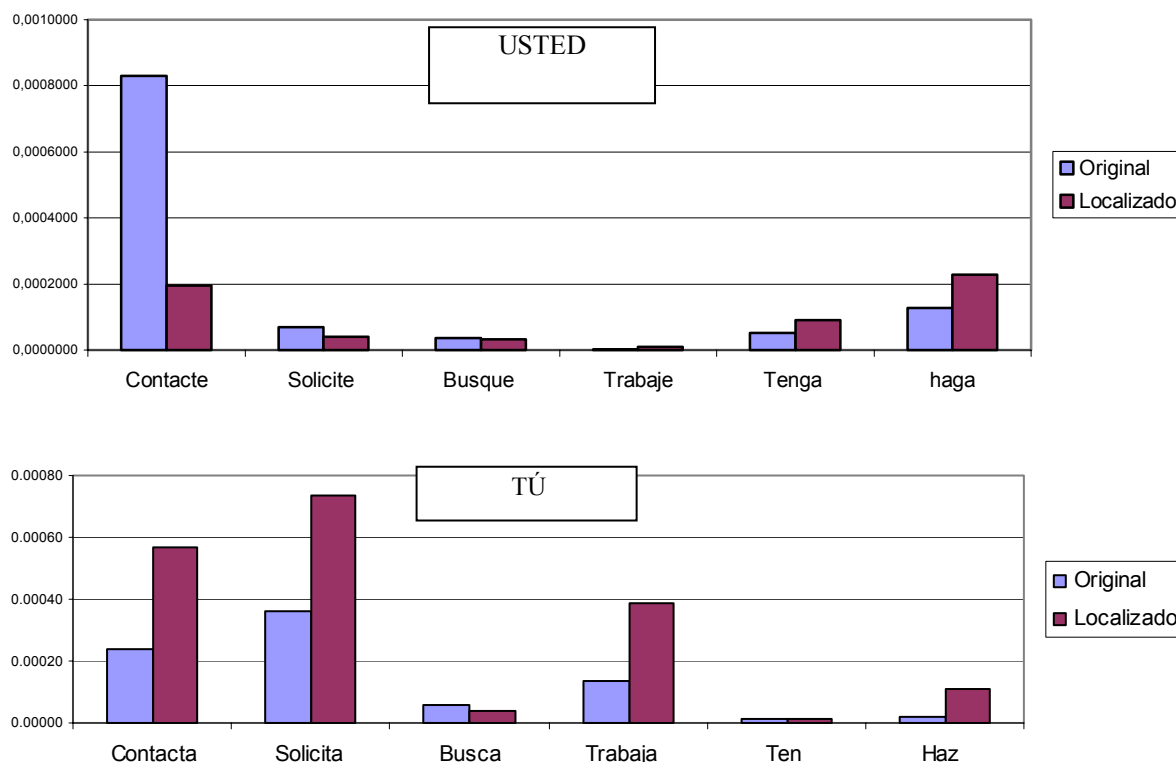


**Figura 43. Análisis contrastivo del tenor digital a través del modo imperativo según frecuencias globales en ambas secciones del corpus.**

Si organizamos estos mismos datos de acuerdo con la forma verbal en ambos corpus, queda patente una menor presencia general de las formas verbales en «usted», y en especial la mayor presencia en el corpus localizado de modo imperativo con tratamiento informal «tú». La diferencia es muy significativa en lo que respecta a los verbos «contactar», «solicitar» y «trabajar». Podemos por lo tanto

<sup>336</sup> El mayor índice de uso de las formas verbales en modo imperativo «haga» y «tenga» en el corpus localizado se corresponde con el uso recurrente de calcos léxicos y sintácticos como «haga su selección», «haga sus ideas realidad», «haga posible [...]», «haga click», «haga su uso», «haga uso del sitio web», «haga su vida más cómoda», «haga todo esto» o «ten por seguro», «tenga a mano». En el corpus original el uso de la forma en modo imperativo «ten» se reduce a las colocaciones «ten presente» y «ten seguro», y la forma «haz» se reduce a «haz click [clic]» y «haz copias». En el análisis de las colocaciones hemos observado un número reducido de uso de la forma «ten» en referencia al número diez en inglés.

concluir que la presencia de formas verbales en modo imperativo con el tratamiento informal «tú» es ampliamente superior en el corpus localizado.



**Figura 44. Análisis contrastivo del tenor digital por formas verbales en modo imperativo en textos globales. Estadísticas extraídas de la frecuencia en la lista global de palabras de ambas secciones del corpus.**

Decidimos realizar un segundo análisis centrado en los verbos que representan el ruego y que modulan el discurso. La pertinencia de este análisis adicional se debió a la posible distorsión que podría haberse producido por la repetición léxica en el texto «interfaz», como los menús de navegación. Con este objetivo seleccionamos los verbos que presentaron una mayor frecuencia en el análisis léxico, que se concretó en este caso en los verbos: «necesitar», «desear» y «querer». En este segundo análisis, y con el fin de controlar la repetición léxica, contabilizamos cada segmento que se repitiese a nivel oracional como una sola entrada por representar una única selección por parte del redactor o traductor.

El análisis de los resultados obtenidos demuestra que la frecuencia de formas verbales representativas de tratamiento formal es superior en la totalidad de los verbos en los textos originales. Por el contrario, los textos localizados muestran una mayor frecuencia en las formas verbales informales en los tres verbos seleccionados. Esta gráfica revela por tanto una ligera preferencia en los textos localizados por un tratamiento informal y distendido con los usuarios, que puede ser reflejo de una mayor *informalidad* (Posteguillo 2003) de los textos web anglosajones con respecto a los españoles. En segundo lugar la gráfica revela que, tal como Muñoz Martín recoge de los estudios lingüísticos previos (1995: 44), el tono se puede mostrar igualmente en la selección léxica. La gráfica muestra que el verbo «desear» y, en menor medida «necesitar», aparecen mayoritariamente en su forma verbal de «usted» en ambas secciones del corpus y, por consiguiente, queda patente una asociación entre el nivel de formalidad escogido y la propia selección léxica. Por contraste, el verbo «querer» aparece mayoritariamente en un tratamiento de «tú», «si quieres...», aunque no obstante la frecuencia es superior en los textos localizados. De hecho, esta variable muestra la mayor diferencia relativa entre ambos corpus, un 12,73%.

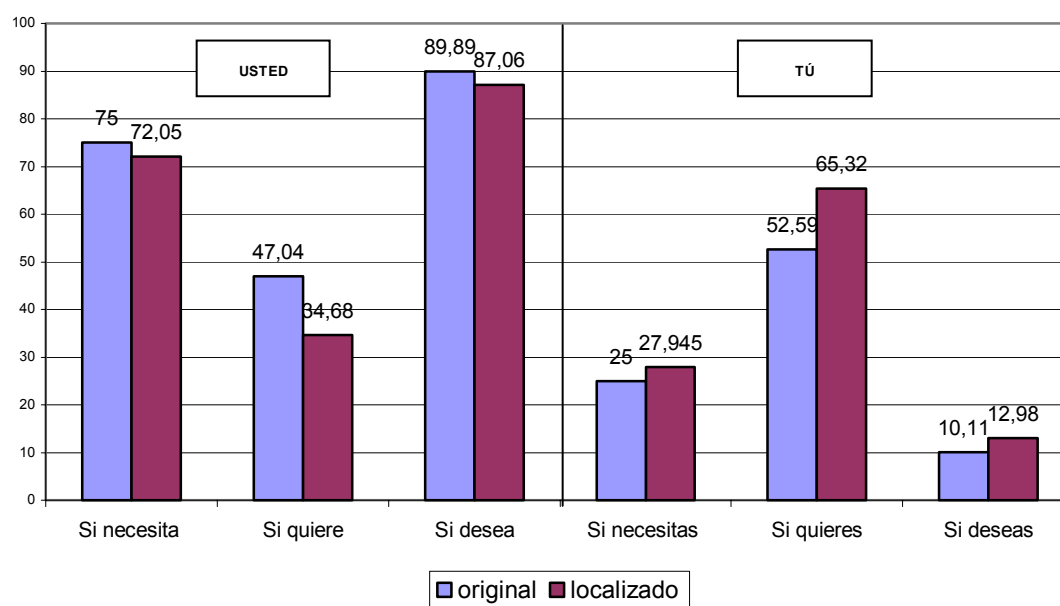


Figura 45. Análisis contrastivo del tenor digital a través de las formas de cortesía.

En un tercer lugar, y con el fin de evitar variables extrañas en el análisis cuantitativo del tenor digital, desarrollamos un análisis sobre unos elementos recurrentes en las páginas web, los menús de navegación. Tras el análisis de estos subtítulos (capítulo 9) seleccionamos los tres bloques o secciones que incorporan una apelación directa al receptor mediante formas verbales en modo imperativo: el bloque textual «contacto» (B), y las secciones «trabajo» (G1) y «registrarse» (H5) (véase 9.1).

Si comenzamos con el bloque comunicativo «contacto» (B), éste se caracteriza por ser el bloque de mayor frecuencia de aparición en este género y, por tanto, sería el más representativo en el uso de formas imperativas (véase 9.1). En su análisis se recopiló una sola forma verbal por sitio web en representación de la selección del tono textual, como «contacte», «contacta», «contáctenos», etc. (véase Figura 59, p. 358). La primera tabla muestra la primacía de la forma verbal del infinitivo, seguido de un 40% de uso de la forma verbal de usted, y un 18% de uso del imperativo en «tú».

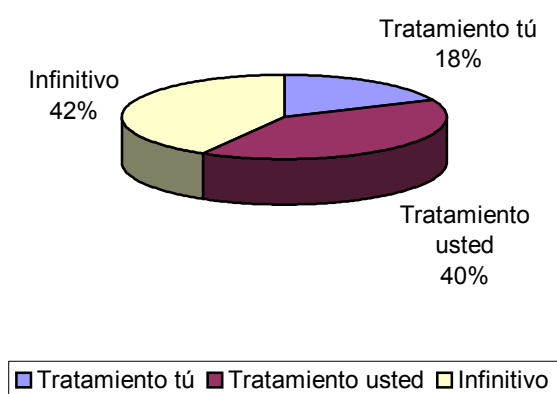


Figura 46. El tenor en los menús de navegación. Extracto de bloque «Contacto».

Si se excluyen las formas verbales infinitivas, podemos observar la primacía de un tratamiento formal con el uso de «usted» (68,57%).

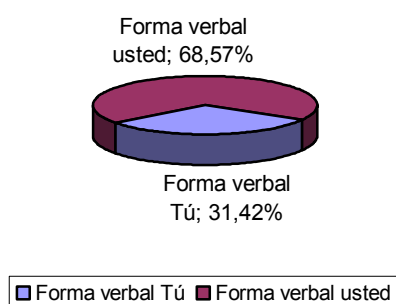


Figura 47. El tenor en los menús de navegación. Formas verbales en tú frente a usted «Contacto».

En un segundo lugar seleccionamos la forma lingüística convencional en los menús de navegación que apela al receptor a registrarse como usuario. Estas unidades léxicas se corresponden con la sección comunicativa «H3» (véase 9.1.2.7.1). En este caso, queda patente un porcentaje similar de uso del modo imperativo con «usted» frente a la informal «tú»<sup>337</sup>:

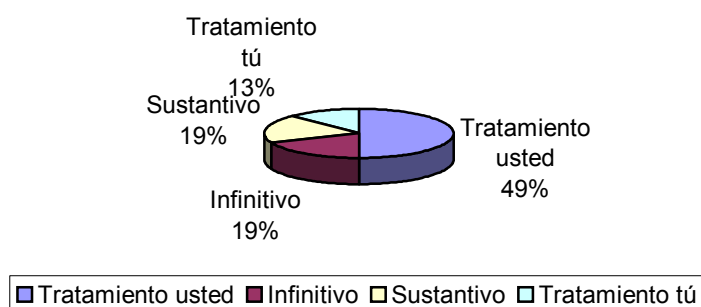
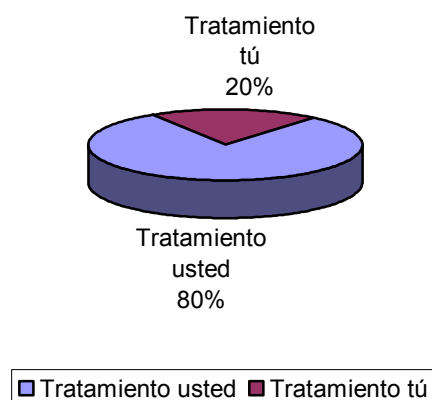


Figura 48. El tenor digital en la sección «Registro».

De igual modo, si analizamos exclusivamente las formas verbales personales, queda patente la preferencia por el tratamiento formal:

<sup>337</sup> Para la lista completa de palabras usadas para la elaboración de esta gráfica, consúltese la Figura 81.





**Figura 49. El tenor digital en la sección «Registro», formas verbales en modo imperativo formales e informales.**

Este caso atestigua, al suponer una apelación directa a que el usuario interactúe con el sitio web, la convencionalización del tono con una clara primacía del tratamiento de usted, el 80%, o el infinitivo, con el 20%. Estos resultados se asemejan por tanto a los obtenidos en el bloque «contacto».

No obstante el análisis del tercer elemento elegido, que se corresponde con la sección de trabajo (G1), muestra una clara diferencia con los anteriores. Concretamente, del total de formas verbales usadas en la denominación de esta sección, el 70% de las mismas se formula mediante el tratamiento «tú», el 20% con el tratamiento de «usted», y el 10% mediante la forma verbal infinitiva. Las siguientes gráficas muestran un claro contraste en la formulación del tenor en el bloque «contactar» y la sección «registro», en los que prima el tratamiento de «usted». En este caso, la necesidad de persuadir al usuario para que trabaje con la compañía lleva a un tratamiento más informal, con lo que se aprecia un cambio en las relaciones de poder. En este caso el emisor y receptor se situarían en un mismo plano en las relaciones de poder<sup>338</sup> que sería congruente con los datos expuestos en la siguiente gráfica:

---

<sup>338</sup> Este aspecto entroncaría con la dimensión cultural «distancia al poder» de Hofstede (1991) aplicados al diseño web por Sighn y Pereira (2005).

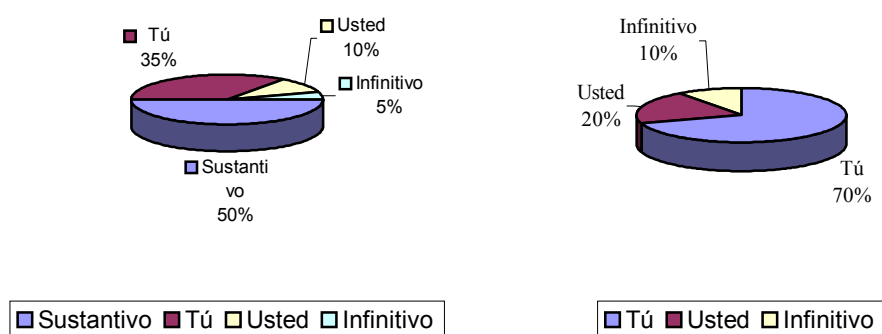


Figura 50. Tenor digital en la sección «Trabajo». Sustantivo o forma verbal<sup>339</sup>.

En resumen, los resultados obtenidos revelan en general la preferencia del tratamiento formal entre la empresa y el cliente en los textos originales, una tendencia compartida con los textos localizados, aunque en menor medida. Nuestros análisis han constatado que estos últimos textos presentan consistentemente un mayor uso de formas informales en comparación con los originales, en los que podemos observar que el tratamiento informal se suele reducir al 20% de casos. No obstante, el análisis de la sección de trabajo (G1) demuestra que el grado de formalidad entre la empresa y el usuario no es único y homogéneo, sino que puede variar dependiendo del destinatario de cada bloque o sección del sitio web. Las implicaciones de estos resultados son de importancia en la labor traductora hacia el español y se analizarán de manera contrastiva posteriormente con los textos localizados.

### 8.1.3. Análisis del contexto sociocultural

Los factores relativos al contexto sociocultural son fundamentales en el análisis de las características extratextuales del género *sitio web corporativo*, puesto que pueden influir en su creación y localización. El objetivo de este apartado será analizar cómo se refleja este aspecto en la evolución y estado actual del mismo. Las

<sup>339</sup> Para la lista completa de palabras usadas para la elaboración de esta gráfica, consúltese la Figura 73.

dos secciones del corpus, aunque parten de entornos socioculturales distintos, se dirigen a unos receptores situados en un contexto de recepción similar y, en virtud de esta divergencia, la influencia del contexto sociocultural de partida de los textos localizados puede manifestarse en el texto final, una de nuestras hipótesis de partida.

De entre los aspectos de interés en el contexto sociocultural, nuestro análisis incide principalmente en el marco normativo, jurídico y socioprofesional. No obstante, y siguiendo de nuevo a Gamero (2001), hemos de mencionar que las normas y convenciones prescritas no han de reflejarse automáticamente en este tipo de textos. Así, a pesar de la existencia de manuales de redacción web, manuales de usabilidad, normas internacionales o legislación que atañe a los sitios web, este conjunto de normas o recomendaciones pueden no cumplirse. El siguiente epígrafe se centrará principalmente en el marco normativo, que hemos dividido según una doble vertiente que coexiste y condiciona los textos web: por un lado se encuentran las normas o estándares funcionales centrados en la interoperabilidad y el intercambio de la información en el medio de Internet y de carácter internacional (Wright 2006), y por otro lado, se encuentran la legislación y normativa vigentes de cada región a la que han de someterse las empresas en cuanto responsables de un sitio web.

### **8.1.3.1. Marco normativo**

En el contexto español, la Agencia Española de Normalización (AENOR) es el organismo encargado de la elaboración de las normas españolas UNE. Ésta agencia quedó establecida por la Administración Pública de acuerdo con el Real Decreto 2000/1995. Es una entidad privada, sin ánimo de lucro e independiente que se encarga de tareas de normalización y certificación en nuestro ámbito geográfico.

La única norma identificada en nuestra investigación que atañe a los sitios web corporativos es la norma UNE 139803: 2004, «Aplicaciones informáticas para personas con discapacidad», desarrollada según las directrices internacionales del consorcio W3C que regula Internet. La agencia ha desarrollado en la actualidad una certificación AENOR de accesibilidad, otorgada hasta la fecha mayoritariamente a sitios web institucionales, a excepción de un solo sitio web corporativo<sup>340</sup> que ha recibido tal certificación (El Mundo, 9-5-2007).

En el marco internacional existen varios organismos de normalización que cuyas labores de regularización influyen en los sitios web corporativos. Wright (2006) resume las diferentes normas entre:

- **Normas de formato o base:** incluyen los diferentes lenguajes de etiquetado (SGML, XML, HTML, etc.), los usos de metainformación, los estándares que gobiernan la codificación de caracteres, los protocolos de acceso y la interoperabilidad de la información. En este sentido es esencial destacar la labor normalizadora que realiza el W3C en el desarrollo y evolución de estos estándares.
- **Normas de traducción:** definen los parámetros para negociar los contratos de traducción, localización y proporcionan elementos métricos para evaluar las traducciones (el Modelo de QA de LISA), así como facilitar el intercambio y segmentación de las unidades de memoria de traducción (LISA Translation Memory eXchange o TMX, SRX, etc.). En nuestro contexto es de importancia el estándar TransWS (*Translation Web Services*) sobre Localización Web, aunque a través de nuestro corpus no podemos obtener información alguna sobre la posible aplicación o no de estos estándares.

---

<sup>340</sup> Se trata del sitio web de Cajastur. Las Cajas de Ahorros se distinguen por su carácter social, de ahí la importancia de la accesibilidad en su sitio web.

- **Normas de terminología y lexicografía** como la norma ISO TC 37 (*Terminology and language content resources*).
- **Normas de taxonomía y ontología:** como las normas para tesauros de la ISO y la ANSI y la Lengua de Ontología Web (OWL), así como la norma del W3C *Simple Knowledge Organization Systems* (SKOS).<sup>341</sup>
- **Las normas sobre corpus:** tratan las convenciones para el etiquetado semántico, sintáctico y lógico de los corpus. Un ejemplo serían las normas del consorcio EAGLES.

En el contexto normalizador de Internet, requiere una mención los organismos como la IANA<sup>342</sup> (Internet Assigned Numbers Authority) que controla las direcciones IP, la asignación de nombre de dominios o los parámetros de los protocolos de conexión. Hemos asimismo de mencionar la IETF<sup>343</sup> (Internet Engineering Task Force) que desarrolla y promueve estándares en Internet junto con el consorcio W3C y la ISO<sup>344</sup>, la Organización Internacional de Normalización.

Si procedemos a analizar en impacto de estos estándares en nuestro corpus, las mayores diferencias cuantificables serían las referentes a la codificación. Los estándares de codificación de caracteres encontrados son UFT-8, ISO-8859-1, y Windows-1252. Si contrastamos las codificaciones entre ambas secciones del corpus, queda patente esta notable diferencia recogida en la siguiente gráfica.

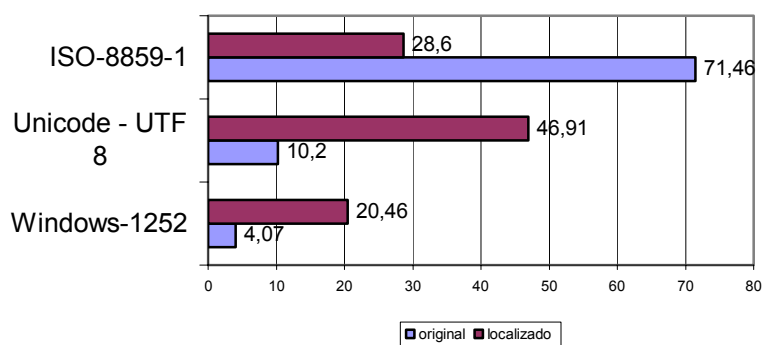
---

<sup>341</sup> Estos estándares son esenciales en el desarrollo de la Web Semántica, Web, 2.0 centrada en la descripción de la información.

<sup>342</sup> <<http://www.ietf.org/>>

<sup>343</sup> <<http://www.iana.org/>>

<sup>344</sup> Fue creada en 1947 destinada a promover el desarrollo de actividades de normalización en todo el mundo y está formada en la actualidad por 109 organismos nacionales de normalización como AENOR en España o DIN en Alemania.



**Figura 51. Porcentajes de uso de codificación en ambos corpus.**

En el corpus original la codificación mayoritaria es ISO-8859-1, con un 71,46% de páginas web, seguido de UTF-8, con 10,2% y 4,07 en Windows-1252. En la sección localizada identificamos una codificación mayoritaria en Unicode, con un 46,91% de inclusión de UTF-8, un 28,6% de inclusión de ISO-8859-1, y un 20,46% de codificación en Windows-1252. El carácter internacionalizador del texto localizado queda patente en la mayor inclusión del estándar de codificación Unicode. Este reciente estándar constituye un sistema de codificación de caracteres diseñada para acomodar la totalidad de posibles caracteres del conjunto de lenguas del mundo<sup>345</sup>.

### 8.1.3.2. Marco jurídico

En nuestro análisis hemos identificado un bloque comunicativo referente a las cuestiones legales que afectan al sitio web (véase 9.1.2.6 y 10.3). En el mismo se han identificado cinco posibles relaciones legales que lo condicionan.

- La relación legal entre el usuario y la empresa por cuanto a la protección de los datos suministrados por el primero. Este aspecto se encuentra regulado en España por la Ley Orgánica de Protección de Datos de Carácter Personal (Ley orgánica 15/1999, de 13 de diciembre). En España existe además una agencia gubernamental

<sup>345</sup> Para más información consúltase el sitio web de Unicode:  
<<http://www.unicode.org/standard/standard.html>>

encargada de velar por la protección de la información personal, la Agencia Española de Protección de Datos<sup>346</sup>. Por su parte, en el contexto estadounidense las consideraciones legales sobre privacidad no se encuentran reguladas por ley como en España, sino que las propias empresas se autorregulan bajo el auspicio de la Federal Trade Commission (Liu y Arnet 2000). En una investigación sobre políticas de privacidad de los sitios web corporativos estadounidenses, los investigadores de Liu y Arnet (ibid) apuntan que la FTC sólo promueve la adecuación de las prácticas empresariales a sus directrices, puesto que éstas no se regulan por imperativo legal:

The Federal Trade Commission (FTC) promotes adherence to these principles to insure effective privacy protection. Currently, the FTC negotiates between businesses that want to self regulate and the legislature that has an obligation to protect privacy, regarding fair information practice. But complete compromise is not likely (Liu y Arnet 2000: 14).

En resumen, los principios propuestos por la Federal Trade Commission estadounidense se centran en: (1) avisar en el momento que se proceda a recoger información personal, (2) el acceso posterior a la información recogida por parte de los usuarios, (3) el consentimiento en el uso que la empresa haga de esta información, y por último (4) la seguridad e integridad de la información (Federal Trade Commission 1999, cf. Liu y Arnet 2000: 14). Esta autorregulación conlleva la necesaria aparición de un bloque textual que recoja las disposiciones legales de cada empresa en los sitios web corporativos estadounidenses. No obstante, este bloque referido a la protección de datos es opcional en sitios web que no recojan información personal alguna (ibid).

- El cumplimiento por parte de la empresa, como prestadora de servicios a través de Internet con carácter de lucro, de la Ley 34/2002, de 11 de julio, «de Servicios de la Sociedad de la Información y de Comercio Electrónico». Las áreas que se ven afectadas por esta ley se recogen en la exposición de motivos del texto legal:

Se acoge, en la Ley, un concepto amplio de «servicios de la sociedad de la información», que engloba, además de la contratación de bienes y servicios por vía electrónica, el

---

<sup>346</sup> <<https://www.agpd.es/index.php>>

suministro de información por dicho medio (como el que efectúan los periódicos o revistas que pueden encontrarse en la red), las actividades de intermediación relativas a la provisión de acceso a la red, a la transmisión de datos por redes de telecomunicaciones, a la realización de copia temporal de las páginas de Internet solicitadas por los usuarios, al alojamiento en los propios servidores de información, servicios o aplicaciones facilitados por otros o a la provisión de instrumentos de búsqueda o de enlaces a otros sitios de Internet, así como cualquier otro servicio que se preste a petición individual de los usuarios (descarga de archivos de vídeo o audio...), siempre que represente una actividad económica para el prestador (BOE 166: 25388-25403).

- El registro legal que le confiere a la empresa su estatus jurídico como tal según la legislación española. Esto conlleva la inclusión del CIF (Código de Identificación Fiscal) de la empresa y sus datos de inclusión en un registro mercantil en España.
- La relación jurídica vinculante entre empresa emisora del texto y el receptor expresada en las «condiciones de uso».
- En el caso de localizaciones desde un país situado fuera de la Unión Europea, la legislación vigente de protección de datos de aplicación a empresas que recojan datos personales dentro de la Unión es la Directiva Europea de Protección de Datos 95/46/CE<sup>347</sup> del Parlamento Europeo y del Consejo del 24 de octubre de 1995. Esta directiva regula la protección y uso de los datos personales que se almacenen tanto dentro como fuera de la Unión Europea. Existe además una Directiva Europea de Privacidad de las Comunicaciones Electrónicas del 2003<sup>348</sup>.
- Finalmente, los sitios web respaldados por fondos federales estadounidenses se encuentran sujetos a una ley referente a la accesibilidad; las enmiendas de 1998 a la Sección 508 de la Rehabilitation Act (McAlpine 2001). Esta enmienda requiere que toda información digital sea accesible a personas con discapacidades: «electronic and information technology developed, procured, maintained, or used by the federal government be accessible to people with disabilities» (ibid: 101). Esta enmienda supuso una mejora exponencial de la accesibilidad y se tradujo, por ejemplo, en un crecimiento exponencial de las etiquetas «Alt» que describen el contenido de imágenes (véase 9.2.2). Las guías de estilo señalan que las grandes empresas han de aspirar a grados similares de accesibilidad en sus sitios web

---

<sup>347</sup> <[http://ec.europa.eu/justice\\_home/fsj/privacy/](http://ec.europa.eu/justice_home/fsj/privacy/)>

<sup>348</sup> <[hmsso.gov.uk/si/si2003/20032426.htm](http://hmsso.gov.uk/si/si2003/20032426.htm)>



corporativos (Jeney 2007), aunque como mostraremos con posterioridad, nuestro análisis intralingüístico no respalda esta aspiración.

#### **8.1.4. Marco socioprofesional**

La pertinencia de analizar el marco profesional surge por la influencia directa que ejerce en los contextos de producción y traducción de los textos recopilados, puesto que las restricciones que éstos imponen se verán reflejadas en sus productos (Baker 1999: 285). En la creación y desarrollo web multilingüe coexisten los desarrolladores técnicos, los diseñadores gráficos y los redactores técnicos, que ulteriormente suelen colaborar con los traductores de las versiones localizadas. La labor de estos últimos se desarrolla inexorablemente en el seno de un equipo multidisciplinar compuesto por los desarrolladores, programadores y diseñadores web. Dado el carácter contrastivo entre textos localizados y originales de la presente investigación, nos es necesario delimitar el contexto de producción en ambos casos. En principio, asumimos que los textos originales han sido producidos por redactores técnicos o por desarrolladores web generalistas, mientras que hemos de asumir que los textos localizados, y puesto las empresas son de reconocido prestigio, han sido producidos por traductores-localizadores profesionales cuya lengua A sería el español de España<sup>349</sup>.

##### **8.1.4.1. El redactor técnico o redactor web**

El redactor técnico o redactor técnico web es un especialista que se encarga de redactar todo tipo de documentos técnicos o textos distribuidos por las empresas. Esta profesión se encuentra muy extendida en el ámbito estadounidense, tal como atestigua la gran cantidad de titulaciones oficiales de pre y pos-grado presentes en

---

<sup>349</sup> En la recopilación del corpus localizado no se encontró ningún texto que se identificara como traducción por parte de un hablante cuya lengua A no fuera el español.

la mayoría de las universidades de este país<sup>350</sup>. La profesión de redactor web surgió desde el redactor técnico tradicional, aspecto este que apoya Jeney (2007) en la introducción a su manual de redacción web al señalar: «most of the writing done for professional Web sites is technical and information-based» (2007: xiv). La labor del redactor web es de primordial importancia y se equipara en valor al resto de componentes del diseño y desarrollo de sitios web (Nielsen y Tahir 2002: 14; Bly 2002. etc.). En lo que respecta al contexto profesional en España, coincidimos con Silvia Gamero (2001: 102) en que esta profesión es casi inexistente en nuestro país, aunque no obstante, se observa un antagonismo relativo a las carencias de centros de formación en este ámbito laboral y la existencia de una gran cantidad de centros universitarios que ofrecen titulaciones en traducción técnica.

La diferencias referentes al grado de profesionalización entre ambos ámbitos geográficos queda reflejada en el mayor número de publicaciones sobre manuales de estilo y redacción para Internet en el contexto anglosajón (Horton 1990; Anderson *et al.* 1998; Kilian 1999; Sammons 1999; Hale y Scanlon 1999; McAlpine 2001; Bly 2002; Garrand 2001; Hammerich y Harrison 2001; Usborne 2002; Price y Price 2002; Lynch y Horton 2002; McGovern *et al.* 2002; Veloso 2005; Mill 2005; Jeney 2007; etc.). Por contraste, los dos únicos libros que tratan exclusivamente la redacción web en el contexto español serían los libros de Díaz Noci y Salaverría (2003), y de Vilamor (2001), aunque dirigidos exclusivamente a aquellos profesionales de la información que se aventuren en el mundo del periodismo digital o ciberperiodismo. Además de estas guías de estilo, hemos de mencionar que se en España se han editado traducidas tres guías de estilo web anglosajonas (Kilian 1999; Price y Price 2002; Lynch y Horton 2004), que en nuestra opinión

---

<sup>350</sup> Existen varios manuales de estilo web concebidos como libro de texto para estos cursos (Anderson *et al.* 1998; Jeney 2007, etc.).

constituyen un vehículo de introducción de convenciones anglosajonas de diverso tipo en los textos españoles<sup>351</sup>.

Es indudable que la producción de textos web en España es muy amplia. Estos suelen producirse por empresas integrales de desarrollo web, principalmente por expertos generalistas en diseño web sin una formación específica para la redacción en este medio. A pesar de la importancia de la redacción web en el desarrollo web y tal como apunta Nielsen y Tahir (2002:14): «effective content writing in one of the most critical aspects of all web design», los desarrolladores web con una formación tecnológica suelen encargarse de esta tarea. De hecho, esta realidad se apunta como uno de los principales responsables de los problemas lingüísticos que se encuentran en estos textos. Al igual que ocurre en la localización, en donde la supremacía de aspectos técnicos conlleva la reducción de la traducción a una simple equivalencia entre palabras (Biau y Pym 2006; Foralón 2006), en el mundo del desarrollo web se asume que cualquier hablante sin importar su competencia lingüística o en convenciones genéricas posee una competencia textual adecuada para la redacción de este tipo de textos:

In the online world, technical proficiency and entrepreneurship are valued, but words are not. The view is that anyone can write words. Writers [...] are considered the wage slaves of the new economy (Bly 2002: 4).

Esta falta de atención hacia los aspectos textuales se observa en el corpus original español, sobre todo en presencia de un lenguaje estereotipado y la inclusión de gran cantidad de errores ortográficos y erratas. De hecho, éstas son especialmente importantes en lo que respecta a la acentuación de palabras. La falta de atención a la calidad del texto constituye una dificultad añadida a nuestra investigación intratextual con herramientas informáticas, ya que la mayoría de palabras acentuadas aparecen duplicadas en su forma con y sin tilde. Como ejemplo, en

---

<sup>351</sup> Una futura vía de investigación de interés sería el análisis de estos manuales de estilo traducidos como vía de entrada de convenciones anglosajonas en los textos web originales en español.

nuestro corpus original observamos una frecuencia similar de la palabra «aquí» con tilde (n. 2113) y «aqui» sin tilde (n. 2069).

En la actualidad no existen en España empresas que se dediquen exclusivamente a la redacción técnica (Gamero 2001) o web, al igual que no existe programa de estudios universitarios aprobado con tal fin. No obstante, desde los estudios de periodismo se preparan a los redactores en las peculiaridades de la redacción periodística para Internet (Díaz Noci y Salaverría 2003: 23-25), ya que como señala Vilamor (2001: 59) se necesitan periodistas «con un adiestramiento complementario o diferenciado del redactor convencional para que pueda dar respuesta a estas nuevas necesidades [de redacción y conocimientos tecnológicos]».

Al igual que el entorno del localizador, la labor del redactor web se encuentra inmersa en un grupo de trabajo en el que cooperan programadores, artistas digitales, administradores de bases de datos, desarrolladores de interfaces gráficos, etc. (Jeney 2007: 2). El carácter heterogéneo del texto web se refleja en las competencias requeridas de los redactores web, ya que se demanda la redacción de elementos interactivos, textos legales<sup>352</sup>, presentaciones Flash, descripciones de productos, *banners* publicitarios, formularios, etc. Tim Garrand (2001) resume en su guía de estilo multimedia este carácter heterogéneo y exigente:

The necessity of writing for many media in the same production is as demanding on the multimedia writer as is dealing with interactivity. [...] the writer of interactive multimedia must be expert in a variety of techniques: writing to be read (journalism, poetry, copy-writing); writing to be heard (radio, narration); writing to be seen (presentations, films, video); plus writing for the special demands of the computer screen. This is because multimedia can easily incorporate many types of media in a single production or even a single screen, and multimedia can manipulate these media in ways not previously possible (Garrand 2001: 17).

---

<sup>352</sup> Tim Garrand (2001) señala que los textos legales en los sitios web deberían ser redactados por un abogado o experto legal.

El carácter heterogéneo de la labor del redactor web y la exigencia de poder asumir gran variedad de tipos textuales, género, campos, etc., se transvasa al campo de la localización web, ya que esta amplia diversidad textual se refleja en el subsiguiente proceso de traducción.

#### **8.1.4.2. El localizador de textos web**

El conjunto de textos localizados de nuestro corpus son el resultado de un proceso de traducción-localización producido por conjunto heterogéneo de traductores o localizadores de difícil delimitación. La profesión del localizador se suele enmarcar dentro del contexto más amplio de la traducción técnica<sup>353</sup> (Byrne 2006; Jiménez Serrano 2002: 186), aunque este perfil especializado y diferenciado se ha implantado durante los últimos años sobre todo en la combinación lingüística desde inglés hacia otros idiomas (Dunne 2006). A su vez, dentro de la localización existen diferencias entre la localización web y la de software en virtud de la mayor competencia tecnológica que requiere la última (Esselink 2006). La localización web puede realizarse por medio de las herramientas informáticas de memoria de traducción<sup>354</sup> y, consecuentemente, no podemos considerar que los traductores de los textos recopilados hayan sido expertos exclusivos en esta área. Como ejemplo, el consorcio eColore<sup>355</sup> realizó una encuesta en el ámbito británico cuyos resultados muestran que el 54% de los traductores usuarios de memorias de traducción traducen regularmente contenido web (Reinke 2004). Podemos por tanto concluir que en el contexto británico la localización web se realiza mayoritariamente por traductores generalistas sin formación específica en localización. Para nuestros

---

<sup>353</sup> Aunque el investigador señala que el carácter técnico se refiere principalmente al soporte, ya que el contenido del discurso puede ser eminentemente general (Jiménez Serrano 2002: 187).

<sup>354</sup> SDL Trados contiene la herramienta TagEditor para la localización de textos etiquetados. La mayoría de las herramientas de memoria de traducción contienen módulos para la localización de textos web.

<sup>355</sup> El consorcio eColore financiado por la Unión Europea se dedica a la facilitación de la docencia en localización en el entorno. El sitio web del consorcio es: <<http://www.ecolore\ecolore.leeds.ac.uk\xml\about\site9ed2.html>>

objetivos, sería de interés un estudio similar en el contexto español, aunque en principio podríamos asumir que los resultados podrían ser similares.

El análisis de nuestro corpus textual demuestra que la localización en general incorpora la traducción de manuales, folletos, anuncios, cartas comerciales, software, sitios web, procesos internos, atención al cliente, términos y condiciones de los contratos de productos, publicidad, etc., un campo de traducción mucho más amplio de lo que se podría pensar en principio. En definitiva, la localización es una especialización de la traducción en la que se engloba prácticamente cualquier tipo de traducción, aunque requiere unas competencias técnicas específicas (Foralon 2006; Quirion 2003; Gouadec 2003).

Es común la falta de atención a aspectos textuales en la localización, ya que en este contexto priman las consideraciones técnicas, económicas y de diseño (Pym 2004a). En general, las empresas suelen asumir que la traducción de contenidos lingüísticos no revierte mayores complicaciones y esta desatención queda patente en el texto final. Yunker (2003) en su libro sobre localización web señala esta falta de concienciación hacia los aspectos textuales:

Too often, companies translate their source-language web sites without giving much thought to the text. Just because text is effective in the source/language site is no guarantee it will be effective in the target-language site, no matter how well it's translated (Yunker 2003: 267).

El contexto de producción textual en el que se desenvuelve la labor localizadora es radicalmente diferente a la redacción web original. El primero se encuentra inmerso en un proceso tecnológico en el que prima el uso de memorias de traducción y bases terminológicas, perdiéndose el carácter «individualista» que ha caracterizado a la forma occidental de traducir durante siglos (Tricàs 2007). Este carácter de proceso tecnológico y económico en el que se encuadra la profesión del localizador repercute en el producto, el texto localizado (Shreve 2006), ya que

entre otras peculiaridades se pierde gradualmente la noción de texto como unidad (Biau y Pym 2006; Bowker 2006; Foralón 2006; Shreve 2006; Mossop 2006). El producto final surge de procesos de traducción múltiples debido tanto al entorno colaborativo y tecnológico, como a la reutilización de segmentos textuales pretraducidos. Este entorno de producción resulta en textos sin un estilo único y una falta de homogeneidad a diferentes niveles, cuyo resultado se ha denominado en publicaciones anteriores «sentence salad» (Bedard 2000) o «train wrecks» (Bowker 2006). A nuestro parecer el término que representa este concepto en español sería «textos *collage*» (Mossop 2006), textos localizados que presentan un estilo y terminología heterogéneos debido al impacto del entorno colaborativo y las memorias de traducción en su contexto de producción<sup>356</sup>. En nuestro estudio hemos identificado empíricamente una falta de homogeneidad terminológica (véase 11.3.1), un aspecto que sería congruente con estos estudios previos.

Gregory Shreve (2006) resume acertadamente el impacto de tal entorno de producción y los problemas que conllevan:

- Los traductores traducen en mayor medida en un procesamiento a nivel de oración (Dragsted 2002, cf. Shreve, 2006; Webb 2000).
- La presegmentación que proporcionan las memorias de traducción suponen que la unidad operativa de traducción es la oración, con lo que los traductores traducen en mayor medida dentro del microcontexto. Esto se combina con la posibilidad del microreuso, el uso a nivel suboracional de colocaciones, unidades léxicas. (Macklovitch y Russell 2002).
- Se pierde la noción de documento al dividir el texto en unidades oracionales (Bowker 2006), o en palabras de Macklovitch y Russel «the very notion of a document is lost» (Macklovitch y Russell 2002).

---

<sup>356</sup> El resultado de esta labor traductora con el uso de memorias de traducción resulta en palabras de Bowker «The general result [...] is a text that is inherently less cohesive or coherent, less readable, and of a lesser overall quality. It may be gramatically correct, but it risks containing oversimplified syntax, monotonous rhythm, and a lack of diversity» (Bowker 2006: 180).

- Escasos recursos y la escasez de tiempo a la hora de alimentar las memorias de traducción y las bases terminológicas, que se producen con unas fuentes de consulta limitadas.
- Los plazos limitados de entrega de los proyectos que reducen el tiempo necesario para la resolución de problemas y la búsqueda de documentación de consulta. Asimismo las memorias de traducción no se suelen actualizar con las correcciones en el proceso de edición o control de calidad.

A los puntos anteriores, hemos de añadir que la labor del localizador web puede estar influenciada por un texto original desarrollado con el usuario internacional en mente (Yunker 2003; McAlpine 2001; Bly 2002). En este caso, la terminología, las referencias culturales específicas, las metáforas o juegos de palabras específicos de una cultura pueden haber sido objeto de un control previo que se corresponde una comunicación en un inglés internacional neutro (Bouffard y Craignon 2006: 809). De igual modo, no hemos de olvidar las restricciones impuestas por el encargo de la traducción, ya que por motivos económicos puede exigir el uso de un «español internacional neutro» (Castro 1996), a pesar de que el texto localizado se presente al receptor final como una producción exclusiva para la región lingüística o *locale* de destino (Karsch 2007)<sup>357</sup>.

Como resumen, asumimos que los textos localizados recopilados han sido producto de un proceso de localización heterogéneo con la participación múltiple de localizadores, cuyas competencias<sup>358</sup> relativas a la redacción y peculiaridades del texto web no son homogéneas.

---

<sup>357</sup> Terminóloga asociada a la empresa Microsoft. Comunicación personal el 11 de diciembre del 2007 tras la conferencia del doctorando en la conferencia «Teaching Localization for Global Business Readiness» organizada por la Localization Industry Standard Association en Amberes, Bélgica.

<sup>358</sup> El concepto de competencia lo definimos siguiendo a Hurtado Albir (2001: 385) como la habilidad de saber traducir que incluye cinco subcompetencias, la lingüística, la extralingüística, la de transferencia o translatoria, la profesional y la estratégica.



#### 8.1.4.2.1. La competencia traductora en localización

Para comenzar, podemos definir la competencia traductora en general como «el sistema subyacente de conocimientos, aptitudes y habilidades necesarios para traducir» (PACTE 2001: 39). Dada la relativa juventud del campo de la localización, son escasos los modelos específicos de competencias requeridas en este ámbito profesional. Éstos se han desarrollado principalmente con objetivos didácticos (Foralon 2006; Auster mühl 2006; Gouadec 2003). A nuestro parecer, el modelo de competencia sobre la localización más completo hasta la fecha es el formulado por Debbie Foralon (2006: 212-216), e incluye:

- a) **Subcompetencias de gestión de proyectos** de localización.
- b) **Subcompetencias tecnológicas** centradas tanto en el uso de herramientas informáticas de traducción como el conocimiento básico de los diferentes lenguajes de programación.
- c) **Subcompetencias lingüístico-culturales** asociadas con la traducción en general.

En nuestra opinión hemos de señalar que, a pesar del nivel de concreción de la propuesta, en cierta manera sería reduccionista si la analizamos desde la perspectiva de modelos holísticos de competencia traductora (PACTE 2000, 2005, 2007). El modelo del grupo de investigación PACTE se compone de cinco subcompetencias<sup>359</sup> y un componente psicofisiológico. En la propuesta de Foralon, el primer grupo de subcompetencias se situaría en el contexto más amplio de la subcompetencia de conocimientos de la traducción<sup>360</sup>, y el segundo grupo dentro

---

<sup>359</sup> Las subcompetencias serían (1) subcompetencia bilingüe, (2) subcompetencia extralingüística, (3) subcompetencia de conocimientos sobre la traducción, (4) subcompetencia instrumental, (5) y subcompetencia estratégica que sería la más importante del conjunto. El componente psicofisiológico (6) consistiría en componentes cognitivos y actitudinales así como mecanismos psicomotrices.

<sup>360</sup> Que se define como «knowledge of the principles that guide translation (processes, methods and procedures, etc.) and the profession (types of translation briefs, users, etc.)» (PACTE 2005: 610).

de la subcompetencia instrumental<sup>361</sup>. Este modelo englobaría el resto de subcompetencias dentro de la denominada «subcompetencia lingüístico – cultural», que equivaldría a las subcompetencias bilingües y extralingüísticas del modelo de PACTE. La subcompetencia estratégica, la de mayor importancia en este último modelo (PACTE 2005: 610), no quedaría recogida en ninguna de las subcompetencias del primer modelo.

En nuestra opinión, el modelo de la investigadora no hace más que corroborar la mayor importancia en subcompetencias instrumentales y profesionales que se requieren del localizador. Esta puesta en relieve de ciertas competencias instrumentales/tecnológicas específicas corroboraría, por tanto, la existencia de dos perfiles socioprofesionales diferenciados y relacionados, el traductor y el localizador, aunque siempre teniendo en cuenta que la traducción forma parte intrínseca del proceso de la localización (véase 2.3). Este aspecto ya fue formulado por Gouadec en 2002:

On dira donc, pour être précis, que le traducteur se prévaloir de l'étiquette de localisateur s'il effectue l'ensemble de traitements du matériel - y compris la traduction - mais qu'il reste un traducteur intervenant dans une prestation de localisation s'il «se contente de traduire» (Gouadec 2002:52).

Además, el distinto grado en esta subcompetencia instrumental y profesional puede suponer la presencia de distintos perfiles laborales que dependen del mismo:

After the advent of the Web, a new division in labor appeared. In the field of Web page translation, one finds IT people who are not translators performing certain tasks, translators with technical knowledge performing other tasks, and finally plain old translators who do minimal work with not-textual elements (Mossop, 2006: 791).

Como reflexión final, hemos de advertir sobre el peligro de identificar la competencia localizadora con la exclusiva adquisición de ciertas competencias

---

<sup>361</sup> Definida como «knowledge related to the use of documentation sources and information technologies applied to translation» (PACTE 2005: 610).

instrumentales inexistentes en otros tipos de traducción, ya que esta suposición puede llevar a un déficit en las subcompetencias bilingües y extralingüísticas asociadas a la traducción de géneros digitales. Una encuesta realizada en el 2005 entre empresas del ramo mostró que las empresas demandan mejoras en sus traductores tanto en competencias tecnológicas (71%) como en las de traducción (67%) (Reineke y Sánchez León 2005). Como ejemplo, nuestros resultados demostrarán cómo el bloque comunicativo sobre información legal muestra los menores niveles de calidad en su traducción, que asociamos parcialmente a una deficiente competencia específica en la traducción jurídica de sus traductores/localizadores. En este sentido, si retomamos que los sitios web corporativos constituyen géneros complejos (Martin 1995) o textos complejos (Reiss y Vermeer, 1984) (véase 5.1.2), hemos de señalar que «plantan mayor dificultad de traducción, ya que requieren de una competencia textual más amplia»<sup>362</sup> (Hurtado Albir, 2001:477), y por tanto, los modelos didácticos de competencia localizadora han de dar cabida a esta realidad.

---

<sup>362</sup> Esta competencia textual se situaría dentro de la subcompetencia bilingüe en el modelo del grupo de investigación PACTE (2005, 2007) al que pertenece la investigadora Hurtado Albir.

## 9. ANÁLISIS INTRATEXTUAL CONTRASTIVO

Tras realizar el análisis extratextual, el siguiente paso en nuestra progresión será un análisis intratextual contrastivo y descriptivo. Nuestro análisis discurre dentro del *continuum* analítico propuesto por Baker (1992), del nivel extratextual al intratextual, y dentro de éste último, se procederá jerárquicamente desde aspectos macrotextuales hacia los microtextuales. Hemos de recordar que, en principio, los rasgos de caracterización extratextuales deben ser similares entre textos pertenecientes a un mismo género. Son precisamente los rasgos intratextuales como las convenciones de género los que pueden presentar variaciones dentro de un mismo género y, por consiguiente, los que nos permitirán cuantificar las posibles diferencias entre ambos conjuntos textuales. Desarrollaremos el análisis intratextual de acuerdo a criterios principalmente cuantitativos, procediendo según una metodología empírica de análisis contrastivo acorde con investigaciones previas (Göpferich 1995; Gamero 1998, 2001). El método de análisis se ha ajustado a las peculiaridades del texto web y, tras un análisis cuantitativo por textos completos, se realizará a partir de la superestructura que establecemos como *tertium comparationis* entre los textos originales y localizados.

Nuestra primera fase se centra así en el aspecto intratextual de mayor rango jerárquico, la superestructura, sobre la que realizaremos un estudio descriptivo del género web *sitio web corporativo*. Este estudio descriptivo producirá un perfil prototípico del género en el contexto sociocultural español, que podrá ser utilizado como herramienta de consulta durante la traducción, como parte del diccionario de género<sup>363</sup> propuesto por Hurtado Albir (2001) y Gamero (1998: 430) o «banco

---

<sup>363</sup> No obstante, nuestro estudio será monolingüe en contraposición con el carácter bilingüe de los trabajos de estas investigadoras.

textual» en terminología de Göpferich (1995: 454). Este perfil convencional se establecerá a través de los bloques textuales de distinto rango jerárquico que componen el género, y el mismo permitirá, en una segunda fase, obtener a partir de los datos obtenidos las convenciones de importancia de acuerdo con su frecuencia estadística. Finalmente, las superestructuras prototípicas de ambos perfiles textuales nos permitirán establecer una base de análisis contrastivo entre ambas secciones del corpus. Realizaremos análisis contrastivos sobre los textos globales o sobre los subtextos de diverso tamaño que identificaremos con los distintos bloques y secciones comunicativas del género en cuestión. Los puntos de convergencia y divergencia podrán ser analizados y constituirán una sonda hacia las peculiaridades del proceso de traducción durante la localización de sitios web.

En resumen, los objetivos del análisis intratextual descriptivo – contrastivo que proponemos se han desarrollado adaptando a nuestros objetivos los propuestos por Gamero (1998; 2001), y se centran en los siguientes aspectos seleccionados por su relevancia desde una óptica del proceso de traducción:

- a) Establecer un perfil cuantitativo del género web desde el análisis por textos globales de ambas secciones del corpus.
- b) Averiguar la superestructura prototípica del género en bloques y secciones a través de análisis de los menús de navegación, que dado su carácter metadiscursivo, representan la estructura hipertextual global. Ésta se corresponde con el modelo mental textual difuso que los usuarios esperan encontrar en este género (véase 5.3). En nuestro estudio hemos equiparado estos menús con los títulos, subtítulos, encabezamiento de secciones y texto en negrita en las investigaciones de géneros textuales lineares (Göpferich 1995; Gamero 1998, 2001; Fernández Sánchez 2004).
- c) Contrastar las superestructuras prototípicas entre los textos originales en español y los localizados a partir de los sitios estadounidenses con el fin de

analizar las áreas de convergencia - divergencia y analizar el impacto de los diferentes contextos de producción textual y las peculiaridades del proceso de traducción en sus superestructuras.

- d) Obtener a través de criterios estadísticos las formas lingüísticas convencionales asociadas a cada bloque o sección del género.
- e) Deducir qué convenciones de interés al traductor son prototípicas en este género a través del análisis descriptivo y contrastivo.
- f) Comprobar si los textos localizados se ajustan a las convenciones establecidas que obtendremos de los textos originales, o en otras palabras, comprobar si los textos localizados se asemejan a los desarrollados originalmente en España, es decir, si se ajustan a los objetivos que el sector de la localización marca para los productos localizados<sup>364</sup>.

### **9.1. La superestructura prototípica del sitio web corporativo español**

Como describimos en nuestro marco teórico, la superestructura textual constituye el aspecto intratextual de mayor rango jerárquico (Göpferich 1995a; Gamero 2001). Tras la recopilación del corpus procedimos a realizar un análisis de la superestructura mediante la identificación manual de los distintos bloques textuales a partir de los menús de navegación. Este análisis nos permitió establecer los distintos bloques y secciones comunicativos constitutivos de este género textual. El procedimiento de recuento que aplicamos fue descrito anteriormente (véase 5.3.1), y en el mismo desempeña un papel fundamental la concepción de que cada página web constituye un bloque textual de distinto rango en el sentido clásico de las investigaciones sobre género textual (Askehave y Nielsen 2005). A su vez, los bloques o secciones comunicativos que identificaremos constituirán nuestras variables de análisis contrastivo dentro de nuestro estudio empírico.

---

<sup>364</sup> Como se señaló anteriormente, si los productos localizados tienen «the look and feel of locally-made products» (LISA 2003: 5).

El establecimiento de toda convención relativa a la superestructura textual depende de la cultura en la que se desarrolle un género (Neubert y Shreve 1992), y puesto que asumimos que existirán diferencias, iniciamos nuestro análisis en el componente original del corpus. A través de nuestra metodología empírica hemos identificado ocho bloques comunicativos característicos del género *sitio web corporativo* español:

1. Páginas de inicio
2. Páginas de contacto
3. Páginas con información sobre la empresa
4. Páginas de productos o servicios
5. Páginas de noticias
6. Páginas con información legal
7. Zonas exclusivas de clientes
8. Páginas basadas en la interactividad con el sitio web

La Figura 52 muestra el nivel de convencionalización de cada uno de estos bloques principales básicos en los sitios web originales. Nuestro análisis pone de manifiesto que la superestructura de los sitios web corporativos españoles se caracteriza por presentar cinco bloques comunicativos de carácter recurrente, es decir, con una frecuencia superior al 50%: las **páginas de inicio** (A), 100%, el bloque **contacto** (B), 86,04%, el bloque de información sobre la **empresa** (C), 85,46%, el bloque en el que se ofrecen **productos o servicios**<sup>365</sup> (E), que se combinan con un 99,41%, y el bloque **noticias** (D), 54,04%. Con carácter ocasional encontramos los bloques comunicativos **avisos legales** (F), 45,34%, **zonas de clientes** (G), 22,09%, y las secciones que se basan en la **interactividad** sitio web-usuario (H), como el **mapa web**, 20,34%.

---

<sup>365</sup> El bloque productos-servicios se combina en los sitios web corporativos y se engloban en una entidad única ya que ambos presentan la actividad que realiza la empresa de interés para el usuario.

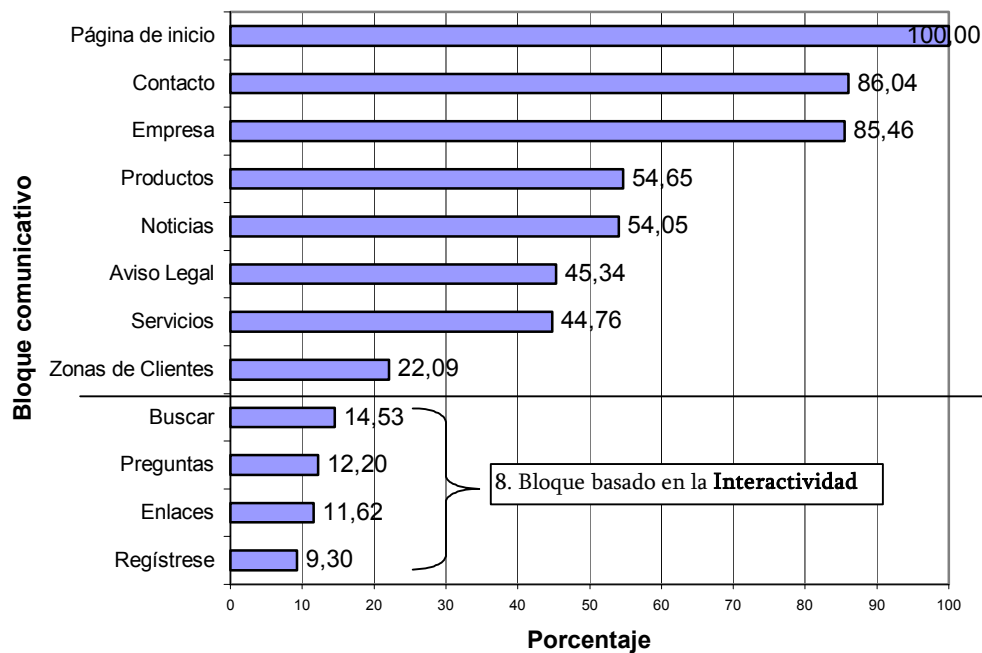


Figura 52. Grado de prototipicidad de bloques comunicativos en el género web sitio web corporativo español.

En un segundo nivel jerárquico, cada bloque comunicativo se puede subdividir en **secciones** con distinto grado de prototipicidad. En este género digital de los ocho bloques identificados, sólo cinco se pueden subdividir en secciones con diverso grado de convencionalización. En su conjunto, hemos identificado 58 secciones comunicativas en el seno de los sitios web corporativos españoles, que recogemos en la siguiente tabla. Esta tabla recoge la superestructura prototípica del corpus del género web, y se ha organizado de manera jerárquica a partir del porcentaje de frecuencia de cada bloque o sección comunicativa. Como se mencionó anteriormente, la metodología empleada para identificar los bloques se basa en la existencia de una página web independiente (Askehave y Nielsen 2005) y, en consecuencia, el porcentaje que recogemos es indicativo de su aparición en el corpus original como página independiente, que no de su posible frecuencia absoluta de aparición en el sitio web (véase 7.1.2):



Bloques y Secciones	% Corpus Original
<b>A. Pagina de inicio</b>	<b>100%</b>
<b>B. Contacto [interacción]</b>	<b>86,04%</b>
1) Formulario	31,42%
<b>C. Información sobre la empresa.</b>	<b>75%</b>
1) Localización	28,48%
2) Experiencia de la empresa	17,44%
3) Misión	10,46%
4) Sistema de calidad	9,88%
5) Historia	8,13%
6) Instalaciones	7,55%
7) Logística	6,49%
8) Medio Ambiente	6,39%
9) Proyectos	4,65%
10) Galerías de Imágenes	4,65%
11) I+D	4,06%
12) Alianzas	4,06%
13) Seguridad	4,06%
14) Divisiones	1,16%
15) Exportación	0,58%
16) Filantropía-fundación	0,58%
<b>D. Noticias/Actualidad</b>	<b>54,05%</b>
1) Novedades	13%
2) Eventos	6,98%
1) Notas de prensa	6,98%
3) Suscripción noticias	2,90%
4) RSS	1,16%
5) Notas de prensa	0,58%
<b>E. Productos -Servicios-Soluciones</b>	
1) Productos	53,48%
2) Servicios	44,76%
3) Ofertas	9,88%
4) Información técnica	5,81%
<b>F. Legal</b>	<b>45,34%</b>
1) Aviso-Nota Legal	27,90%
2) Declaración de privacidad	13,37%
3) Condiciones Generales	4,65%
4) Derechos de propiedad	2,90%
<b>G. Zona de Clientes</b>	<b>22,09%</b>
1) Trabajo	19,76%
2) Consejos	9,30%
3) Formación	5,81%
4) Tarifas	5,81%
5) Ayuda/atención técnica	5,81%
6) Pedidos	5,23%
7) Publicaciones	4,65%
i. Manuales/instrucciones	4,06%
ii. Folletos	2,90%
8) Inversores	4,06%
9) Profesionales	2,90%
10) Presupuestos	2,90%
11) Franquicia	2,32%

Bloques y Secciones	% Corpus Original
12) Regalos	0,58%
13) Financiación	0,58%
<b>H. Interactividad con sitio web</b>	
1) Mapa web	20,34%
2) Buscar	14,53%
3) Preguntas o ayuda [faq]	12,20%
4) Enlaces	11,62%
5) Regístrese	9,30%
6) Imprimir	6,97%
7) Glosario	5,23%
8) Descargas	4,06%
9) Comunidades	2,90%
10) Accesibilidad	2,32%
11) Su opinión	2,32%
12) Juegos	1,16%

Tabla 10. Frecuencia de aparición en el corpus original de las variables de análisis que responden a los bloques y secciones.

Con el afán de identificar cada bloque o sección en nuestra tesis, hemos asignado a cada bloque una letra por orden alfabético descendiente dependiendo de su frecuencia. Las distintas secciones comunicativas dentro de cada bloque también se han ordenado jerárquicamente por su frecuencia, aunque las hemos identificado mediante la combinación de la letra su bloque comunicativo y un número que denota su posición. Como ejemplo, la sección **privacidad** se identifica como «F2»: se sitúa en el bloque F, **legal**, y constituye la segunda sección en frecuencia tras **avisos legales** (F1).

La decisión de identificar las secciones como tales queda patente en el carácter ocasional (< 50%) de todas ellas. Así, las frecuencias de aparición recogidas en la tabla muestran que para todas las secciones son inferiores al 50%. Por su carácter como sección poseen una función textual parcial subsidiaria o secundaria. Las secciones con un mayor índice de frecuencia son los **formularios de contacto** (B1), 31,42%, la **localización** de la empresa (C1), 28,48%, **avisos legales** (F1), 27,90%, la sección comunicativa **trabajo** (G 1), 19,76%, y los **mapas web** (H1), 20,34%.

Este primer nivel analítico revela cómo ciertos bloques y secciones difieren en lo relativo a su **representación única** o **múltiple**. Estos conceptos se refieren a la formulación de un bloque o sección por medio de (1) una sola página web (representación única), o (2) la posibilidad de que un bloque se formule por medio de más de una página, como ocurre en el bloque **productos** (representación múltiple). En este género de carácter único y, por lo tanto, se representan por una sola página web: el bloque comunicativo **inicio** (A), el bloque **contacto** (B), el bloque **empresa** (C), el bloque **legal** (F), y la mayoría de las secciones en el bloque **interactividad** (H). Por su parte, son de carácter múltiple y se formulan mediante varias páginas: el bloque **noticias** (D), el bloque **productos-servicios** (E), y **zonas de clientes** (G). Estos dos bloques se caracterizan por estar normalmente representados por *e-textos* (véase 3.3.1), textos escritos para otro medio que se presentan en un sitio web. En este respecto, los porcentajes de frecuencia de nuestras variables no se encuentran ligados al número de páginas o extensión de cada bloque, sino a su mera presencia frecuencia de aparición.

Si comparamos estos resultados con los obtenidos sobre sitios web estadounidenses por Nielsen y Tahir (2002: 44-48), los investigadores identifican siete bloques comunicativos básicos: la navegación (*navigation*), la información corporativa (*about us*), la información de contacto (*contact us*), la política de privacidad (*privacy policy*), las ofertas de empleo (*Job info*), el mapa del sitio web (*site map*) y la función de ayuda (*Help*). Los autores basan su estudio en el análisis de 50 sitios web, y en nuestra opinión la estructura que presentan sería incompleta. En la misma se mezclan bloques de la estructura interna de cada página, como la navegación (véase 9.2), con los bloques textuales que incluimos en la superestructura como la información de contacto. No obstante, puesto que el estudio parte del contexto sociocultural estadounidense, es de interés para nuestro estudio cómo identifican el bloque textual legal con el término «política de privacidad» que, como veremos, constituye una sección comunicativa

sobrerrepresentada en la superestructura de los textos localizados en comparación con los originales (véase 9.1.1).

Contenido	Diseño
<b>1. Información de la empresa</b> -Biografías -Historia -Mensaje del Director General -Misión -Organización de la empresa -Prensa -Video de presentación	<b>Presentación</b> -Animación -Marcos -Gráficos -Sonido -Video
<b>2. Apoyo y Ayuda</b> -Teléfono de la empresa -Correo electrónico -Preguntas Más Frecuentes -Dirección de la sede principal -Chat de ayuda	<b>Navegación</b> -Enlaces a otros sitios -Contenidos protegidos -Búsqueda -Mapa del sitio
<b>3. Actualidad</b> -Contenidos actualizados -Indicador de la última actualización	<b>Seguridad</b> -Acceso seguro
<b>4. Información financiera</b> -Informes anuales -Información financiera básica	<b>Velocidad</b> -Tiempo de descarga de la página de inicio -Tiempo de descarga entre páginas
<b>5. Empleo</b> -Descripción de empleos en la empresa -Ofertas de empleo	<b>Control</b> -Uso de <i>cookies</i>
<b>6. Aspectos sociales</b> -Información sobre <i>cookies</i> -Sensibilidad cultural -Traducciones -Privacidad -Responsabilidad social	

Figura 53. Variables de análisis contrastivo de la superestructura (*content features*) y la forma (*design features*) de Robbin y Stylianou, (2003: 207).

En nuestra búsqueda bibliográfica sólo hemos identificado un estudio contrastivo que compare interlingüísticamente a través de procedimientos cuantitativos la superestructura prototípica de los sitios web corporativos multinacionales (Robbin y Stylianou 2003). Esta investigación supone un estudio exploratorio preliminar y surge desde la disciplina de las Ciencias de la Información, y en el mismo se establecen veintitrés posibles variables de análisis de la superestructura de este género. Los investigadores dividen estas variables entre las relativas al «contenido»

(en la primera columna), diferentes de las referentes a la forma o diseño (en la segunda)<sup>366</sup>.

En el establecimiento de las variables de análisis en esta investigación observamos un carácter eminentemente deductivo y cualitativo ya que no se incluye uno de los bloques obligatorios de mayor relevancia en este género: el bloque sobre los productos o servicios de la empresa. Además, ciertos bloques pertenecen lógicamente a un apartado superior, como sería en el caso de la información financiera que en nuestro estudio incluiremos dentro del bloque global sobre la empresa. Finalmente, el establecimiento de las variables de análisis no sería exhaustivo, aunque este aspecto queda justificado por los objetivos del estudio. En contraposición, las variables de análisis de nuestra tesis se obtuvieron de manera deductiva a partir de un análisis completo de todos los textos del corpus siguiendo el procedimiento de recuento que especificamos anteriormente.

A pesar de las posibles deficiencias que hemos observado en este estudio, una de sus conclusiones principales apoya la pertinencia de nuestro propio análisis: en el estudio se concluye que las diferencias interlingüísticas son mínimas en los sitios web corporativos en lo referente a la forma o diseño (gráficos, animaciones, etc.), mientras que las divergencias interculturales son significativas en el contenido del género *sitio web corporativo*, una variable que se correspondería con la superestructura en nuestro estudio.

### **9.1.1. Análisis contrastivo de la superestructura**

El siguiente paso en nuestra investigación consistió en obtener el perfil superestructural prototípico de los sitios web localizados siguiendo la misma

---

<sup>366</sup> El estudio incorporaría por tanto en su marco teórico el esquema de caracterización de Shepherd y Watters (1998) y de Schmitt-Isler (2000).

metodología que empleamos en el análisis de textos originales. A continuación realizamos un análisis contrastivo de carácter cuantitativo con el objetivo de identificar los puntos de convergencia o divergencia entre ambos perfiles textuales prototípicos. Las conclusiones de investigaciones interlingüísticas de género concluyen que la superestructura es dependiente tanto del género del que se trate (Göpferich 1995), como del entorno sociocultural en el que se desarrolle (ibid; Gamero 1998). La hipótesis de partida de esta fase de análisis se basa en la relación existente entre la superestructura genérica prototípica y su entorno sociocultural de origen (Neubert y Shreve 1992). Esta dependencia se refleja en posibles diferencias interlingüísticas en un mismo género (Göpferich 1995; Gamero 2001), que según nuestra hipótesis se verán recogidas en divergencias entre las superestructuras entre textos originales y localizados. En caso de que existan estas diferencias, las podremos atribuir a las diferencias entre ambos contextos de producción (Baker 1995).

Hasta la fecha no hemos identificado investigación alguna cuyo objetivo se centre en las diferencias superestructurales entre textos originales y traducidos en una misma lengua, aunque sí que se han investigado las diferencias superestructurales intralingüísticas entre subgéneros específicos. Las conclusiones de estas investigaciones apuntan que las diferencias superestructurales suelen ser más significativas entre subgéneros en una misma lengua, que entre un mismo género en varias lenguas (Vázquez del Árbol 2002, 2006; Fernández Sánchez 2004).

En el caso específico de la superestructura, defendemos que las limitaciones impuestas por el contexto tecnológico de los textos web podrían dificultar la realización de cambios durante el proceso de localización, ya que ésta se encuentra fijada y cualquier cambio requeriría un proceso técnico de mayor complejidad. En este sentido, pensamos que estos cambios revierten una mayor complejidad en

comparación con los textos impresos, ya que requieren una competencia técnica que puede no poseer el traductor-localizador.

La siguiente tabla contrastiva muestra la frecuencia de la totalidad de variables superestructurales identificadas que se corresponden con los bloques y secciones comunicativos posibles. La frecuencia en los textos originales se presenta a la izquierda, mientras que a la derecha se presenta la frecuencia en los textos localizados:

<b>Bloques y Secciones</b>	<b>Corpus Original</b>	<b>Corpus Localizado</b>
<b>A. Pagina de inicio</b>	<b>100%</b>	<b>100%</b>
<b>B. Contacto [interacción]</b>	<b>86,04%</b>	<b>94,73%</b>
Formulario	31,42%	26,31%
<b>C. Información sobre la empresa.</b>	<b>75%</b>	<b>86,31%</b>
1) Localización	28,48%	57,89%
2) Experiencia de la empresa	17,44%	30,52%
3) Misión	10,46	44,21%
4) Sistema de calidad	9,88%	11,57%
5) Historia	8,13%	33,25%
6) Instalaciones	7,55%	4,21%
7) Logística	6,49%	6,31%
8) Medio Ambiente	6,39%	13,61%
9) Proyectos	4,65%	0%
10) Galerías de Imágenes	4,65%	7,36%
11) I+D	4,06%	24,21%
12) Alianzas	4,06%	5,26%
13) Seguridad	4,06%	7,36%
14) Divisiones	1,16%	3,15%
15) Exportación	0,58%	1,05%
16) Filantropía-fundación	0,58%	6,31%
<b>D. Noticias/Actualidad</b>	<b>54,05%</b>	<b>79,94%</b>
2) Novedades	13%	9,47%
3) Eventos	6,98%	25,26%
4) Notas de prensa	6,98%	10,52%
5) Suscripción noticias	2,90%	5,26%
6) RSS	1,16%	6,31%
<b>E. Productos -Servicios-Soluciones</b>		
1) Productos	53,48%	84,21%
2) Servicios	44,76%	68,42%
3) Ofertas	9,88%	14,73%
4) Información técnica	5,81%	5,26%
<b>F. Legal</b>		
1) Aviso-Nota Legal	27,90%	47,36%
2) Declaración de privacidad	13,37%	70,52%

Bloques y Secciones	Corpus Original	Corpus Localizado
3) Condiciones Generales	4,65%	38,94%
4) Derechos de propiedad	2,90%	17,89%
5) Marcas registradas	0%	11,57%
<b>G. Zonas de Clientes</b>	<b>22,09%</b>	<b>38,94%</b>
1) Trabajo	19,76%	58,94%
2) Consejos	9,30%	23,15%
3) Formación	5,81%	21,05%
4) Tarifas	5,81%	3,15%
5) Ayuda/atención técnica	5,81%	29,47%
6) Pedidos	5,23%	25,26%
7) Publicaciones	4,65%	17,89%
i. Manuales/instrucciones	4,06%	5,26%
ii. Folletos	2,90%	4,21%
8) Inversores	4,06%	35,78%
9) Profesionales	2,90%	49%
10) Presupuestos	2,90%	2,1%
11) Franquicia	2,32%	0%
12) Regalos	0,58%	0%
13) Financiación	0,58%	5,26%
14) Recursos adicionales	0%	7,36%
<b>H. Interactividad</b>		
1) Mapa web	20,34%	60%
2) Buscar	14,53%	67,36%
3) Preguntas o ayuda [faq]	9,30%	35,78%
4) Enlaces	11,62%	15,78%
5) Regístrese	9,30%	22,10%
6) Imprimir	6,97%	6,31%
7) Glosario	5,23%	6,31%
8) Descargas	4,06%	17,73%
9) Comunidad	2,90%	5,26%
10) Accesibilidad	2,32%	2,1%
11) Su opinión	2,32%	9,47%
12) Juegos-	1,16%	5,26%
13) Su opinión	0,58%	9,47%
14) Herramientas	0%	9,47%

Tabla 11. Análisis contrastivo de la superestructura de los textos originales y localizados.

Este análisis contrastivo superestructural deja patente que ambos perfiles textuales se componen y comparten un mismo conjunto de bloques y secciones comunicativos, con un mínimo número de bloques exclusivos de uno u otro perfil textual.

De las sesenta y dos variables que recogemos, sólo hemos identificado tres secciones comunicativas de presencia exclusiva en los textos localizados: **marcas**



**registradas** (F5), **recursos adicionales** (G14) y **herramientas** (H14). En los textos originales solo hemos identificado tres secciones que no se encuentran representadas en los textos localizados: **proyectos** (C9), **franquicias** (G11) y **regalos** (C12). En este sentido, dada la escasa frecuencia de aparición de estas seis secciones, podemos concluir que no existen amplias diferencias en el número de bloques textuales posibles. No obstante, las diferencias más significativas serían las referentes a la frecuencia de aparición de ciertos bloques o secciones, por lo que en principio podemos concluir que existen diferencias entre sendos perfiles prototípicos, aun cuando se corresponden con un mismo género.

Nuestra primera matización sobre los resultados de esta tabla sería la similitud en la divergencia entre ambos perfiles textuales en lo que respecta a los dos bloques de mayor frecuencia: **contacto** (B) y **empresa** (C). Ambos difieren en un porcentaje cercano al 10%. Puesto que consideramos estos dos bloques obligatorios por su alta frecuencia, y el porcentaje de divergencia entre ambas secciones del corpus en estos dos bloques es similar, estableceremos el 10% como el porcentaje que atribuiremos a la pertenencia exclusiva de los textos localizados en un subgénero específico, los *sitios web corporativos de grandes empresas*. Estas diferencias ya quedaron patentes en el primer análisis cuantitativo sobre estos textos (véase 7.4), en donde observamos que arrojó una diferencia significativa en el número de páginas por sitio, 224,95 en localizados frente a 111,5 en originales.

Como ya apuntamos, el corpus de sitios web originales es representativo de la totalidad de los sitios web corporativos con independencia de la importancia económica de la empresa. Consecuentemente, hemos de asumir que esta diferencia en el número de páginas por sitio conlleva una mayor posibilidad de que en los textos localizados cada bloque o sección se formule mediante páginas web independientes. Como señalamos, hemos fijado esta divergencia en el 10% y

cualquier porcentaje superior será considerado como indicativo de diferencias superestructurales entre ambos perfiles textuales.

Si atendemos a los porcentajes de aparición de aquellos bloques y secciones comunicativos superiores al 50%, podemos obtener los perfiles superestructurales prototípicos de ambas secciones. Estos quedan recogidos en la siguiente tabla que incluye esta superestructura prototípica de base:

	Original		Localizado	
Bloques y secciones <b>altamente recurrentes</b>  > 70%	A. Página de Inicio	100%	A. Página de Inicio	100%
	B. Contacto	86,4%	B. Contacto	94,73%
	C. Empresa	75%	C. Empresa	86,31%
	E. Productos-Servicios <sup>367</sup>	99,41%	D. Noticias	79,94%
			E.Productos-Servicios	100%
		F2. Decl. Privacidad	70,52%	
Bloques y secciones <b>recurrentes</b> <70%, >50%	D. Noticias	54,05%	H2. Buscar	67,36%
			H1. Mapa Web	60%
			C.3 Trabajo	58,94%
			C1. Localización	57,89%

**Tabla 12. Análisis contrastivo de bloques y secciones recurrentes.**

El presente análisis muestra que el género *sitio web corporativo* español se caracteriza por cinco los bloques textuales prototípicos: la **página de inicio** (A), el **bloque de contacto** (B), el bloque **empresa** (C), el bloque **productos y servicios** (E) y el bloque **noticias** (D).

El perfil textual de los textos originales cuenta con cuatro bloques altamente recurrentes (>70%) y un bloque adicional recurrente (>50% <70%), mientras que la superestructura modular de los sitios web localizados incorpora cinco bloques con una frecuencia superior al 70%, junto con una sección, **privacidad** (F2), que se complementa con cuatro secciones con una frecuencia superior al 50%: **localización** (C1), **trabajo** (C3), **mapa web** (H1) y **buscar** (H2). A nuestro parecer, la

<sup>367</sup> Productos y Servicios se han combinado en un mismo bloque puesto que gran cantidad de empresas suelen ofrecer ambos.

presencia recurrente de estas dos últimas secciones de carácter metadiscursivo apunta a mayores niveles de interactividad en los sitios web localizados<sup>368</sup>. Por su parte, la alta frecuencia de las secciones **localización** (C1) y **trabajo** (C3) podría apuntar a la mayor envergadura económica de las empresas tras corpus localizado.

No obstante, el carácter altamente recurrente de la sección **privacidad** (F2) es indicativo de la introducción en los textos localizados de un rasgo convencional en el ámbito sociocultural estadounidense. Esta sección presenta una baja frecuencia en los textos originales (13,37%), y por su importancia contrastiva se tratará con mayor detenimiento en el análisis específico de este bloque textual (véase 9.1.2.6 y 10.3).

Una vez contrastados ambos perfiles prototípicos y, tras haber constatado las divergencias que podríamos atribuir a los diferentes contextos de producción, nuestro siguiente paso fue realizar un análisis contrastivo ampliado que pusiera de relieve aquellos bloques y secciones con un mayor diferencial de frecuencia en las variables de análisis. La siguiente tabla recoge este diferencial en porcentaje de prototipicidad entre los textos originales y los localizados. Los bloques y secciones están organizados jerárquicamente según el porcentaje de diferencia en su frecuencia de aparición representado en azul. La parte roja representa el porcentaje de aparición de los bloques o secciones en los textos originales, mientras que el valor global representa su frecuencia en los textos localizados.

Si analizamos detenidamente esta gráfica, la diferencia más representativa en la superestructura prototípica de ambos perfiles textuales es la mayor frecuencia de la sección comunicativa **privacidad** (F2) en los textos localizados, con un diferencial del 57,15%. Este hecho refleja claramente la influencia del contexto sociocultural del TO en los textos localizados, en especial por la autorregulación propia del

---

<sup>368</sup> En parte, como apuntamos anteriormente, por la mayor media de páginas web por sitio web en los sitios localizados.

contexto legal estadounidense (Liu y Arnett 2000), frente a la regulación por ley existente en España. Por su importancia, su impacto se analizará posteriormente.

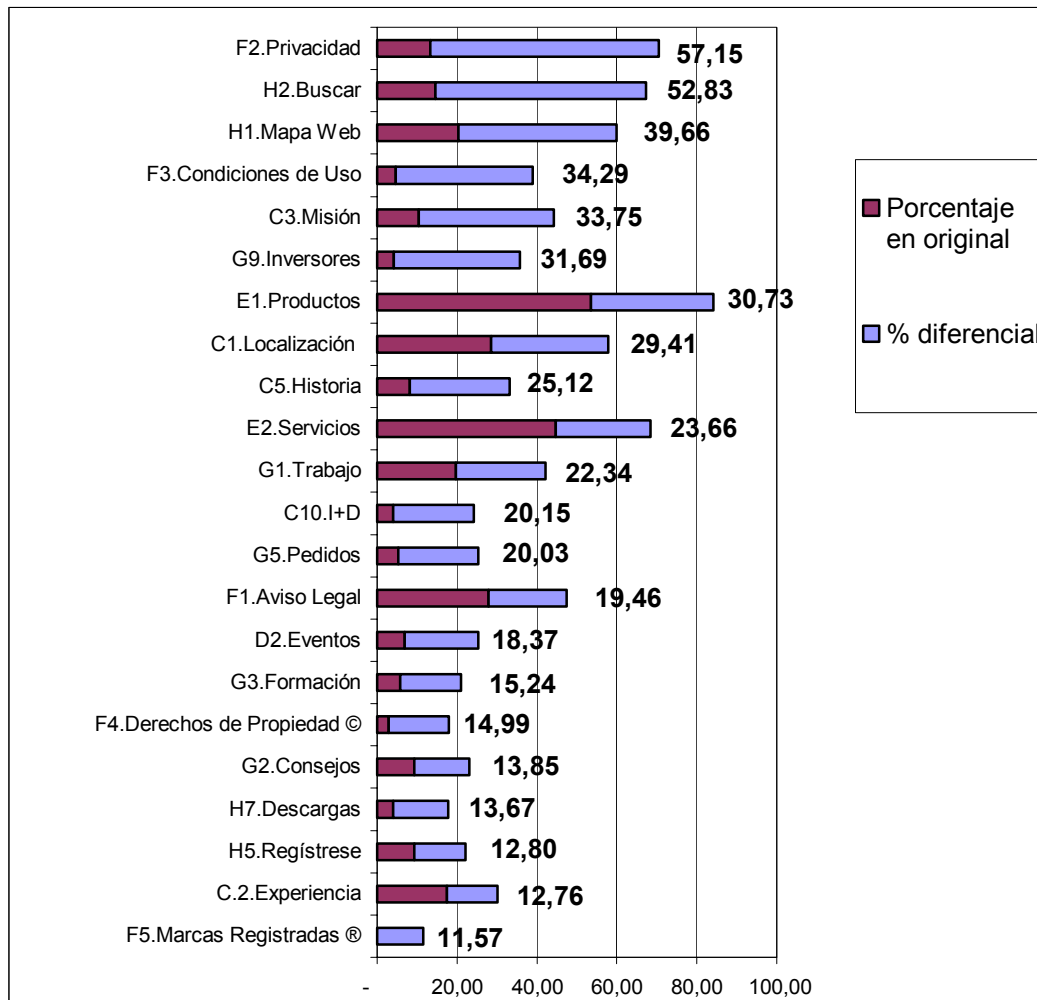


Figura 54. Bloques y secciones con mayor índice diferencial de frecuencia.

En general, la totalidad de las secciones de las que se compone el bloque **legal** (F) se encuentran sobrerrepresentadas en los textos localizados: **privacidad** F2 (+ 57,15%) **condiciones de uso** F3 (+34,29%), **avisos legales** F1 (+19,46%), **derechos de propiedad** F4 (+14,99%), y **marcas registradas** F5 (+11,57%), éste último inexistente en los textos originales. Esta mayor explicitación de los aspectos legales en los textos localizados es congruente con la conclusión principal de la investigación sobre sitios web corporativos de Robbin y Stylianou (2003): la mayor explicitación en los aspectos legales de los sitios web corporativos estadounidenses frente a los

de otros países. Las conclusiones de este estudio conllevarían que la diferencia que hemos identificado no sería exclusiva de la combinación lingüística que nos ocupa. No obstante, y como revelan nuestras estadísticas, esta diferencia entre textos originales estadounidenses y españoles se refleja en los textos localizados al español, con lo que se confirmaría en parte la hipótesis de partida según la cual son escasos los cambios superestructurales durante el proceso de localización web y, consecuentemente, los textos producto de este proceso muestran una superestructura clonada de los TO<sup>369</sup>.

Las diferencias más significativas en segundo lugar se encuentran en secciones pertenecientes al bloque comunicativo **interactividad** (H), tales como **buscar** (H2), 52,83% **mapa web** (H1), 39,66%, **descargas** (H7), 13,67% o **regístrese** (H5), 12,80%. En este sentido, cabe señalar que la mayor inclusión de estas secciones se debería parcialmente al mayor número de páginas por sitio web de los localizados. De hecho, la explicitación de la superestructura en un mapa web y la inclusión de una opción de búsqueda es más frecuente conforme aumenta el número de páginas: los manuales de desarrollo web indican que la necesidad de un mapa web y opciones de búsqueda a partir de 100 páginas (Nielsen y Loranger 2006: 138). La media en los sitios web originales es de 111, con lo que este diferencial podría inicialmente atribuirse a la mayor extensión de los sitios web localizados. La mayor importancia económica de las empresas emisoras de estos textos podría asimismo reflejarse en el mayor porcentaje de prototipicidad de las secciones que reflejan cuestiones económicas como **inversores** G9 (+31,69%), **productos** E1 (+30,73%), **servicios** E2 (+23,66%), **trabajo** G1 (+22,34%), **I+D** C11 (+20,15%), o **pedidos** G5 (+20,03%).

Finalmente, aparecen en tercer lugar un grupo de secciones referentes a la descripción de la empresa, como **misión** C3 (+33,75%) o **historia** C5 (+25,54%).

---

<sup>369</sup> El mantenimiento de la superestructura en un mismo género en el proceso de traducción no sería exclusivo del género que nos ocupa. Nobs (2006) muestra unas conclusiones similares en su investigación sobre folletos turísticos en la combinación alemán-español.

Estas divergencias podrían ser indicativas de diferencias socioculturales relativas a la inclusión de elementos de valoración subjetiva en el posicionamiento de la empresa frente a la competencia.

### **9.1.2. Análisis descriptivo-contrastivo de los bloques y secciones comunicativas**

Tras el análisis contrastivo global de las superestructuras prototípicas, pasamos a realizar un análisis contrastivo-descriptivo de los distintos bloques y secciones que componen en este género textual.

#### **9.1.2.1. Análisis del bloque «página de inicio» (A)**

Esta página se considera la de mayor importancia en este género (Garrand 2001; Nielsen y Tahir 2002), y en ella se presenta la información más importante para el usuario y un resumen de los contenidos que puede encontrar en el sitio web (Kilian 1999: 19). Este bloque sería obligatorio en los sitios web corporativos y queda recogido en el 100% de los sitios recopilados en ambas secciones del corpus. Si atendemos a su **función textual**, la página principal cumple dos funciones básicas: una expositiva en la que se comunica la afiliación y contenidos del sitio web, y otra exhortativa o instructiva con ciertos elementos textuales y visuales que nos animan a quedarnos en el sitio web o visitar ciertas partes (Askehave y Nielsen 2005: 123; Bolaños 2002a). Esta página sirve de «puerta de entrada» al sitio web, si bien los usuarios pueden acceder al sitio web directamente desde cualquier página sin pasar por este bloque<sup>370</sup>, un fenómeno que se conoce como *deep linking* (Janoschka 2003; Nielsen 2000: 179).

---

<sup>370</sup> Este hecho se debe a la concepción de cada página web como unidad de contenido y su posible indexación independiente en este medio.

La concepción de este bloque comunicativo se ha identificado con ciertas metáforas, como la portada de una revista<sup>371</sup>, la puerta de entrada, la tabla de contenidos, el recepcionista de la empresa, folleto informativo, etc., pero no todas se corresponden con la funcionalidad de este nuevo género (Nielsen 2000: 19).

De entre las posibles formas en que se puede presentar esta página, encontramos (1) la página principal que se asemeja a una página de presentación Flash caracterizada por la brevedad y la concisión, (2) la página de inicio que se asemeja a un anuncio y (3) finalmente aquellas que se asemejan a la portada de un periódico, con resúmenes por medio de párrafos cortos de los contenidos del sitio (Bly 2002: 134). Como en cualquier otro tipo de páginas, la brevedad, concisión y claridad son esenciales para que usuario la comprenda de forma adecuada, siendo preferible una presencia mínima de gráficos (Yunker 2003).

Las páginas de inicio de sitios web corporativos han sido objeto de investigaciones previas, sobre todo por que se identifica como la página más representativa en este género (Bolaños 2002, 2003; Bolaños *et al.* 2005; Renau 2004; Askehave y Nielsen 2005). Éstas presentan diferencias dependiendo del contexto sociocultural del que surgen (Rennau 2004), y poseen una estructura altamente convencionalizada (Nielsen y Tahir 2002). Dado que existen varias investigaciones sobre este bloque textual en nuestro país, nuestro análisis contrastivo se reducirá al grado de convencionalización de la terminología empleada en los menús de navegación.

#### **9.1.2.1.1. Análisis contrastivo de títulos en bloque «página de inicio»**

En el corpus original queda patente la convencionalización de la unidad léxica «Inicio» en la denominación de este bloque, con un 60% de frecuencia de uso. En

---

<sup>371</sup> La diferencia esencial que indican Nielsen y Tahir (2002) en este caso es que el usuario accede al sitio web deliberadamente y en cambio, una portada de revista necesita competir con el resto de publicaciones. En el 4% de los sitios web localizados y el 1,67% de los originales esta sección se denomina «portada» en clara referencia a esta metáfora.

segundo lugar se encuentra el anglicismo «*Home*» con un 24% de uso, que reflejaría una influencia directa del género estadounidense en los sitios españoles. Hemos detectado que esta influencia se debe en varios casos al uso de herramientas informáticas de desarrollo web localizadas que afectan directamente al diseño de los sitios web. Específicamente, se identificaron relaciones metafóricas dependientes del inglés tales como la caracterización de este bloque con el icono de casa, «home» en inglés, el icono más extendido por aparecer en el navegador Internet Explorer de Microsoft. Como ilustra la siguiente imagen, aparte de esta relación metafórica en el contexto sociocultural estadounidense, encontramos, la relación entre el icono de una lupa y la sección comunicativa «localización»:



Figura 55. Uso de iconos anglosajones en página web de inicio española. [www.caditesting.es](http://www.caditesting.es).

Otras unidades léxicas con menor nivel de convencionalización que representarían posibles alternativas en el corpus original son: «página principal» (6%), «principal» (4%), «portada» (4%) o «página de salida» (2%)<sup>372</sup>.

El análisis contrastivo muestra que los textos localizados comparten como término más frecuente «inicio» (44,64%). De igual manera, en los textos localizados aparece «home» (24,00%), «página principal» (8,93%), «principal», «Homepage» y «Página inicio» (3,75%) y «menú principal», «página de inicio» y «portada» (1,79%).

<sup>372</sup> En la investigación de Belda (2003) sobre la terminología en informática e Internet se señala «página principal» como traducción de «homepage». Nuestro análisis revela que el compuesto sintagmático nominal «página de inicio» sería el más frecuente con el 60% frente al 10% de uso de unidades léxicas que incluyen la forma adjetival «principal». Asimismo, recoge que en distintas publicaciones se muestran once posibles traducciones de «homepage»; en nuestro corpus representativo de los sitios web corporativos españoles se encontraron sólo cinco posibles unidades léxicas que pueden aparecer abreviadas, como «página principal» o «principal». Las distintas formas que recoge el autor son «página inicial», «página raíz», «página de casa», «página principal», «página frontal», «página de inicio», «portada», «página de bienvenida», «página de entrada», «página de presentación», «página de origen», «página hogar». Sólo tres de estas posibles traducciones fueron encontradas en nuestro corpus.



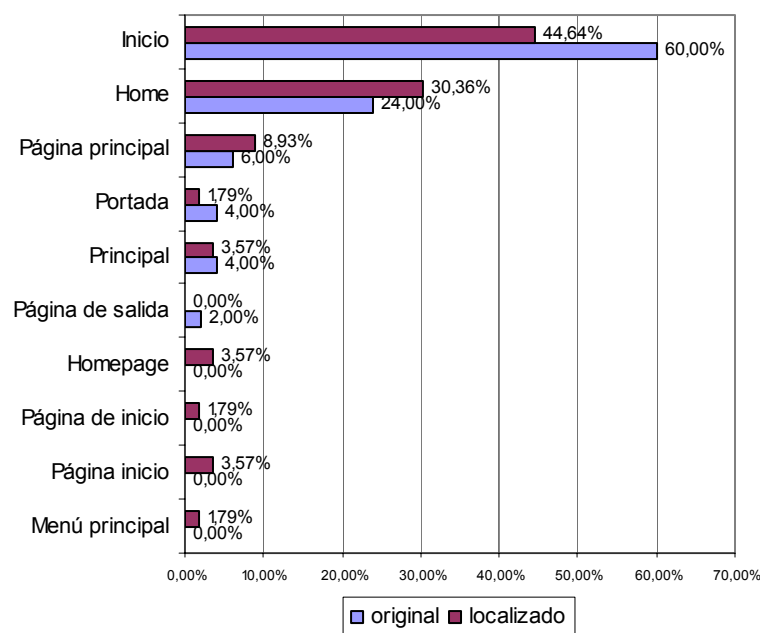


Figura 56. Análisis contrastivo de términos empleados en el bloque página de inicio (A).

La unidad léxica más convencionalizada en este bloque, «Inicio», respondería a una forma abreviada del compuesto sintagmático nominal «página de inicio». Este recurso se debe a necesidad de concisión en los menús de navegación y se manifiesta además en otros compuestos sintagmáticos abreviados como «principal» o «home». Dentro de los textos localizados hemos identificado otros préstamos o anglicismos como el compuesto sintagmático ortográfico *homepage* (Belda 2003: 35), o formas erróneas como «página inicio», 3,57%, ya que este compuesto nominal requiere la preposición «de» en español<sup>373</sup>. No obstante, la presencia de anglicismos es abundante en los textos originales, con un 24% de uso de «home» o la aparición del anglicismo «página de salida», inexistente en los textos localizados y que se corresponde con la unidad léxica «start page» en inglés.

<sup>373</sup> No obstante, en la terminología convencionalizada en otras secciones se observa que esta preposición en compuestos sintagmáticos nominales no es obligatoria en español, como el caso de «mapa web» que es la unidad léxica más convencionalizada en la sección H1.

### 9.1.2.1.2. Las páginas web de presentación

Dentro de las páginas web de inicio se suelen situar otro tipo de páginas con ciertas características propias: las páginas web de presentación o *Splash pages* en inglés que se incluyen en este primer bloque (A). Estas páginas se presentan cuando el usuario accede al sitio web y suelen consistir en animaciones con un concepto gráfico que pretenden atraer al usuario y mostrarle la alta calidad del sitio y, por ende, de la empresa. El formato de estas presentaciones suele ser Macromedia Flash. A pesar de lo extendido de su uso, las guías de usabilidad desaconsejan usarlas ya que ralentizan la acción del usuario que se dirige al contenido que busca (Brinck *et al.* 2002), y en conexiones lentas pueden incluso llevar al usuario a abandonar el texto (Nielsen y Loranger 2006). En la actualidad estas páginas han prácticamente desaparecido en el contexto de los sitios web corporativos de grandes empresas (Nielsen y Loranger 2006: 90). En caso de incluirlas, se aconseja incluir la opción de saltar o pasar la introducción, que en español se suele indicar con «Saltar introducción», «Saltar intro», calcos del inglés *skip intro*<sup>374</sup>.

En la recopilación del corpus original las presentaciones en Flash causaron ciertos problemas al impedir la descarga con la herramienta informática empleada, Htrack. Este hecho motivó la eliminación de 19 páginas de Splash<sup>375</sup> para poder descargar los sitios web completos<sup>376</sup>. En la sección localizada no identificamos ninguna página de presentación, con lo que podríamos concluir que se ajustan a las recomendaciones que desaconsejan su uso desde el punto de vista de la usabilidad, guías de estilo web y localización (Nielsen 2000; Nielsen y Loranger 2006; Brinck *et al.* 2003; Bly 2002; Yunker 2003; Lingo Systems 2004). Como mencionamos anteriormente, su localización supone una mayor inversión económica debido a las

<sup>374</sup> En español sería más correcto «pasar a contenidos».

<sup>375</sup> El término «página de presentación» se puede usar como sinónimo de página de inicio o *homepage*

<sup>376</sup> Aun cuando en nuestra definición del texto web forman parte del hipertexto que se recopila (véase 3.4).

peculiaridades de su proceso de traducción que se aleja de simplicidad del texto html (Macromedia 2004).

Uno de los usos más extendidos de este tipo de páginas en España es el uso como selector de idiomas, como se muestra en el ejemplo de [www.rianxeira.es](http://www.rianxeira.es). Estas presentaciones multimedia pueden incluir sonido o narraciones orales en la presentación<sup>377</sup>, como en el caso de [www.ceylan.es](http://www.ceylan.es). En la primera presentación no existe la opción de pasar al contenido principal, mientras que en el sitio web de la empresa Ceylan esta opción se recoge en el término en inglés *skip*, un anglicismo que podría dificultar el uso del sitio para usuarios españoles.



Figura 57. Pantallas animadas *Splash* en formato Macromedia Flash.

De entre las desventajas de este formato se encuentran los problemas de accesibilidad, un mayor tiempo de descarga, o la pérdida de tiempo del usuario ya que se interpone entre éste y la información que busca (Yunker 2003; Bly 2002).

### 9.1.2.2. Análisis del bloque «contacto» (B)

El bloque **contacto** (B) es altamente recurrente en el género de estudio, con una frecuencia de 86,04% en textos originales y 94,73% en localizados. Su alta

<sup>377</sup> Yunker (2003) recomienda reducir al mínimo la inclusión de sonidos en las presentaciones con Flash.

frecuencia de aparición nos apunta a la importancia clave del proceso de interacción comunicativa entre el usuario y la empresa emisora del texto. El nuevo medio permite que se establezca una comunicación con mayor rapidez, por medio de correo electrónico, formularios o incluso chats de contacto y atención al cliente (Janoschka 2003; Price y Price 2002). Además, es preciso añadir la información de contacto mediante canales tradicionales como números de teléfono, fax y dirección de correo postal, con lo que la aparición de este género ha multiplicado los medios de comunicación posibles. La evolución del género ha supuesto la reciente aparición de los formularios y chats de contacto que agilizan el intercambio comunicativo; los primeros aparecen en el 32,2% de los sitios web originales y 26,31% de los localizados. Por su parte, los chats presentan una escasa frecuencia en textos originales (0,58%) y no se encuentran en los textos localizados.

Contacto

Ingrese los datos requeridos en el siguiente formulario.

Los campos con ( \* ) son obligatorios

- Nombre y Apellido:\*
- Teléfono:\*
- Fax:\*
- E-mail:\*
- Consulta:

Enviar

Figura 58. Formulario de contacto con una función textual primaria exhortativa. [www.adea.es](http://www.adea.es).

La **función textual** de este bloque comunicativo es expositiva en descripción, al ser su objetivo principal presentar la información básica para comunicarse con la empresa. Su **función secundaria** es exhortativa con alternativa y queda recogida en los enlaces interactivos, como los del correo electrónico, que convidan al usuario a su activación (McGovern *et al.* 2002). Dentro de este bloque, los formularios de contacto presentarían una función textual diferenciada marcada por la interactividad propia del medio. Su objetivo básico es condicionar la «acción» del

usuario para que los cumplimente, por lo que argumentamos que su función primaria es exhortativa con alternativa. En estos formularios interactivos se redefine en parte la tipología textual exhortativa clásica, ya que están compuestos por formas sustantivas con los campos a cumplimentar en el formulario representando el componente apelativo. Por su interés y alto grado de convencionalización, analizaremos las peculiaridades textuales de estos formularios en el análisis exhaustivo formas lingüísticas convencionales en el próximo capítulo (véase 10.2.2).

La **funcionalidad** de este bloque se basa en la interacción entre el usuario y el emisor por medio de correo electrónico, formularios de contacto y en un pequeño número de casos, chat de contacto y ayuda. La interactividad con el medio forma también parte del intercambio comunicativo, ya que la dirección de correo electrónico suele producir una ventana de correo emergente al activarla mediante la etiqueta html <mailto: >, así como la posibilidad de interactuar con los formularios y páginas.

#### **9.1.2.2.1. Análisis contrastivo de títulos en el bloque «contacto»**

Si procedemos a analizar contrastivamente la terminología empleada para denominar este bloque en los menús de navegación, observamos que en los textos originales la unidad léxica más convencionalizada y frecuente es la forma sustantiva «contacto» (49,66%), seguida de la forma verbal «contactar» (14,48%).

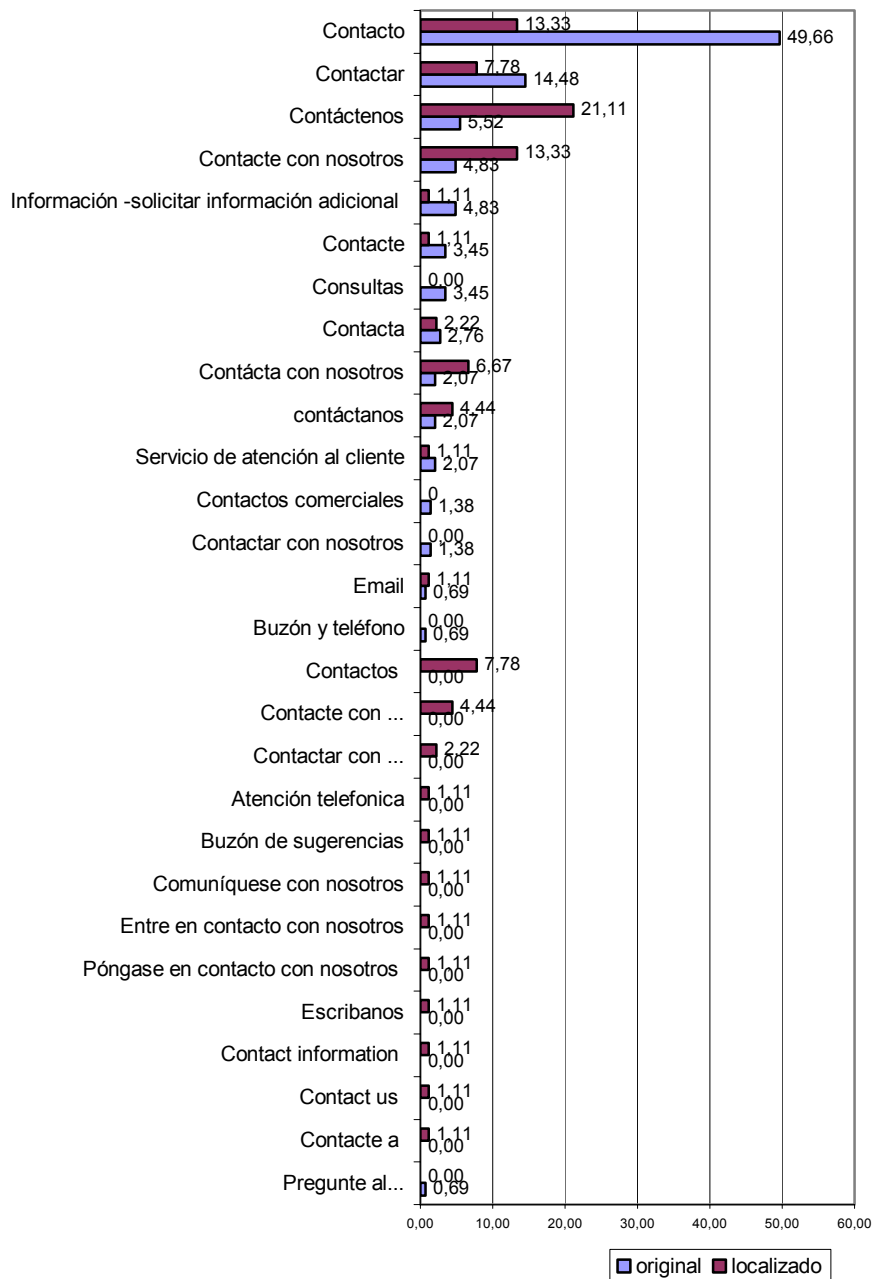


Figura 59. Análisis contrastivo de terminología en bloque contacto (B).

Sin embargo, este alto nivel de convencionalización de «contacto» no se comparte con los textos localizados, puesto que la unidad léxica de mayor frecuencia es la forma verbal en modo imperativo con pronombre enclítico «contáctenos» (21,11%), seguida de otra forma en modo imperativo, «contacte con nosotros» (13,33%) y el sustantivo «contacto» (13,33%). En esta discrepancia en la unidad léxica más frecuente entre ambas secciones del corpus se puede vislumbrar el alto

grado de convencionalización de la unidad léxica de carácter verbal «contact us» (89%) en los sitios web corporativos estadounidenses (Nielsen y Tahir 2002).

El análisis contrastivo muestra que los textos localizados presentan una mayor frecuencia de las cuatro unidades léxicas que incorporan el pronombre «nosotros» de forma enclítica o unido sintagmáticamente por medio de la preposición «con»: «contáctenos», «contacte con nosotros», «contacta con nosotros», y «contáctanos». Globalmente, el 51,1% de unidades léxicas localizadas incorporan una forma pronominal, mientras que sólo el 15,7% de términos usados en textos originales la incluye. Es igualmente significativa la mayor presencia de formas verbales en las unidades léxicas localizadas (71,11%) frente a las originales (41,38%), que se decantan por forma nominales en su formulación. Esta diferencia respondería en nuestra opinión a una influencia del carácter verbal de «contact us».

A pesar de la relativa convencionalización en los textos originales de la forma léxica «contacto», el hecho más significativo sería el escaso seguimiento de esta convención en los textos localizados, que a su vez muestran una gran variación denominativa en la expresión de esta forma lingüística convencional. Dentro del conjunto de unidades léxicas localizadas hemos identificado una considerable cantidad de alternativas inexistentes en el corpus original y, de entre ellas, varias son contrarias a las recomendaciones de brevedad y concisión en los menús de navegación (Price y Price 2002; Nielsen y Loranger 2006): «entrar en contacto con nosotros», «póngase en contacto con nosotros» o «comuníquese con nosotros». La aparición de estas combinaciones léxicas inexistentes en los textos originales podría apuntar a un proceso de explicitación<sup>378</sup> en su traducción. En total, hemos identificado recogido 23 posibles alternativas en la denominación de este bloque en 95 sitios web, frente a 14 en la sección original en 172 sitios web. Si asumimos como apuntan Nielsen y Tahir (2002) que el 89% de los sitios web estadounidenses

---

<sup>378</sup> El concepto de explicitación se define como «an overall tendency to spell things out rather than leave them implicit» (Baker 1996: 180).

presentan «contact us», podríamos concluir que las selecciones terminológicas de los traductores durante el proceso de localización serían las responsables de la variación observada, con doce posibles alternativas que no existen en el corpus original (24,73%).

Otro dato de interés en las unidades localizadas lo constituye la presencia de ciertos sinónimos de «contacto» inexistentes en el corpus original, como «comunicarse» o «escribanos». Por último, hemos de destacar el posible carácter de anglicismo léxico y sintáctico de las unidades léxicas formadas a partir del lexema [contact\*], en consonancia con Bouffard y Craignon (2006) relativo a su uso en francés. De ahí que creamos preciso resaltar la presencia del término «consultas» en los textos originales (3,45%), ya que el objetivo de los usuarios tras el acceso a esta sección no es contactar *per se* con la empresa, sino realizar una consulta. Como síntesis final, queda patente que existen diferencias tanto en lo relativo a las unidades léxicas de mayor frecuencia como en el número posible de alternativas para expresar esta forma lingüística convencional, en parte por la influencia de la forma convencional en el contexto sociocultural estadounidense.

### **9.1.2.3. Análisis del bloque «empresa» (C)**

Como medio de comunicación, en Internet se presentan en un plano igualitario todas las empresas (Nielsen y Tahir 2002: 46). En este contexto, toda empresa que emita un texto web debe establecer sus credenciales y credibilidad, el objetivo principal de este bloque comunicativo (Mill 2005: 85; Nielsen y Loranger 2006). En descripciones previas se ha defendido que este bloque puede asemejarse a un folleto informativo sobre la empresa condensado en uno o varios párrafos, aunque se ha de optimizar para su lectura en pantalla (Bly 2002: 183). Sin embargo, defendemos que no se ha de identificar necesariamente con un género impreso, ya que su evolución ha distanciado estos dos géneros. Además, la distribución textual



no se circunscribe a un entorno geográfico específico debido al carácter globalizador de Internet, sino que los posibles receptores pueden encontrarse en un contexto sociocultural distinto y, consecuentemente, puede que no posean elementos valorativos sobre la empresa (McAlpine 2001: 118). A pesar de que este bloque se suele caracterizar por un bajo porcentaje de lectura, la importancia de establecer la credibilidad del emisor en este medio hace imperativa su aparición. Esta necesidad se confirma por su carácter altamente recurrente: 75% de frecuencia en textos originales y 86,31% en localizados.

El objetivo de este bloque se dirige a consolidar la reputación y credibilidad del emisor del texto: la **función** de este bloque es expositiva en descripción, con un foco secundario exhortativo. El foco principal sería la exposición que se formula en ciertas secciones como **localización** (B1), **historia** (B5), **instalaciones** (B6), **divisiones** (B13), mientras que ciertas secciones como **experiencia** (B2) o **calidad** (B4) pretenden potenciar la imagen comercial de la empresa y que el usuario confíe tanto en la veracidad del texto como en la empresa en sí. Adicionalmente, existe cierta variación y las secciones **experiencia** (B2) o **misión** (B3) poseen un foco contextual exhortativo con un foco secundario expositivo. Este hecho se debe el propósito tras su presencia se centra en la publicidad de la empresa y en mostrar su superioridad frente a la competencia (Bly 2002: 8), con un lenguaje claramente publicitario, una información más breve y una apelación directa a la reacción positiva del receptor. En este sentido, las guías de estilo recomiendan no abusar del estilo publicitario en este bloque en particular: el usuario se dirige a las secciones **experiencia**, **misión** o **historia** en caso de necesitar información de fondo con objetivos valorativos, como una situación de posible inversión (Nielsen y Loranger 2006).

La siguiente tabla contrastiva recoge las distintas secciones posibles y sus diversas frecuencias de uso, que podemos identificar con los posibles contenidos

proposicionales presentes en este bloque. Las secciones más significativas en ambas secciones del corpus son **localización** (C1), **experiencia** (C2), y **misión** (C3). Contrastivamente, las mayores diferencias en la frecuencia de aparición se recogen en la sección **misión** (C3), un hecho que apuntaría a la mayor importancia de atribuir a la empresa ciertos valores subjetivos compartidos por la comunidad de discurso estadounidense. Igualmente, hemos identificado una diferencia significativa en las secciones **localización** (C1) e **historia** (C5). La primera responde a la necesidad de situar a una empresa multinacional en el contexto español, y la segunda tendría como función presentarse como una empresa de prestigio frente a la competencia y, en este sentido, su función podría ser similar a la sección de **experiencia** (C2).

<b>C. Información sobre la empresa.</b>	<b>Corpus Original 75%</b>	<b>Corpus Localizado 86,31%</b>
1) Localización	28,48%	57,89%
2) Experiencia de la empresa	17,44%	30,52%
3) Misión	10,46	44,21%
4) Sistema de calidad	9,88%	11,57%
5) Historia	8,13%	33,25%
6) Instalaciones	7,55%	4,21%
7) Logística	6,49%	6,31%
8) Medio Ambiente	6,39%	13,68%
9) Proyectos	4,65%	0%
10) Galerías de Imágenes	4,65%	7,36%
11) I+D	4,06%	24,21%
12) Alianzas	4,06%	5,26%
13) Seguridad	4,06%	7,36%
14) Divisiones	1,16%	3,15%
15) Exportación	0,58%	1,05%
16) Filantropía-fundación	0,58%	6,31%

**Tabla 13. Análisis contrastivo de secciones comunicativas en el bloque empresa (C).**

Es asimismo significativa la diferencia en la frecuencia de la sección **I+D** (C10), que a nuestro parecer responde a la mayor envergadura económica de las empresas del corpus localizado. Finalmente, la mayor presencia de la sección **filantropía** (C15) en el corpus localizado, con una frecuencia 5,73% superior, se corresponde a la

necesidad de aportar elementos valorativos positivos que mejoren la imagen corporativa.

Las dos únicas secciones de mayor frecuencia en el corpus original son **instalaciones** (C6) y **proyectos** (C8). En el contexto de empresas de menor importancia económica, las instalaciones que ocupen o los proyectos encomendados se pueden valorar como aspectos que proporcionarían un mayor prestigio para la empresa y, en consecuencia, suponen un hecho diferencial frente a la competencia, uno de los objetivos principales de este bloque.

Tras la revisión de las diferencias superestructurales, pasamos a analizar contrastivamente la traducción de los títulos dentro de cada bloque.

#### **9.1.2.3.1. Análisis contrastivo de títulos en el bloque «empresa»**

Si comenzamos con los textos originales, nuestro análisis revela la primacía de la palabra «Empresa» (37,2%) como la unidad léxica más convencionalizada para denominar este bloque. Ésta puede aparecer con cierta variación mediante la inclusión de determinantes o pronombres posesivos: «La empresa» y «Nuestra empresa», o el sinónimo «Compañía», que incluimos como variante en el anterior pero que sólo supone el 3,8% del total. En segundo lugar aparece, «Quiénes somos» (22,48%), que representa una presentación del bloque desde el punto de vista del usuario. En tercer lugar aparece «presentación» (8,52%), y en cuarta y quinta posición se encuentra las unidades léxicas de mayor interés desde un punto de vista contrastivo, «Sobre [nombre de la empresa]» (6,97%) y «Acerca de...» (3,87%). Si sumamos la frecuencia de estas dos preposiciones sinonímicas su frecuencia ascendería a 10,85% del total. Contrastivamente, en los textos originales estadounidenses se encuentra convencionalizada la unidad léxica compuesta que

incluye la traducción de esta preposición: *About.../About us* (71%) (Nielsen y Tahir 2002).

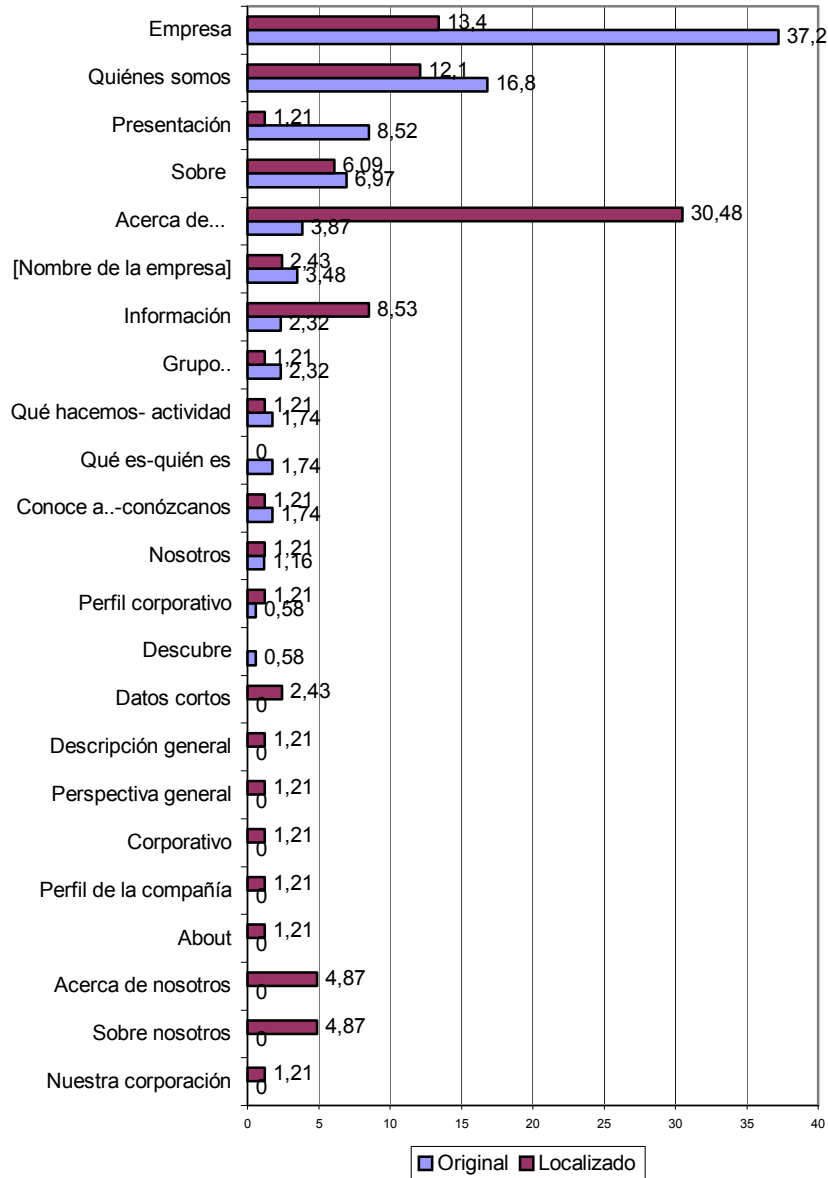
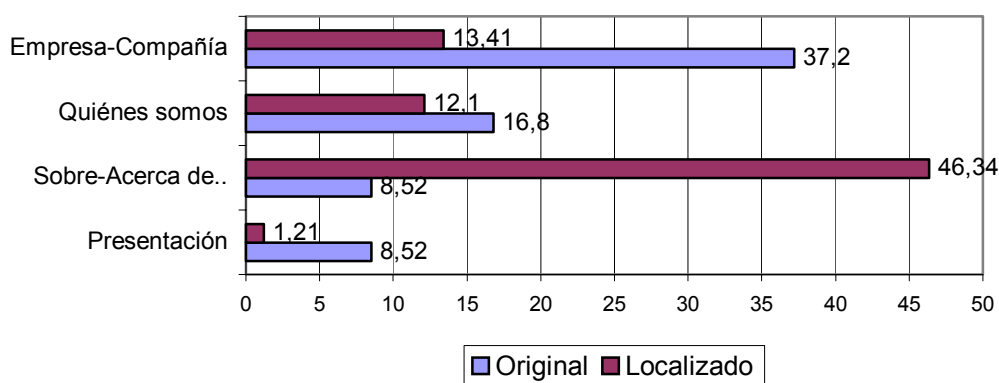


Figura 60. Análisis contrastivo de terminología en bloque empresa (C).

El análisis contrastivo en nuestro corpus revela la primacía de la unidad léxica compuesta por la preposición más sustantivo «Acerca de [nombre de la empresa]» (30,48%), en la que se refleja la unidad léxica más convencionalizada en los textos originales estadounidenses. Esta divergencia en la frecuencia de uso es más

pronunciada al incluir la relación de sinonimia entre las preposiciones «sobre» / «acerca», y al añadir las unidades léxicas que incorporan estas preposiciones: «sobre [nombre de la empresa]», «sobre nosotros», «acerca de nosotros» y «about us». La suma total asciende al 47,51% de frecuencia, mientras que en el original supone el 10,58%.

Es igualmente significativa la preferencia de la forma preposicional «acerca» frente a «sobre» en los textos localizados, que se invierte en los originales al ser mayoritario el uso de «sobre», con una frecuencia del 6,97% frente al 3,89% de uso de «acerca». Asimismo, nos hacemos eco de la alta frecuencia de uso exclusivo en los textos localizados de «sobre nosotros» y «acerca de nosotros» (4,87%), que consideramos anglicismos léxicos al ser inexistentes en el corpus original. La frecuencia combinada de las formas inexistentes en el corpus original asciende a 19,43%.



**Figura 61. Análisis contrastivo de la terminología más frecuente en bloque empresa (C) según sinónimos.**

Estos anglicismos no serían los únicos identificados en este caso, puesto que las unidades léxicas «datos cortos»<sup>379</sup>, «corporativo» y «nuestra corporación» no aparecen tampoco en el corpus original y constituyen anglicismos léxicos.

<sup>379</sup>Esta traducción supone un calco de la unidad léxica *quick facts*.

**(1) Sección «localización» (C1)**

Esta sección presenta la localización física de la empresa al usuario. En el corpus original los términos más frecuentes son «localización» y «oficinas» (20,4%), seguidos de «¿dónde estamos/está?» (18,37%). Aparece igualmente el sinónimo de localización «Situación» (4,08%).

En un gran número de casos la localización se facilita mediante un mapa de ubicación, y este hecho se recoge en la constitución de las unidades léxicas empleadas. De hecho, el 10,09% incluyen implícitamente los términos «mapa» o «plano»: «Ver mapa de situación» [n. 1], «Ver plano de situación» [n. 1] / «mapa de situación» [n. 1]», «Localízanos en el plano» [n. 1].

De igual manera, en el 10,05% de formas se incluyen buscadores de oficinas o tiendas, que apuntan a la existencia de posibles instalaciones próximas al usuario: «Busca tu tienda», «Buscador de oficinas», «Localizador de tiendas», «Red de distribuidores», «Concesionarios [SEAT]», «Dónde comprar». El análisis contrastivo muestra la presencia de gran cantidad de unidades léxicas de uso exclusivo en una u otra sección del corpus, en parte por los distintos tipos de instalaciones posibles: «oficinas», «delegaciones», «distribuidores», «tiendas», «concesionarios», «partners», «locales», «centros», «plantas». Además, la variación se manifiesta a través de la forma léxica que conceptualiza el procedimiento de búsqueda: «ver plano», «dónde...», «red...», «localizador», «buscar», «buscador», «localizar», «encontrar», o «presencia». El dato más significativo sería la inexistencia en el corpus localizado del término de mayor índice de prototipicidad en el corpus original, «localización» (20,41%), aunque sí que se ha identificado «localizaciones» (1,82%), y el sinónimo «ubicación»<sup>380</sup>(5,45%), formas inexistente en el corpus original.

---

<sup>380</sup> En el corpus original aparece el sinónimo «situación» (4,08%).

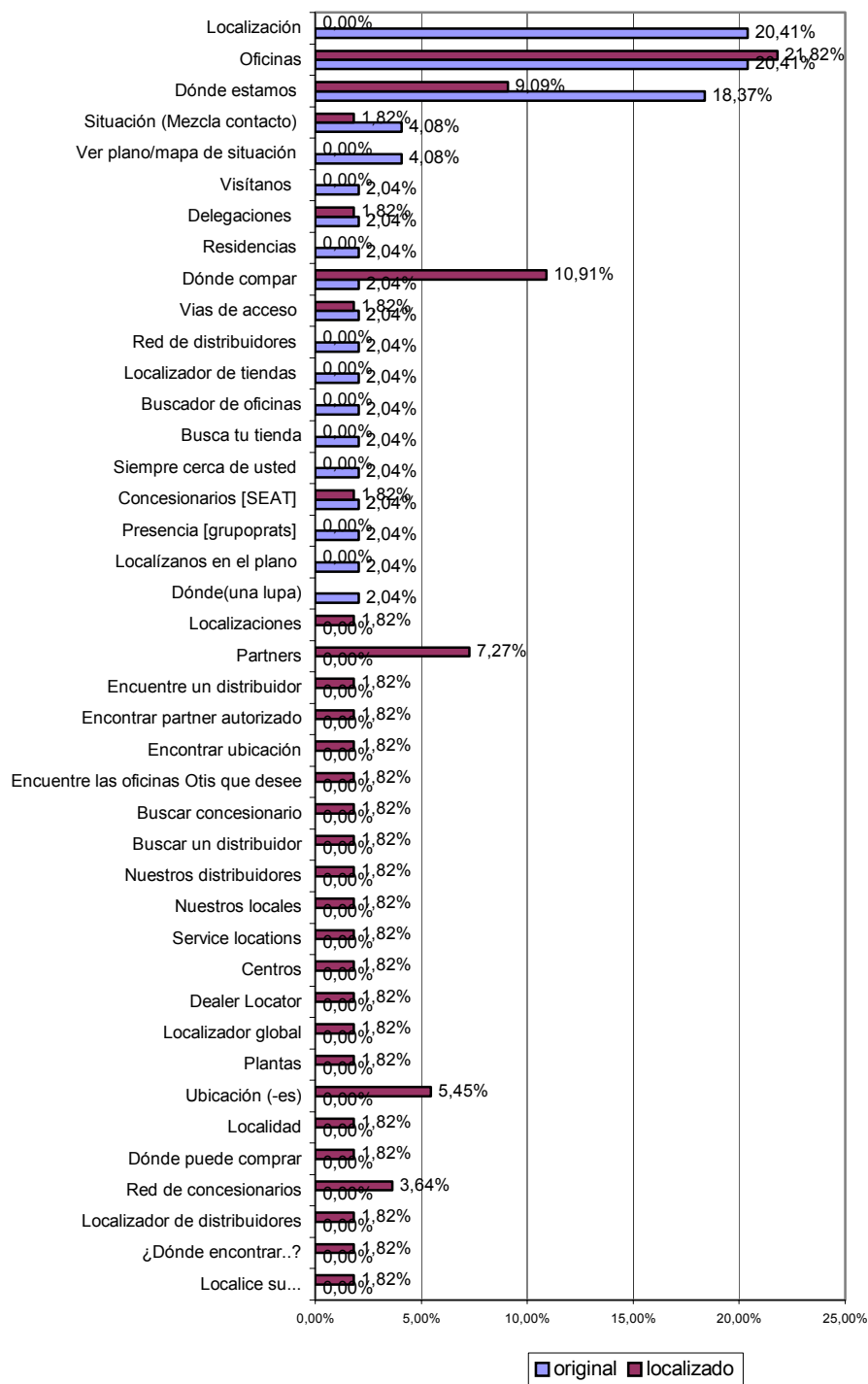


Figura 62. Análisis contrastivo de terminología en sección localización (C1).

(2) Sección «experiencia» (C2)

Durante el proceso de establecimiento de la credibilidad del emisor, es primordial mostrar la experiencia acumulada para las empresas que no gocen de un amplio

reconocimiento social. Esta sección queda recogida en el 17,44% de los sitios web originales y en el 30,52% de los localizados.

En el corpus original el término más usado es «clientes» (36,66%). Este hecho se debe a que la empresa trata de mostrar los clientes de prestigio con los que mantiene una relación comercial. En segundo lugar por frecuencia se encuentran las unidades léxicas «trabajos/servicios realizados» (30%), que apuntan igualmente a la experiencia de la empresa. El resto de unidades léxicas identificadas en el corpus original son: «trabajos realizados» [n. 3], «servicios realizados» [n. 2], «instalaciones» [n. 3], «dossier» [n. 1]. En el análisis de esta sección se observó la polisemia del término «instalaciones», que en este caso se refiere a los trabajos realizados, como por ejemplo murales o instalaciones publicitarias, ([www.mpelectronics.info](http://www.mpelectronics.info)), y en el epígrafe anterior se refiere las oficinas o presencia física de la propia empresa. En un 10% de los casos aparece directamente el término «Experiencia», que en esta sección constituye el elemento diferencial de la empresa. Este aspecto queda recogido en dos casos: «La diferencia» o «elemento diferencial».

Si nos adentramos en el análisis contrastivo, podemos constatar una considerable divergencia entre la terminología empleada. En los textos localizados se prefiere mostrar los posibles beneficios de entablar una relación comercial por medio de las experiencias previas de clientes. Éstas quedan reflejadas en la unidad léxica «casos de éxito», que consideramos un calco. En el corpus localizado el 37,07% de expresiones incluyen la palabra «casos»: «casos de éxito» (18,51%), «casos de clientes», «casos concretos», «casos prácticos», «estudio de casos», «case studies». Estas unidades léxicas que incluyen la palabra «casos» constituirían anglicismos léxicos, que hemos establecido como tal al no estar presentes en el corpus original<sup>381</sup>. En un segundo lugar, en los textos localizados aparecen los premios,

<sup>381</sup> Sólo se identificó una aparición de esta unidad léxica en el cuerpo de una página web original.



reconocimientos y galardones recibidos, que se encuentran representados en un 29,62% de expresiones. Por último aparecen «historias de éxito» y «ventajas de...» ambas con un 7,40% de uso. Durante el análisis se observó una dualidad en las secciones tituladas «premios» o «galardones» que en este caso se refieren a la experiencia de la empresa, aunque en otros casos se refiere a premios otorgados por la propia empresa, como en [www.ey.com](http://www.ey.com) [Earnest & Young]. Por su parte, los textos originales se decantan por mostrar a los clientes previos que han confiado en la empresa o sus trabajos realizados, mientras que los localizados se decantan por mostrar los posibles beneficios de esta experiencia desde el punto de vista del usuario, por medio de testimonios de clientes previos y premios/galardones recibidos<sup>382</sup>.

### (3) Sección «misión» (C3)

En esta sección se presentan los valores subjetivos que representan a la empresa o a los que aspira. Las guías de estilo recogen que los usuarios tienden a ignorar esta sección debido a su estilo exagerado y publicitario (Price y Price 2002). De nuevo, la **función** textual sería expositiva descriptiva con un foco contextual secundario exhortativo, puesto se pretende condicionar al usuario para que respete y valore a la empresa. Esta sección presenta un bajo índice de convencionalización en los textos originales (10,46%). Sin embargo, en los textos localizados esta sección presenta una mayor frecuencia (44,21%) y, de hecho, esta sección presenta uno de los mayores índices de diferencia relativa a la frecuencia entre ambas secciones del corpus (+ 33,7%, véase Figura 54).

Las unidades léxicas «Misión / nuestra misión» (29,41%) y «Filosofía / nuestra filosofía» (17,45%) son las más frecuentes en los textos originales. Asimismo, son

---

<sup>382</sup> Este recurso retórico de mostrar los aspectos positivos de la empresa desde el punto de vista de los posibles beneficios al usuario se recomienda en la totalidad de las guías de estilo anglosajonas (Price y Price 2002; Jeney 2007; etc.).

varias las palabras empleadas para denominar el concepto al que se alude en esta sección, que presentamos en orden descendente de uso: «misión», «filosofía», «enfoque», «visión», «compromiso», «responsabilidad», «política» o «trabajamos para ti». La **función** pragmática de esta sección, en la que es esencial la inclusión de elementos de valoración subjetivos y la apelación a la empatía del usuario, hace que aparezca el pronombre «Nuestra-o...» en el 44,4% de unidades léxicas del corpus original. Este constituye el mayor índice de mención implícita del emisor en la totalidad de unidades léxicas del corpus original. Sorprendentemente, esta forma pronominal solo aparece en el 13,63% de las unidades léxicas localizadas, en claro contraste con la presencia mayoritaria de las formas pronominales de primera persona del plural en las unidades léxicas de otros bloques como **contacto** (B).

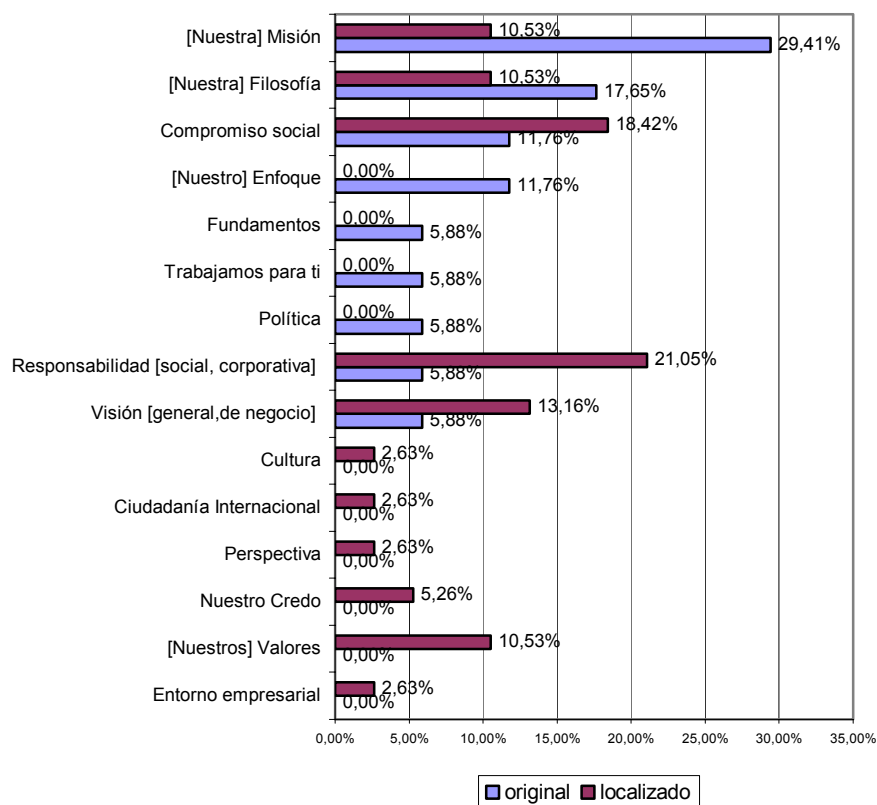


Figura 63. Análisis contrastivo de terminología en sección misión (C3).

El análisis contrastivo deja patente la preferencia mayoritaria del uso de la palabra «misión» en los textos originales y de «responsabilidad» en los localizados, aunque en ambos aparece en menor medida «visión» y «filosofía». No obstante, las unidades léxicas localizadas presentan ciertas formas sustantivas sin equivalentes en el corpus original como «valores», «credo», «perspectiva» y «cultura». En particular, «credo» constituye un error de traducción por su exclusiva pertenencia al dominio de la religión según el diccionario de la RAE<sup>383</sup>.

#### **(4) Sección «calidad» (C4)**

La convencionalización de esta sección responde al imperativo de presentar elementos diferenciales y de prestigio de la empresa. Esta sección aparece en el 9,30% de los sitios web originales y 11,57% de los localizados. En los primeros el término «calidad» aparece directamente en el 57,25% de las unidades léxicas de esta sección, con la posibilidad de presentar un compuesto sintagmático con sustantivos como «compromiso de...» o «certificados de...»<sup>384</sup>. Dado que la calidad se puede certificar por medio de normas internacionales, en el 47,75% de las unidades léxicas aparece la palabra «certificación» o «acreditación». Asimismo, hemos identificado un enlace denominado directamente con el término relativo a las normas de calidad, las normas internacionales ISO, específicamente la norma ISO 9001: 2000, y la norma ISO 14001: 2004 [Diode.es]<sup>385</sup>.

El análisis de los textos localizados revela que la palabra «calidad» aparece en el 72,2% de las unidades léxicas y, por tanto, al igual que en la sección original sería

---

<sup>383</sup> La definición de credo del diccionario de la RAE es: 1. m. Oración en la que se contienen los principales artículos de la fe enseñada por los apóstoles. 2. m. Conjunto de doctrinas comunes a una colectividad.

<sup>384</sup> Los textos web pueden mostrar un subtexto sobre «certificados» referentes a los certificados de seguridad en línea que protegen la información que se proporciona en los sitios web como los ofrecidos por [www.verisign.com](http://www.verisign.com). Los certificados se referirían a la seguridad del sitio web.

<sup>385</sup> En los sitios web localizados se contabilizaron dos páginas web independientes referidas a las normas ISO 9001: 2000 en [www.lexmark.com](http://www.lexmark.com) y [www.fedex.com](http://www.fedex.com) aunque los enlaces se encuentran dentro del cuerpo de la página web y no como parte de los menús de navegación, y por tanto, no fueron contabilizados según nuestra metodología de análisis.

la forma más frecuente. Hemos identificado igualmente «certificados» y «certificación-es», con lo que podemos concluir que no existen diferencias significativas en esta sección entre ambas secciones del corpus.

### (5) Sección «instalaciones» (C6)

La sección que presentan las instalaciones que ocupa la empresa aparece en el 7,55% de los sitios web originales y en el 4,21% de localizados. El término más utilizado en los textos originales es directamente «instalaciones» (61,53%). Otros términos usados con una aparición cada uno son «fábricas», «establecimientos», «espacio», «laboratorios».

En esta sección se podrían incluir varias de las apariciones contabilizadas de la unidad léxica «oficinas» que incluimos en la sección **localización** (C1) puesto que, aparte de indicar la localización de la empresa, en ciertos casos se describen las oficinas en sí. Estas dos secciones se pueden presentar combinadas en una misma página. En ciertos casos nuestra asignación a una u otra sección no se puede realizar con criterios totalmente nítidos, ya que se pueden combinar varias secciones en una misma página. No obstante, como indicamos en nuestro marco metodológico, procedimos al análisis de su función comunicativa principal o dominante para asignarlas a una sección o bloque determinado. Un caso que ejemplifica la necesidad del contexto para la desambiguación de la terminología empleada es la aparición de «establecimientos» en el sitio web de [www.americanexpress.es](http://www.americanexpress.es). En este sitio web constituye un área reservada para establecimientos que acepten esta tarjeta de crédito en lugar describir los establecimientos de la propia empresa, de ahí que la asignáramos a la sección **área de clientes** (G).

#### 9.1.2.4. Análisis del bloque «noticias» (D)

En la actualidad los sitios web corporativos se han convertido en el repositorio de las noticias que se desarrollan en el seno de las empresas, tanto en su estructura como en los productos y servicios que ofrecen. Como tal, éstos constituyen la primera fuente de información sobre las empresas para el público general. Los sitios web ofrecen una plataforma de comunicación efectiva que permite difundir noticias con gran rapidez o responder a situaciones concretas por medio de comunicados de prensa, de ahí que se hayan constituido en el referente básico de información sobre cada empresa (Price y Price 2002). Dado que este bloque suele actualizarse regularmente, el sitio web puede potencialmente ofrecer a los usuarios nuevos contenidos que promueven las visitas recurrentes al texto web (Nielsen y Loranger 2006). La noción de inmediatez a la que se asocia este bloque se manifiesta en la aparición de unidades léxicas con compuestos adjetivales que denotan actualidad como «últimas noticias», «última hora» o «noticias recientes».

La **función** textual en este bloque es expositiva, tanto en la descripción como en narración. Dependiendo del tipo de noticia, puede aparecer un foco contextual secundario exhortativo. Por ejemplo, en ciertos casos se han encontrado en esta sección folletos informativos publicitarios o artículos comerciales que poseen un foco secundario exhortativo (Gamero 2001: 81).

Este bloque y sus secciones son de carácter **múltiple**, ya que cada noticia se suele presentar en una página web diferente. No obstante, a pesar de su adaptación a un estilo más apropiado para la web (Price y Price 2002) las noticias incluidas en este bloque suelen constituir *e-textos* (véase 3.3.1). Por consiguiente, su análisis contrastivo desde un punto de vista intratextual queda fuera de los objetivos de la presente tesis al no constituir formas textuales diferenciadas por el medio<sup>386</sup>.

---

<sup>386</sup> Consultar Bustos *et al.* (2005) para un monográfico sobre la traducción periodística.

La **funcionalidad** que presenta este bloque comunicativo se articula en torno a la interactividad del usuario con el texto, tanto en la activación de vínculos para obtener noticias como en su lectura selectiva, en la búsqueda de noticias por medio de la fecha de emisión o por palabras clave y, por último, en la suscripción a boletines o noticias sindicadas RSS.

<b>D. Noticias/Actualidad</b>	<b>Corpus original 54,05%</b>	<b>Corpus localizado 79,94%</b>
1) Novedades	13%	9,47%
2) Eventos	6,98%	25,26%
3) Notas de prensa	6,98%	10,52%
4) Suscripción noticias	2,90%	5,26%
5) RSS	1,16%	6,31%

**Tabla 14. Análisis contrastivo de secciones comunicativas en el bloque noticias (D).**

En análisis superestructural contrastivo revela una notable diferencia en la frecuencia de aparición de este bloque: 54,05% en el corpus original y 79,92% en el localizado. En parte atribuimos esta divergencia a la mayor envergadura de las empresas de la sección localizada. Si contrastamos estas frecuencias con las presentadas en una investigación previa (González y Ruiz de Valbuena 2006: 279), ésta última sitúa en el 92% la presencia de esta sección en los sitios web originales de grandes empresas. Los resultados que hemos obtenido son inferiores (79,94%), y esta diferencia del 12,08% entre ambas investigaciones podría apuntar a una menor presencia de este bloque en los textos localizados por la frecuente actualizaciones de este bloque y su necesaria localización posterior<sup>387</sup>.

En el análisis del corpus hemos identificado la presencia de cinco secciones posibles junto con las noticias tradicionales:

<sup>387</sup> De hecho, es común la inclusión de noticias en inglés con la indicación expresa del idioma en el que se expresa, tal como se observa en la Figura 40.

- Las **novedades** (D1) suelen referirse a noticias relacionadas exclusivamente con los productos o servicios ofrecidos. Estas noticias sobre los productos o servicios se pueden encontrar asimismo dentro de noticias generales. El porcentaje de aparición en el corpus localizado (13%) es superior que en el corpus original (9,47%).
- En segundo lugar se encuentran los actos organizados por la empresa en la sección **eventos** (D2). La descripción de eventos se produce tanto para las actividades ya realizadas, como para las que se producirán en el futuro. Contrastivamente, la frecuencia de esta sección es bastante superior en los textos localizados, 25,26% frente a 6,98%. En su formulación hemos identificado una relación metafórica con la inclusión de las unidades léxicas «agendas de eventos o actividades» o «calendarios de eventos». Es común su incorporación dentro de un bloque general de noticias y, de hecho, la unidad léxica combinada «noticias y eventos» aparece un 3,48% de los sitios originales. Por último, el análisis de las unidades léxicas revelan que los eventos pueden ser entre otros: «ferias», «congresos» o «actos».
- El tercer lugar lo ocupan las **notas de prensa** (D3), que constituyen una comunicación oficial en respuesta a una situación específica y cuyos receptores son principalmente periodistas (Price y Price 2002). Una nota o comunicado de prensa queda definido por la RAE (XXII edición) como una «nota, declaración o parte que se comunica para conocimiento público». Para ilustrar este punto presentamos un fragmento en el que queda patente este carácter específico de las notas de prensa, cuyo emisor como entidad única es la empresa que responde oficialmente a una situación determinada: «Desde esta sección podrá consultar las últimas Notas de Prensa de Adecco **emitidas por las empresas** del Grupo» [Adecco.es].

- En cuarto lugar se encuentra la suscripción a boletines<sup>388</sup>, una acción que ha de realizarse activamente por parte del usuario. Estos boletines se reciben a través del correo electrónico y, por tanto, la página en la que se formula esta sección constituye un mero formulario de inscripción. Adicionalmente, es común la presencia de un archivo ordenado por fecha de emisión. El ofrecimiento de boletines se recomienda por constituir un mecanismo que vuelve a atraer al usuario al texto y que contribuye a su mayor uso (Nielsen y Loranger 2006: 54). Este carácter de inmediatez asociado con los boletines se refleja en la palabra «avisos», identificado como alternativa para denominar esta sección.




Figura 64. Ejemplo de enlace a sección D5, RSS. De [www.gruporh.com](http://www.gruporh.com).

- En último lugar, y como nuevo desarrollo en la evolución constante de la WWW, aparecen las noticias sindicadas o RSS, *Really Simple Syndication*<sup>389</sup> (Holzner 2006). Este formato de suscripción a noticias surgió en 1999, pero no fue hasta el 2005 cuando se extendió su uso. Las noticias RSS se encuentran en formato XML, requieren normalmente de un lector específico y sus mayores ventajas son (1) la posibilidad de personalizar la información que se recibe y (2) su descarga a PDA o teléfonos móviles (ibid). La frecuencia de esta sección en sitios corporativos atestigua su creciente uso, un 6,31% de frecuencia en sitios

<sup>388</sup> Los boletines se definen como «news bulletins produced for specific readers, such as employees [...] or interest groups. Usually these are linked and archived on the company or organization Web site. Newsletters tend to be brief (four to eight pages), are produced in-house, and ordinarily do not sell advertising space on their pages. Some of these are even distributed via e-mail rather than on the Web» (Jeney 2007: 146).

<sup>389</sup> No obstante, han existido diferentes formatos XML para los RSS. Estas siglas responden dependiendo de la versión a *Rich Text Summary*, versión 0.99, o en la versión 2 de Microsoft *Real Simple Synchronization* (Holtzner 2006).



web localizados y un 1,16% en los originales<sup>390</sup>. Su identificación se ha convencionalizado por medio de un icono normalizado, , y aunque en la actualidad la versión 2.0 es la más extendida, se suelen incluir en otros formatos como se observa en la imagen del sitio web [www.gruporh.com](http://www.gruporh.com):

#### 9.1.2.4.1. Análisis contrastivo de títulos en bloque «noticias»

El análisis de la terminología en la denominación de este bloque muestra una relativa similitud en la frecuencia de uso entre ambos componentes del corpus. En ambas secciones la unidad léxica más frecuente para este título es «noticias» (40,63% en original y 35,11% en localizado). La diferencia más significativa se encuentra en la mayor frecuencia de la unidad léxica «sala de prensa» en el corpus localizado (17,02% frente a 9,38%), que puede responder a la traducción del término «press room» anglosajón. Es igualmente significativa la mayor frecuencia en el corpus localizado del sustantivo «prensa» y del compuesto adjetival «últimas noticias». En la Figura 65 hemos incluido las unidades léxicas del bloque **noticias** (D) y las secciones **novedades** (D1) y **notas de prensa** (D3).

La mayor presencia de la unidad léxica «sala de prensa» en los sitios localizados puede reflejar la unidad léxica original «press room», mayoritaria en los sitios web anglosajones según la investigación de González y Ruiz de Valbuena (2006)<sup>391</sup>. Asimismo, es de interés la relación metafórica entre este bloque y las redacciones de periódicos, que se pone de manifiesto en el uso de compuestos nominales cuyo primer elemento de sustantivos es «sala» [n. 9 original, n. 16 localizado], «centro» [n. 2, n. 2], «gabinete» [n. 2, n. 1], «información» [n. 0, n. 2], y con una segunda

---

<sup>390</sup> González y Ruiz de Valbuena (2006: 273) sitúan en un 8,3% el porcentaje de inclusión de noticias en formato RSS en los sitios web corporativos españoles.

<sup>391</sup> Esta investigación se centra en la presencia de secciones de noticias en sitios corporativos y presenta como posibles unidades léxicas para denominar esta sección *press room*, *press center*, *media center*, *media* y *press*. (González y Ruiz de Valbuena 2006: 270).

forma sustantiva entre «prensa» [n. 16, n. 29], «información» [n. 1, n. 0], «comunicación» [n. 2, n. 1], «noticias» [n. 0, n. 1] y «medios» [n. 0, n. 1].

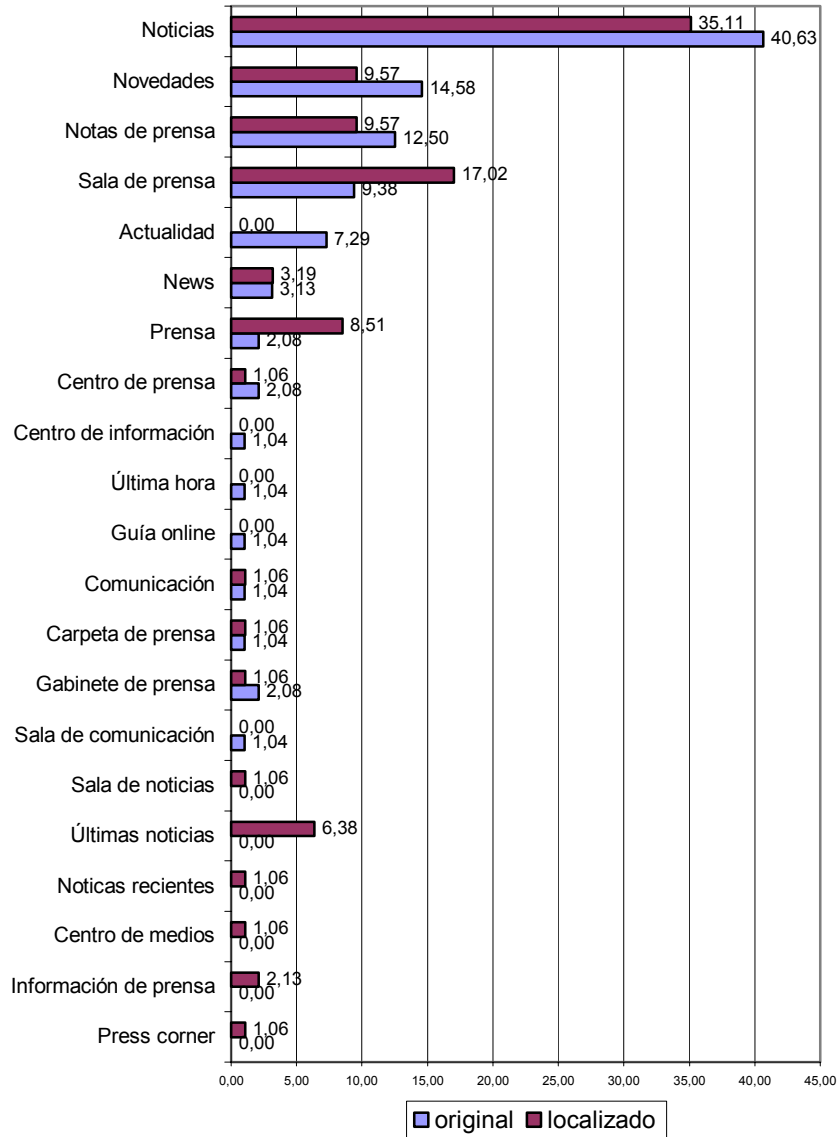


Figura 65. Análisis contrastivo de terminología en bloque «noticias» (D1).

Ciertas unidades léxicas localizadas no se encuentran presentes en el corpus original: «sala de noticias», «centro de medios», «información de prensa», «press corner», «últimas noticias», «noticias recientes». Las cuatro primeras combinaciones no quedan recogidas en el corpus CREA, un hecho que potencialmente podría apuntar a su carácter de calcos léxicos. Como ejemplo, «sala de noticias» se corresponde con *newsroom* en inglés, cuya traducción óptima

podría ser «sala de prensa-redacción». Los textos originales presentan además dos anglicismos puros en este bloque, «news», (3,13%) y «newsletter» (0,58%). La frecuencia de estos préstamos en los textos localizados sería similar, con lo que en este caso no podríamos atribuir su presencia en los textos localizados exclusivamente al proceso de traducción.

### 9.1.2.5. Análisis del bloque «productos/servicios» (E)

Las empresas basan su actividad económica en la venta de productos o servicios, y por consiguiente, los usuarios acceden al sitio web fundamentalmente para obtener información sobre los mismos: tanto en la fase de pre-venta, en su compra o contratación, como en pos-venta. Ciertas investigaciones sitúan como objetivo fundamental de los sitios web corporativos el apoyo a sus actividades comerciales en estas tres fases (Liu *et al.* 1997). En sus orígenes, las empresas emisoras del texto utilizaban su sitio web como plataforma adicional de presentación y publicidad de sus productos o servicios<sup>392</sup>, aunque dada la evolución del Internet y las ventajas de su funcionalidad, podemos afirmar que los sitios web se han constituido en la plataforma primaria de información sobre productos o servicios.

	Corpus original	Corpus localizado
<b>E. Productos -Servicios-Soluciones</b>		
1) Productos	53,48%	84,21%
2) Servicios	44,76%	68,42%
3) Ofertas	9,88%	14,73%
4) Información técnica	5,81%	5,26%

**Tabla 15. Análisis contrastivo de secciones comunicativas en el bloque productos-servicios (E).**

La **función** de este bloque es expositiva descriptiva con un foco contextual secundario exhortativo. El objetivo del bloque se centra en la descripción de aquellos productos y servicios que ofrece la empresa, realizada a través de una

<sup>392</sup> Para ciertas empresas sin presencia física el sitio web es su única plataforma de presentación de productos o servicios, como [www.amazon.es](http://www.amazon.es).

presentación atractiva dirigida a la reacción positiva del usuario. En ciertos casos se presentan descripciones muy breves con una gran carga publicitaria, y en éstos la función es exhortativa con un foco contextual secundario expositivo, ya que supondrían una «descripción publicitaria informativa». De hecho, Trosborg (1997: 14) propone este caso como ejemplo de la posible multifuncionalidad de los textos (véase 5.5.2), puesto que la publicidad de productos puede ser ante todo persuasiva y, por lo tanto, con una función primaria exhortativa o esencialmente descriptiva, sobre todo en descripciones más técnicas o especializadas de los productos.

Corpus original			Corpus localizado		
Lista lematizada de palabras por frecuencia global					
1	Ser	0.60	1	Ser	0.59
2	Com	0.34	<b>2</b>	<b>Productos</b>	<b>0.55</b>
<b>3</b>	<b>Servicios</b>	<b>0.34</b>	3	Nuevos	0.51
<b>4</b>	<b>Productos</b>	<b>0.30</b>	4	Información	0.47
5	Información	0.26	<b>5</b>	<b>Servicios</b>	<b>0.44</b>
6	Empresas	0.24	6	Vehículos	0.42
7	Precio	0.21	7	Ford	0.34
8	Euros	0.20	8	Todos	0.32
9	Nuestros	0.20	9	Hp	0.32
10	Tv	0.19	10	Poder	0.27

Tabla 16. Palabras de mayor frecuencia global en ambos corpus. Estadística realizada con Wordlist.

Este bloque es de carácter **múltiple**, es decir, que puede plasmarse en gran cantidad de páginas dependiendo del número de productos o servicios ofrecidos por la empresa. Este carácter múltiple y la importancia de este bloque en el género quedan patentes en el análisis intralingüístico de la lista global de palabras lematizada<sup>393</sup> del corpus. En ambas secciones del corpus las palabras «productos» y «servicios» aparecen dentro de las primeras posiciones: en el corpus original la palabra «servicios» ocupa el tercer lugar y «productos» el cuarto, mientras que en el corpus localizado ocupan la quinta y la segunda respectivamente.

<sup>393</sup> Realizada con la herramienta Wordlist en el conjunto global del corpus.

La **funcionalidad** que presenta este bloque comunicativo se fundamenta en la interacción que proporciona el medio, búsquedas de productos en catálogo, gráficos interactivos o de diferentes tamaños, videos publicitarios de los productos, así como la posibilidad de configurar ciertos productos con la presentación final del modelo elegido y/o el precio total.

Aparte de la adaptación a la presentación en pantalla y la funcionalidad que proporciona el medio, la redacción de los textos en este bloque, y consecuentemente su traducción, se asemeja a la redacción y la traducción de textos técnicos:

The principles of writing **product copy** for the Web are pretty similar to writing print brochures, catalogs, and sell sheets, with a few exceptions based on the Web user's shorter attention span and the interactivity of the medium<sup>394</sup> (Bly 2002: 172).

Asumimos así que este tipo de textos, al igual que los del bloque **noticias** (D), está constituido principalmente por *E-textos*. Consecuentemente, pensamos que su descripción y análisis contrastivo se ha analizado ampliamente en otras investigaciones sobre traducción técnica (Göpferich 1995; Gamero 1998, 2001; Jiménez Serrano 2002; Byrne 2006), de ahí que no profundicemos en sus características intratextuales más allá del análisis contrastivo sobre las formas lingüísticas convencionales en los menús de navegación.

#### **9.1.2.5.1. Análisis contrastivo de títulos en el bloque «Productos-Servicios»**

Procederemos en primer lugar a analizar la terminología de la sección **productos** (E1). En ambas secciones del corpus es mayoritario el uso del término «productos», 68,47% en el original y 56,25% en el localizado. El aspecto de mayor interés lo constituiría la mayor inclusión expresa del nombre del producto para denominar esta sección en el corpus localizado, 35% frente a 1,08% en el original. Esta

---

<sup>394</sup> Énfasis propio.

incorporación del nombre del producto se correspondería con un mayor carácter descriptivo explícito de estas unidades léxicas en los textos localizados<sup>395</sup>. Identificamos asimismo una mayor frecuencia del término «catálogo» dentro de los textos localizados. De entre las unidades léxicas presentes en el corpus original y sin presencia en el localizado, resaltamos la pregunta retórica «¿qué hacemos-ofrecemos-realizamos?» (16,8%), que se corresponde con «¿quiénes somos?» en el bloque **empresa** (C).

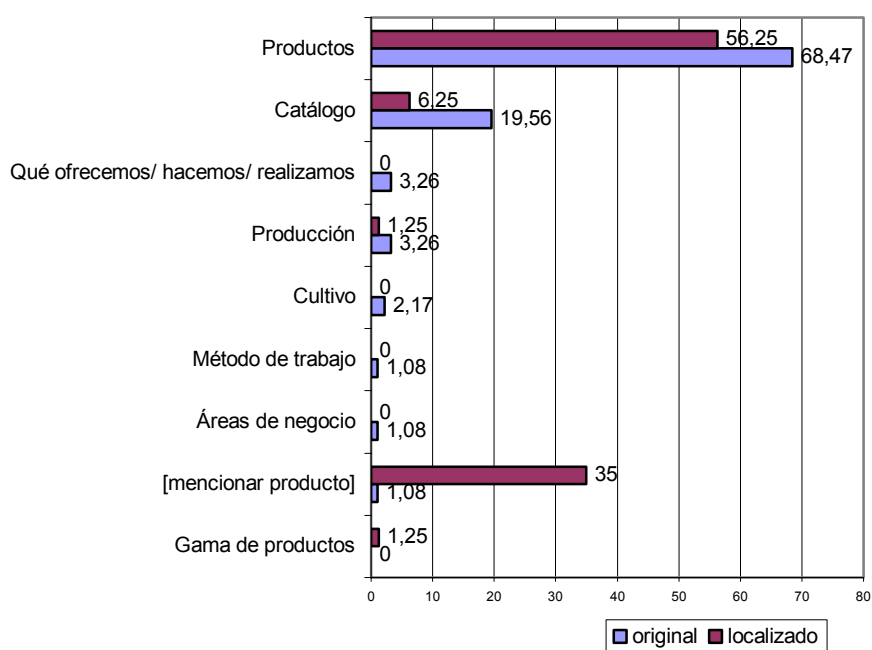
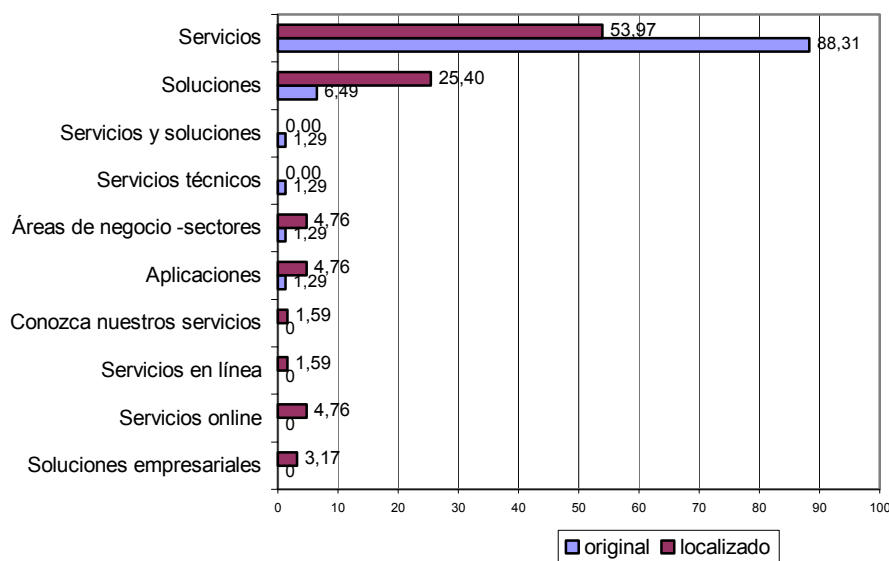


Figura 66. Análisis contrastivo de terminología en sección productos (E1).

En la siguiente sección (E2), el término de mayor frecuencia en el corpus original lo constituye «servicios» (88,1% y 53,97% en localizado). En la sección localizada del corpus los servicios ofrecidos se describen en mayor medida como «soluciones» o «aplicaciones», unos términos con una frecuencia significativamente mayor. Esta mayor frecuencia apuntaría al carácter de préstamo léxico de ambos.

<sup>395</sup> Este aspecto estaría en consonancia con las recomendaciones de las guías de estilo web (Bly 2002).



**Figura 67. Análisis contrastivo de terminología en sección servicios (E2).**

En el corpus localizado prima igualmente la forma «servicios online» (4,76%) frente a «servicios en línea» (1,59%), ambas sin equivalentes en el corpus original. En estos dos compuestos sintagmáticos queda patente la preferencia del anglicismo puro «online» sobre su traducción al español «en línea», aunque éste constituye un anglicismo adaptado<sup>396</sup>.

Por último, la sección E3 se corresponde con las posibles ofertas y promociones de la empresa. El término de mayor frecuencia en el corpus original sería directamente «oferta» (70,58%) en el corpus original, aunque su uso es menor en el corpus localizado (28,57%). En éste último el término de mayor frecuencia es «promoción» (50%), que constituiría un anglicismo léxico puesto que en español esta palabra se asocia con las campañas publicitarias según el diccionario de la RAE (XXI edición)<sup>397</sup>, y no exclusivamente con ofertas como tales.

<sup>396</sup> Se ha propuesto «interactivo» como traducción más apropiada en este caso (Belda 2003).

<sup>397</sup> La definición que ofrece el diccionario de la RAE (edición XXII) de promoción es «4. f. Conjunto de actividades cuyo objetivo es dar a conocer algo o incrementar sus ventas».

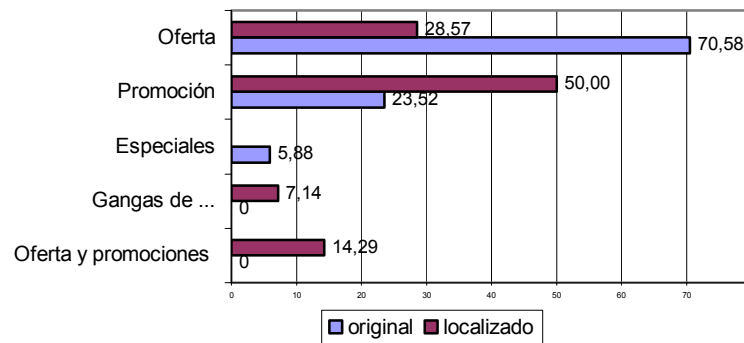


Figura 68. Análisis contrastivo de terminología en sección ofertas (E3).

Entre los términos localizados aparece «gangas de...», que no aparece en los textos originales debido a su carácter coloquial. Se identificó asimismo la polisemia del término «oferta», que en el contexto empresarial se puede referir además a la oferta de prestación de servicios<sup>398</sup>, reflejada en las «peticiones de oferta» [Arrow Ibérica].

### 9.1.2.6. Análisis del bloque «legal» (F)

La actividad de las empresas emisoras del texto se encuentra sometida a la legislación vigente en España, el contexto sociocultural al que se dirigen los textos del corpus. Este aspecto de la situación comunicativa y sociocultural incide en la aparición de este bloque comunicativo que recoge tanto el estatus y las responsabilidades jurídicas de la empresa emisora como el conjunto de derechos y deberes de los usuarios. Podemos sintetizar sus objetivos principales en: (1) establecer el estatus del emisor como entidad jurídica, y (2) garantizar que la información suministrada directa o indirectamente<sup>399</sup> permanecerá confidencial (Liu y Arnet 2000; Jeney 2007; Price y Price 2002).

<sup>398</sup> Las peticiones de oferta suelen representarse interactivamente por medio de un formulario.

<sup>399</sup> Liu y Arnett (2000: 17) dividen este aspecto entre información personal y no personal. La información no personal se refiere a las páginas visitadas, tiempo de visita, dominio desde el que se accede, navegador usado, sistema operativo empleado, la dirección IP del usuario con la posibilidad de instalar una *cookie* en el ordenador del usuario.



En general, este bloque comunicativo debería ser redactado o revisado por expertos legales (Garrand 2001), de ahí que pueda presentar un estilo jurídico distintivo que puede dificultar su comprensión (Bly 2002). No obstante, a pesar de ser un bloque obligatorio por la garantía que ofrece (McAlpine 2001: 120), los usuarios raramente leen estos subtextos por su complejidad.

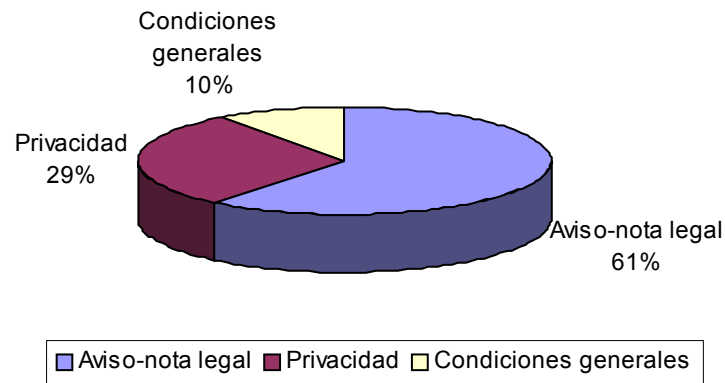
La **función** textual es exhortativa sin alternativa, ya que son jurídicamente vinculantes. De hecho, como veremos en el análisis detallado de esta sección (véase 10.3.5.2), el mero acceso y uso del sitio web implica la aceptación tácita del usuario de las provisiones estipuladas en esta sección.

Este bloque comunicativo se compone de las siguientes secciones constitutivas: **avisos legales** (F1), **declaraciones de privacidad** (F2), **condiciones generales** (F3), **derechos de propiedad** (F4), y **marcas registradas** (F5).

	Corpus original	Corpus localizado
<b>F. Legal</b>		
1) Aviso-Nota Legal	27,90%	47,36%
2) Declaración de privacidad	13,37%	70,52%
3) Condiciones Generales	4,65%	38,94%
4) Derechos de propiedad	2,90%	17,89%
5) Marcas registradas	0%	11,57%

**Tabla 17. Análisis contrastivo de secciones comunicativas en el bloque legal (F).**

Estas secciones no aparecen divididas nítidamente, sino que los contenidos jurídicos de ciertas secciones pueden agregarse en las secciones de frecuencia o rango superior. Como ejemplo, en nuestra tabla se recoge la presencia de la sección «derechos de propiedad», cuya frecuencia según nuestra metodología es indicativa de la existencia de una página web independiente que contenga estos contenidos. Ahora bien, en todos los sitios se ha convencionalizado la aparición de la forma lingüística convencional: «© Todos los derechos reservados». La existencia de estas páginas específicas independientes la asociamos con el otorgamiento tácito de una mayor importancia por parte del emisor del texto.



**Figura 69. Porcentaje de distribución de secciones en bloque «legal» en textos originales.**

En este bloque se hallan las mayores divergencias de frecuencia entre ambos corpus. En especial, la sección **declaración de privacidad** (F2) presenta el mayor diferencial porcentual del conjunto de bloques y secciones en este género (+57,15%; véase Figura 54). Estas diferencias son igualmente significativas en la sección **condiciones de uso** (F3), de escasa frecuencia en el corpus original (4,65%) y presente en el 38,94% de los sitios web localizados. Estos resultados estarían en consonancia con los obtenidos por Robbin y Stylianou (2003) en referencia a la sobrerrepresentación de los contenidos legales en los sitios web estadounidenses respecto a los de otras áreas geográficas.

Este aspecto se confirma igualmente en nuestro análisis de «palabras clave» o palabras cuya frecuencia es estadísticamente relevante en comparación con la frecuencia de aparición en un corpus general<sup>400</sup>. El análisis, en el que usamos sección original del corpus como corpus de referencia, muestra la mayor representación en los textos localizados de los términos relativos a esta sección: «privacidad», «privacy», «condiciones», «términos». La siguiente lista muestra estas posiciones predominantes:

<sup>400</sup> Realizado con la herramienta *Keywords* entre los textos originales y localizados.

N	WORD	FREQ <sub>loc</sub>	%	FREQ <sub>nat.</sub>	%	Keyness
20	términos	9,871	0.12	476		6,193.8
22	privacidad	14,090	0.16	1,535	0.03	6,019.5
36	uso	14,886	0.17	2,481	0.05	4,429.4
40	privacy	4,442	0.05	3		4,053.5
81	condiciones	12,529	0.15	2,681	0.05	2,730.8
86	declaración	3,291	0.04	46		2,641.6
89	política	7,541	0.09	1,112	0.02	2,532.9

**Tabla 18. Términos legales en lista de palabras clave del corpus localizado realizado con la herramienta *Keywords*.**

Podemos concluir que la sobrerrepresentación en la superestructura de estas consideraciones legales, especialmente de las políticas de privacidad, constituye el elemento diferencial más significativo entre los textos originales y los localizados. La justificación de estas diferencias hay que buscarla en el contexto legal estadounidense, que como ya describimos el análisis extratextual (véase 8.1.3), no regula por ley el ámbito legal de la privacidad como en España, sino que se autorregulan por las propias empresas bajo el auspicio de la Federal Trade Commision (Liu y Arnett 1999).

El análisis de las posibles relaciones legales recogidas en el corpus puede esclarecer este punto, ya que este primer análisis refleja cómo el entorno sociocultural de emisión del TO condiciona este bloque y lo distancia de los textos originales. Estas relaciones se apuntaron anteriormente (véase 8.1.3.1) y son:

- a) El registro legal que le confiere a la empresa su estatus jurídico como tal según la legislación española, con la inclusión del código CIF (Código de Identificación Fiscal) de la empresa y sus datos de inclusión en un registro mercantil en España. Esta relación se incluye normalmente en los **avisos legales** (F1). Los detalles específicos de inscripción en el registro mercantil aparecen en el 26,74% de los sitios web originales y en el 10,52% de los localizados.

- b) El cumplimiento de la Ley 34/2002, de 11 de julio, de Servicios de la Sociedad de la Información y de Comercio Electrónico que regula, entre otras, cualquier actividad económica realizada mediante Internet. El 10,46% de los sitios web originales menciona expresamente el cumplimiento de esta ley y, contrastivamente, esta mención expresa sólo aparece en el 3,15% de los sitios web localizados.
- c) La responsabilidad legal de la empresa en cuanto a la confidencialidad de los datos suministrados por el usuario. Este aspecto se encuentra regulado en España por la Ley Orgánica de Protección de Datos de Carácter Personal. En el análisis contrastivo arroja que el 33,13% de sitios web originales mencionan expresamente esta Ley frente al 11,5% de localizados. Esta diferencia considerable en la mención implícita a esta ley española entre textos originales y localizados asciende a casi el 20%<sup>401</sup>. Liu y Arnet (2000) mostraron que en el 2000 sólo el 51,40% de los sitios web estadounidenses de las grandes empresas<sup>402</sup> incluían una sección de privacidad. En nuestro corpus, recopilado en el 2006, los sitios web localizados muestran una cifra de inclusión de política de privacidad del 70,52% en sitios web de grandes empresas, un aumento del 19,12% en las versiones localizadas con respecto a las cifras recogidas en textos originales estadounidenses por estos investigadores.
- d) La relación jurídica vinculante entre la emisora del texto y el receptor del mismo en cuanto al texto web, expresada en las «condiciones de uso». Al igual que las consideraciones de privacidad, los aspectos legales referidos al uso, copia, conflictos legales, etc., se encuentran sujetos a la legislación vigente en España. En los sitios localizados su adaptación al contexto legal español es desigual. Por un lado, se pueden acoger a la legislación española como [www.chevrolet.es](http://www.chevrolet.es), o pueden acogerse a la legislación del país desde el que se

---

<sup>401</sup> En muchos otros países están vigentes leyes similares como la *Data Protection Act* de 1998 en Gran Bretaña (Mill 2005: 124).

<sup>402</sup> Examinaron todas las empresas de la lista Fortune 500.

controla el sitio web, como [www.BlackandDecker.es](http://www.BlackandDecker.es) que se rige de acuerdo a la legislación inglesa:

Este Sitio ha sido creado y funciona de acuerdo con la legislación española. Para la resolución de cualquier controversia en relación con el Sitio, serán aplicables exclusivamente las leyes españolas. [Chevrolet.es]

Esta página web está controlada y supervisada por Black & Decker desde sus oficinas en Inglaterra [...], cualquier conflicto en lo referente a los materiales contenidos en esta página web será amparada por la ley inglesa<sup>403</sup>. [Black&Decker.es]

- e) Sólo un texto localizado, en el sitio web de Electronic Arts, menciona expresamente el cumplimiento de la Directiva Europea de Protección de Datos 95/46/CE<sup>404</sup> del Parlamento Europeo y del Consejo del 24 de octubre de 1995, que regula la protección y uso de los datos personales recogidos tanto dentro como fuera de la Unión Europea.
- f) Por último, encontramos la protección legal que ampara el registro de los derechos de propiedad de los contenidos del sitio web, así como de los derechos de protección de las marcas comerciales. Esta protección se articula en las secciones **derechos de propiedad** (F4) y **marcas registradas** (F5), si bien este último bloque es exclusivo del corpus localizado (11,57%). La aparición de esta sección podríamos relacionarla en parte con la mayor frecuencia de aparición de la sección **I+D**, (C10), con un 24,21% de aparición en los textos localizados, frente a un 4,65% en textos originales, puesto que una mayor inversión en investigación conllevaría la protección de sus resultados.

#### 9.1.2.6.1. Análisis contrastivo de títulos en bloque «legal»

La primera aproximación al análisis contrastivo de la terminología de las diferentes secciones del bloque **legal** (F) nos muestra un alto grado de variación entre los textos originales y localizados. De hecho, en este bloque se manifiesta la mayor

---

<sup>403</sup> Errores de traducción puesto que se refiere a la «legislación inglesa», y los contenidos no son los referentes a «esta página web», sino al «sitio web» en su conjunto.

<sup>404</sup> <[http://ec.europa.eu/justice\\_home/fsj/privacy/](http://ec.europa.eu/justice_home/fsj/privacy/)>

variación denominativa entre ambos componentes del corpus. Como ejemplo, la sección **avisos legales** (F2) original muestra una alto grado de convencionalización de la unidad léxica «aviso legal» para denominar esta sección, mientras que el corpus localizado presenta trece unidades léxicas que no se encuentran en el corpus original. Este grado de variación denominativa queda patente en la amplia variación en el primer elemento nominal de la unidad léxica, «aviso», «nota», «información» y «advertencia», comunes a los originales y localizados, así como los sustantivos de presencia exclusiva en los textos localizados: «condiciones», «descarga», «declaración», «política», «acuerdo» y «comunicado». Asimismo, nos es de interés la aparición en la sección localizada de la forma plural «avisos legales» (11,36%). En su conjunto, el porcentaje de uso de unidades léxicas inexistentes en el corpus original en esta sección asciende a 54,4% del total.

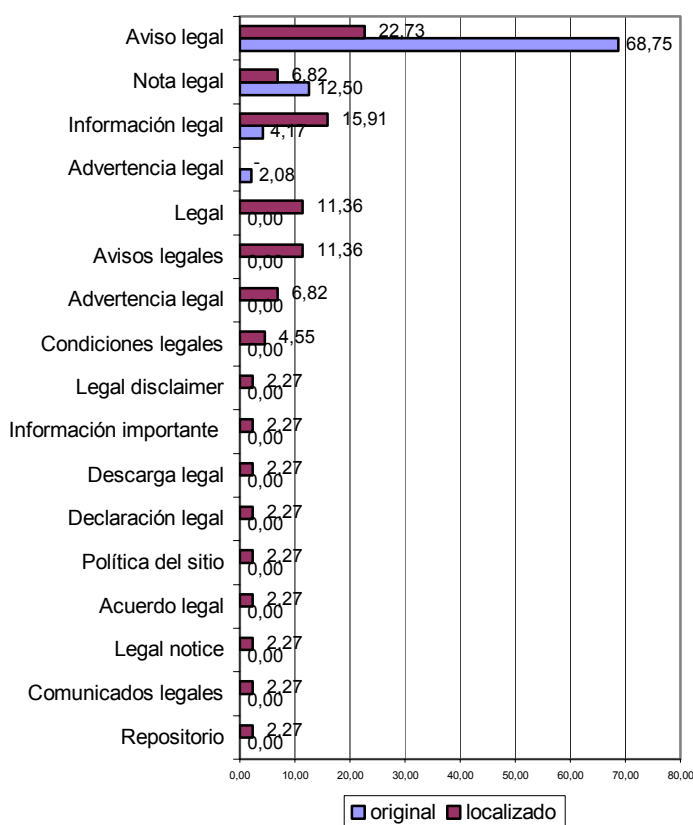
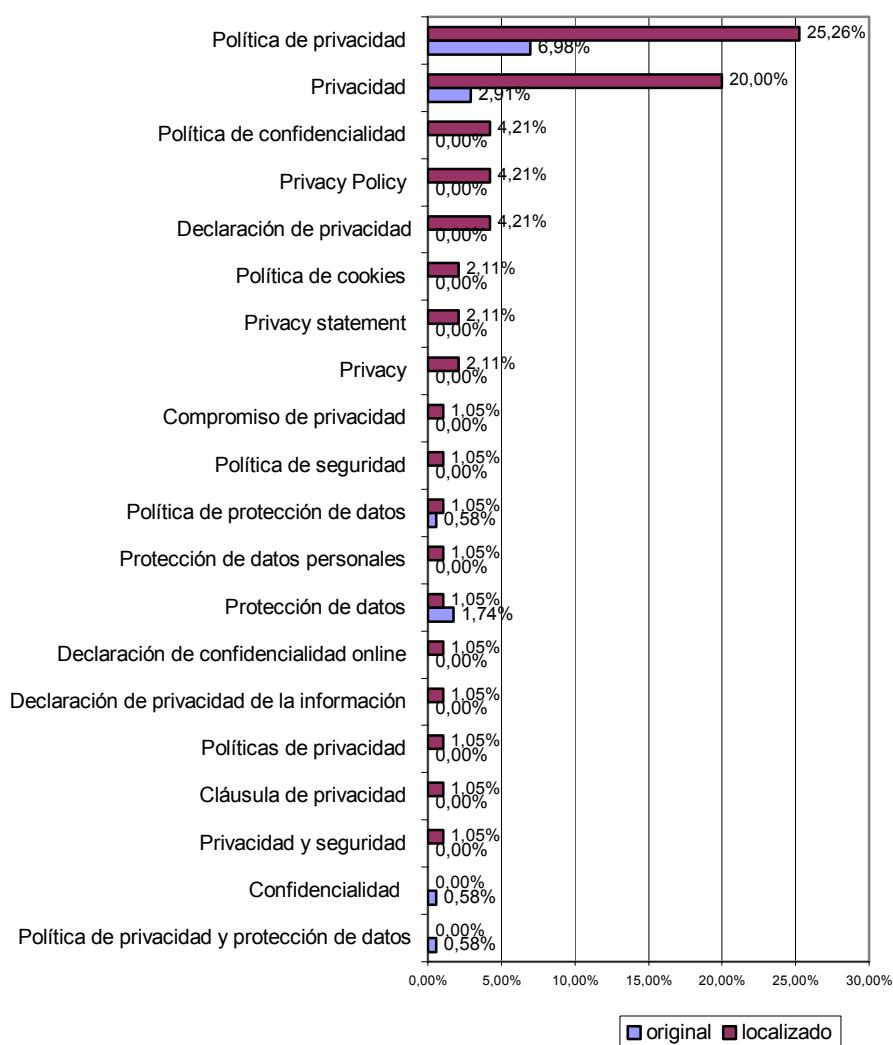


Figura 70. Análisis contrastivo de terminología en sección aviso legal (F1).

Esta misma tendencia se recoge en la terminología de la sección **privacidad** (F2). Con el objetivo de mostrar el efecto de los distintos niveles de convencionalización

de los bloques textuales en la mayor presencia de unidades léxicas localizadas, hemos elaborado una tabla de acuerdo al porcentaje de uso en el conjunto de los sitios web y, por consiguiente, se no se refiere al porcentaje relativo de uso entre las unidades léxicas encontradas en esta sección específica.



**Figura 71. Análisis contrastivo de terminología en sección privacidad (F2).**

Al igual que observamos en la investigación piloto (Jiménez 2005a), las diferencias principales se refieren a la diferente frecuencia de uso de las unidades léxicas, así como una mayor variación denominativa en aquellas secciones cuya diferencia en

su frecuencia de aparición sea significativa entre un mismo género en culturas diferentes. De hecho, en esta sección queda reflejado el mayor diferencial en la frecuencia de aparición entre ambas secciones del corpus. En nuestra opinión este hecho demuestra que el carácter altamente convencionalizado de esta sección en los TO estadounidense, junto con su escaso grado de prototipicidad en los textos originalmente producidos en España, se refleja en las grandes diferencias en cuanto al uso terminológico se refiere.

Si nos adentramos en la variación denominativa observada, en el corpus localizado aparecen catorce unidades léxicas alternativas sin equivalentes en el corpus original. La variación denominativa surge tanto en el primer elemento de la unidad léxica, «política», «declaración», «compromiso», «cláusula», como en el segundo elemento «privacidad», «confidencialidad», «confidencialidad online», «seguridad» «protección de datos [personales]», «seguridad». Este análisis terminológico demuestra el escaso impacto de la terminología presente en la ley vigente en España, «Protección de datos», con un 17,39% relativo de uso global en el corpus original y un 2,82% relativo en el localizado. No obstante, queda patente que la influencia del título oficial de la ley española es mayor en el componente original del corpus.

El uso del inglés en este bloque es significativo en el corpus localizado (11,27%), representado en las unidades léxicas «privacy policy» o «privacy statement». Se incluyen igualmente dos anglicismos puros en la formación de los compuestos nominales como «online» y «cookie»<sup>405</sup>. Estamos de acuerdo con Belda (2003: 127) al señalar el carácter de anglicismo del término «privacidad», que no obstante se recoge en la última edición del DRAE (edición XXII). A pesar de su carácter de

---

<sup>405</sup> Las *cookies* constituyen pequeños archivos que inserta un servidor web en los ordenadores que visitan los sitios web y que almacenan información sobre tiempo de visita, hábitos y preferencias de los usuarios, datos de registro etc. y por tanto constituyen uno de los aspectos sobre la protección de la intimidad del usuario de mayor importancia. Su traducción al español es problemática (Belda 2003), y se ha generalizado el uso del anglicismo puro, «cookie», propuesto por Microsoft, a falta de acuerdo sobre su traducción.



anglicismo, este término supone uno de los neologismos más extendidos en el campo de la informática e Internet.

Al igual que las anteriores, la sección **condiciones de uso** (F3) presenta una diferencia significativa de frecuencia de aparición entre ambos corpus, de ahí que constituya un bloque de interés para observar el posible impacto del distinto grado de convencionalización de un bloque textual en la terminología localizada. Su análisis contrastivo revela la existencia de cuatro posibles formas en los textos originales «condiciones de uso», «términos y condiciones», «términos de uso» y «condiciones generales», mientras que en el corpus localizado hemos identificado catorce alternativas, diez de ellas inexistentes en el corpus original. No obstante, ambos casos muestran como forma más convencionalizada «condiciones de uso», aunque esta forma aparece en el 11,58% del global de sitios localizados y sólo en el 1,74% de originales.

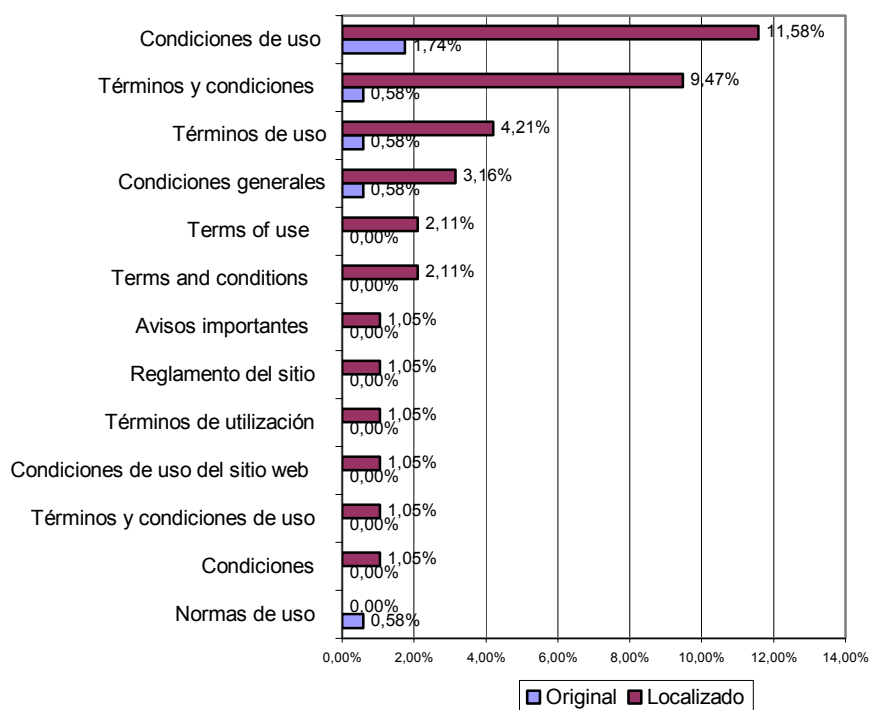


Figura 72. Análisis contrastivo de terminología en sección condiciones generales, F3.

Argumentamos que las unidades léxicas que aparecen en el corpus original se pueden considerar anglicismos léxicos y sintácticos, ya que constituyen calcos directos de las siguientes unidades léxicas: *terms of use* - «términos de uso», y *terms and conditions* - «términos y condiciones». Las unidades léxicas localizadas poseen una frecuencia del 9,74% y 4,21% respectivamente en el corpus localizado, y presentan una sola aparición en el corpus original (0,58%).

En el corpus localizado aparece además el sustantivo «utilización», sinónimo de la palabra «uso», que no está presente en el corpus original y que argumentamos responde a la búsqueda de alternativas durante el proceso de traducción. Por último, en el corpus original hemos identificado una unidad léxica que consideramos sería la que mejor se ajusta al español, el compuesto sintagmático «normas de uso».

En último lugar, el bloque **legal** (F) contiene las secciones **derechos de propiedad** (F4), y **marcas registradas** (F5). Esta última sección es inexistente en el corpus original. En los textos localizados responde a las siguientes unidades léxicas: «marcas registradas», «marcas», «marcas comerciales», «Trademark [information]» y «logotipos y marcas». La presencia exclusiva de esta sección sobre marcas registradas en los textos localizados queda recogida en la inclusión del símbolo anglosajón «™» *trademark*, en el 5,4% de las páginas localizadas y sólo en el 0,24% de las originales. A pesar de ser distintos, creemos pertinente apuntar que en los textos originales aparece con mayor frecuencia el símbolo de la marca registrada «®».

La peculiaridad más significativas desde el punto de vista léxico y terminológico en esta sección sería la posible especificación de todas las marcas comerciales registradas legalmente por la empresa. Como ejemplo, el sitio web de American Express contiene una página con la mención expresa de sus 75 marcas registradas. Las marcas registradas pueden tener tal estatus en España o estar registradas en

otro país. Como ilustración de este hecho podemos observar el siguiente texto que muestra esta posible dualidad: «Marcas de Symantec pendientes de registro y marcas sujetas al sistema jurídico anglosajón "Common-Law"» [Symantec].

Por último, en la sección **derechos de propiedad** (F4) la forma más convencionalizada en ambas secciones es el anglicismo puro «copyright», apareciendo además las siguientes alternativas: «derechos de propiedad», «reservados todos los derechos», «agente de copyright» y «titularidad».

### 9.1.2.7. Análisis del bloque «zonas especiales» (G)

El bloque comunicativo «Zonas especiales» se compone de secciones que se destinan a ciertos usuarios y cuyo acceso puede quedar restringido a usuarios registrados. Este bloque, una zona reservada con contenidos especiales para clientes (G), se presenta en el 22,09% de sitios originales y el 38,94% de localizados. La **función** global de este bloque es de índole expositiva, si bien cada posible sección cumple un objetivo expositivo específico destinado a satisfacer los diversos objetivos que motivan el acceso al texto, como la búsqueda de **trabajo** (G1), la solicitud de **presupuestos** (G9), los **consejos** de distinta índole (G2) o la **formación** en el campo específico de cada empresa (G3). La siguiente tabla muestra el grado de prototipicidad de las siguientes secciones, entre las que destacan con un mayor nivel de frecuencia las secciones **zonas de clientes** (G) (22,9%), y **trabajo** (G1) (19,76%).

Este bloque es de carácter **facultativo**, ya que no supera el 50% de frecuencia, y también es de carácter **múltiple**, puesto que cada sección se formula mediante un número indeterminado de páginas web que puede variar dependiendo del sitio y/o sección. Éste sería el caso de las diferentes páginas que describan cada empleo

disponible o los diferentes consejos relacionados con los productos o servicios ofrecidos.

<b>G. Zona de Clientes</b>	<b>Corpus original</b>	<b>Corpus localizado</b>
	<b>22,09%</b>	<b>38,94%</b>
1) Trabajo	19,76%	58,94%
2) Consejos	9,30%	23,15%
3) Formación	5,81%	21,05%
4) Tarifas	5,81%	3,15%
5) Ayuda/atención técnica	5,81%	29,47%
6) Pedidos	5,23%	25,26%
7) Publicaciones	4,65%	17,89%
i. Manuales/instrucciones	4,06	5,26%
ii. Folletos	2,90	4,21%
8) Inversores	4,06%	35,78%
9) Profesionales	2,90%	49%
10) Presupuestos	2,90%	2,1%
11) Franquicia	2,32%	0%
12) Regalos	0,58%	0%
13) Financiación	0,58%	5,26%
14) Recursos adicionales	0%	7,36%

**Tabla 19. Análisis contrastivo de secciones comunicativas en el bloque zonas de clientes (G).**

En virtud de la mayor importancia económica global de las empresas responsables de los sitios localizados, éstos presentan un mayor índice de frecuencia en estas secciones facultativas, aspecto este ligado al mayor número de páginas web por sitio que señalamos anteriormente. En este sentido, hemos observado cómo los sitios web de grandes empresas se convierten en el repositorio de la totalidad de contenidos que cualquier usuario potencial pueda necesitar (Price y Price 2002). A través del análisis de la terminología empleada en este bloque podemos clarificar el tipo de relación existente entre la empresa y el usuario; en estas zonas quedan recogidos veintiún posibles términos asociados con los usuarios potenciales (véase 8.1.2). A continuación analizaremos los resultados obtenidos para cada una de las secciones de este bloque:

- La sección **trabajo** (G1) es de carácter ocasional en los sitios web originales (19,76%) y de carácter recurrente en los textos localizados (58,94%). A través de esta diferencia podríamos concluir que la presencia de esta sección sería directamente proporcional a la importancia económica de la empresa, aunque no obstante, esta sección presenta los puestos de trabajo ofrecidos a escala internacional en los sitios localizados, y no exclusivamente los ubicados en el contexto geográfico español.
- La sección **consejos** (G2) presenta una mayor frecuencia de uso en los textos localizados (23,15%) frente a los originales (9,30%). El análisis de la terminología empleada para denominar este bloque revela que en los sitios web originales los consejos están relacionados mayoritariamente con la calidad de vida como: recetas [n. 6] y salud [n. 2]. En el corpus localizado los consejos se encuentran en los dominios de: la salud [n. 9]<sup>406</sup>, las actividades en el hogar [n.2].
- Las empresas realizan labores **formativas** (G3) en relación con sus productos o servicios, y éstas se dirigen tanto a usuarios en general como sus empleados. Estos cursos, seminarios o programas de prácticas pueden quedar reflejados en los calendarios de la sección **eventos** (D2). En ciertos casos hemos identificado la relación de las empresas con centros universitarios<sup>407</sup>.
- La sección **tarifas** (G4) aparece con una ligera mayor frecuencia en los textos originales, por lo que podríamos asociar esta sección con pequeñas empresas. Esta sección también puede denominarse con el término «precios». El bajo índice de frecuencia de esta sección se debe a que las tarifas y precios son

---

<sup>406</sup> Se incluyen términos como «diabetes», «paciente», o «dolor de cabeza».

<sup>407</sup> Como por ejemplo la relación entre la empresa Sun y la Universidad Carlos III de Madrid: «Sun patrocina el máster de la Universidad Carlos III en Ingeniería de la Web» [SUN].

obligatorios principalmente en sitios web de venta en línea, aunque como indicamos con anterioridad, éstos suponen un subgénero diferenciado.

- La sección de **ayuda** o **asistencia técnica** (G5) se presenta en empresas cuyos productos revisten una mayor complejidad técnica de uso, por lo que aparece especialmente en productos digitales o de software. Esta sección está presente en mayor medida en los textos localizados (29,47%) que en los originales (5,81%), y este hecho lo asociamos con la alta presencia de empresas de alta tecnología tras los textos localizados: Sun, Cisco, AMD, Microsoft, Electronic Arts, Lexmark, Adobe, HP, etc. Es de interés la inclusión mayoritaria del anglicismo «soporte» en el corpus localizado referido a la asistencia o atención técnica.
- La sección de **publicaciones** (G6) responde tanto a la presentación de las publicaciones emitidas por la empresa, como a un centro para descargar textos de distribución general, como manuales de instrucciones y folletos en formato «.pdf». Dado el contexto económico en el que se sitúan las empresas, esta sección puede incluir publicaciones dirigidas a receptores específicos, tales como los informes y balances para inversores.
- La sección **inversores** (G7) se sitúa como la segunda en este bloque según el porcentaje diferencial de frecuencia entre ambos corpus, 4,06% en el original frente al 35,78% del localizado. Este hecho se debe a que las grandes empresas que cotizan en bolsa han de hacer pública la información financiera, y los sitios web corporativos se han convertido en el principal medio de distribución de este tipo de información (Price y Price 2002). En el corpus localizado esta información se mantiene prioritariamente en inglés dado el carácter de *lingua franca* de este idioma en los mercados económicos internacionales.

### 9.1.2.7.1. Análisis contrastivo de títulos en el bloque «Área de clientes» (G)

El grado de convencionalización de la terminología propia de las distintas secciones presenta una gran variación dependiendo de la sección, con ciertas secciones representadas por una única unidad léxica o con escasa variación como «franquicia» (G11), «inversores» (G8) o «presupuestos» (G10), y otras secciones en las que presentan una gran variación denominativa como la sección **trabajo** (G1), o **consejos** (G2), en las que no se ha identificado ningún término con un porcentaje de uso superior al 12%.

Si comenzamos con el bloque en su conjunto, **zona de clientes** (G), éste se puede denominar con el tipo de receptor hacia la que se dirige y está directamente asociado a la relación entre la empresa y el usuario. En el corpus original éste puede ser «cliente» (45,90%), «distribuidores» (21,62%), o con una sola aparición cada uno «afiliados», «mayoristas», «minoristas» o «particulares». El área o el proceso por el que se accede presenta distintas denominaciones en los textos originales: «Acceso...» (29,72%), «Área» (21,72%), «Zona...» (16,37%), o «Club...» (5,40%). En el corpus localizado aparece la palabra «cuenta...» (5,4%) para indicar esta área reservada. Globalmente, las unidades léxicas más frecuentes en el corpus original son «Acceso clientes» (24,32%), «Clientes» (10,81%), «Área de distribuidores» (10,81%), «Área de clientes» (8,01%) y «Zona de clientes» (8,01%), y en el localizado «log in» (10,81%) y «clientes» (8,10%).

La sección **trabajo** (G1) presenta la mayor variación denominativa del conjunto de las secciones del corpus original. La unidad léxica de mayor frecuencia en este corpus, «Bolsa de empleo», presenta un mero 12,50% de frecuencia y no se identificó en el corpus localizado. Las siguientes unidades léxicas del corpus original con un 8,33% de frecuencia son: «empleo», «trabaja con nosotros», «trabaje con nosotros», «trabajo», «envíanos tu C.V.», «recursos humanos», y «ofertas de

empleo». Como indicamos anteriormente, en las formas verbales que forman parte de las unidades léxicas en esta sección se encuentra el mayor índice de uso del tratamiento de «tú» en la totalidad de las secciones. Este aspecto es de gran interés puesto que señala la posible relajación del tenor dentro de un mismo texto dependiendo del destinatario final de una sección textual (véase 8.1.2.3). En este caso las relaciones de poder cambiarían al incorporar al receptor como futuro miembro del colectivo emisor del texto.

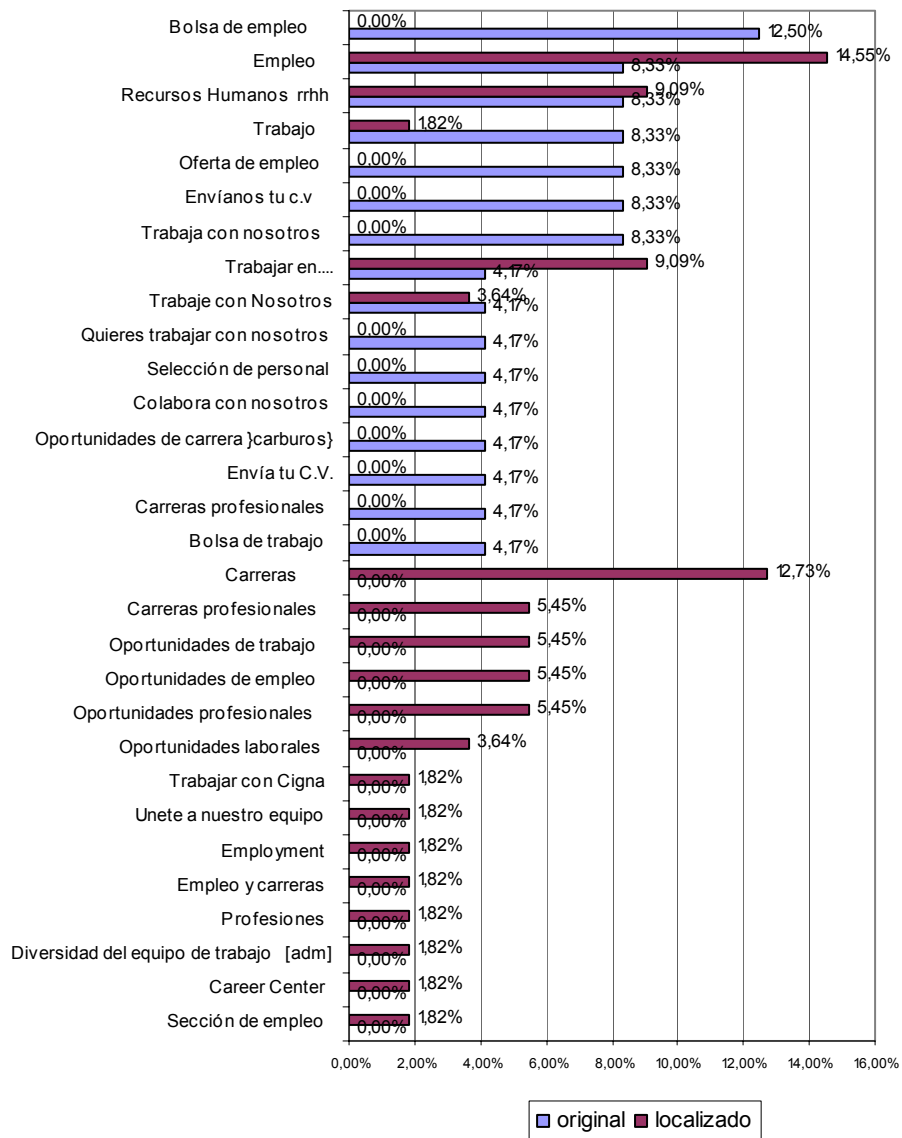
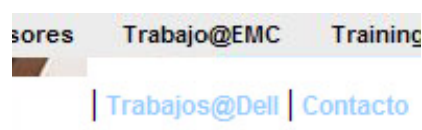


Figura 73. Análisis contrastivo de terminología en sección trabajo (G1).



Por su parte, la unidad léxica de mayor frecuencia en el corpus localizado es «empleo» (14,55%), seguido del anglicismo léxico «carreras» (12,73%), y «recursos humanos [rrhh]» (9,09%). La mínima correspondencia de la terminología entre ambas secciones del corpus refleja un escaso nivel de convencionalización de la terminología asociada a esta sección: sólo cinco de las dieciséis unidades léxicas empleadas en el corpus original se encuentran representadas en el localizado. Podemos caracterizar al conjunto de unidades léxicas localizadas por la presencia de gran cantidad de anglicismos léxicos: «carreras», «oportunidades de empleo / trabajo», «oportunidades laborales/ profesionales». En el corpus localizado, la forma nominal o adjetival de la unidad léxica que denota el concepto «empleo» puede ser: «empleo», «trabajo», «profesional», «laboral» «recursos humanos» o «carrera». De estos últimos, el adjetivo «laboral» y el sustantivo «carreras» son inexistentes en el corpus original.

Por último, en el corpus localizado hemos identificado dos casos de préstamos tipográficos propios de estos textos digitales (Bouffard y Craignon 2006). En «trabajos@Dell» se emplea el símbolo de la arroba @ que en inglés equivale a la preposición «at» o «en» en español, y que por lo tanto se interpretaría en inglés como «work at Dell». En el compuesto localizado se mantiene la arroba, aunque se pierde el significado original debido a que el signo @ no está asociado a preposición locativa alguna en español. De hecho, se encuentra asociada mayoritariamente con el correo electrónico en español. El compuesto resultante es ineficiente y presenta dificultades por su complejidad fonética, y aunque es comprensible en español, supone un cripticismo contrario a su objetivo de claridad y simplicidad. La siguiente imagen muestra los dos casos de calcos tipográficos encontrados en esta sección:



**Figura 74.** Calco tipográfico en la sección trabajo (G1). Sitios web de EMC y Dell.

La siguiente sección de interés desde un punto de vista contrastivo sería **inversores** (G8). La unidad léxica de mayor frecuencia en el corpus original es directamente «inversores» (33%, 21,1% en localizado). Se encontró igualmente una mayor variación en el corpus localizado con la inclusión del sinónimo «inversionistas» que tras su análisis en el corpus español CREA, podemos asociar a las variedades dialectales de Sudamérica, especialmente México, Venezuela y Colombia. Tal como apuntamos anteriormente, el carácter de lengua de especialidad del inglés en el campo de la economía se refleja en el alto índice del anglicismo puro «investor relations» en el corpus original (14,71%), una frecuencia similar a la de su forma en español, «relaciones con los inversores» (14,71%). Éste constituiría un préstamo léxico en español.

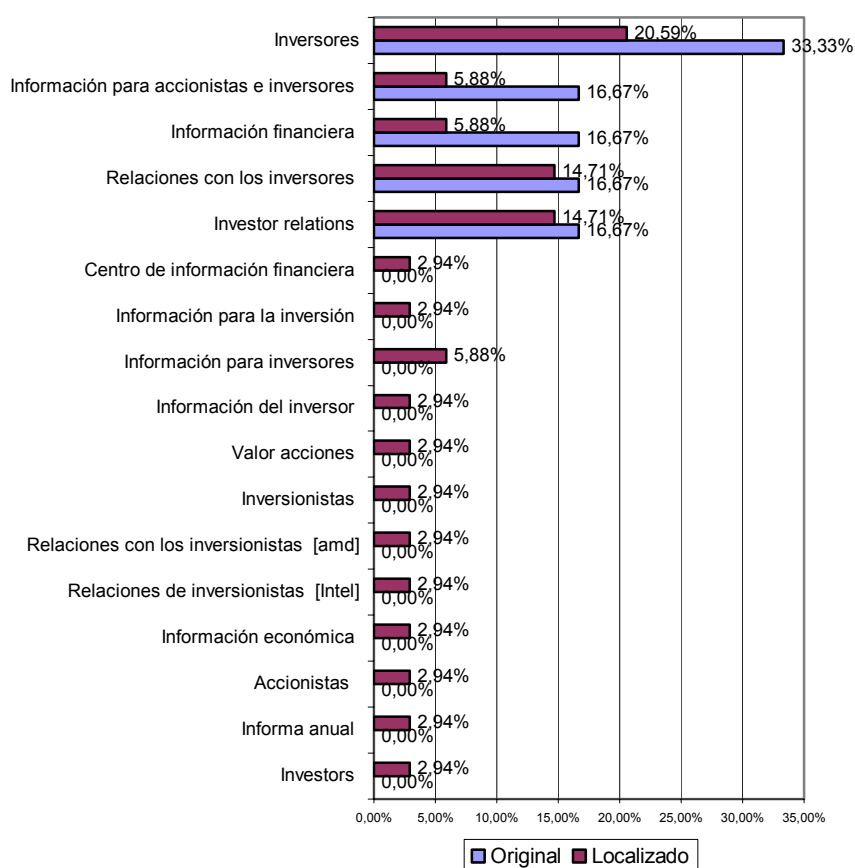


Figura 75. Análisis contrastivo de terminología en sección inversores (G8).

### 9.1.2.8. Análisis del bloque «interactividad con el sitio web» (H)

Este último bloque comunicativo engloba el proceso de interactividad entre el usuario y el sitio web como ente autónomo. La interactividad en Internet proporciona en su conjunto el carácter distintivo del género frente a textos similares en cualquier otro medio (Janoschka 2003), y constituye el rasgo principal por el que hemos asociado a estas secciones bajo un mismo bloque. Este intercambio comunicativo está limitado a las posibilidades de **funcionalidad** que incluya el desarrollador web, tales como búsquedas de distinta índole, ayuda con el uso del sitio web o la posibilidad de registrarse como usuario para acceder a contenidos reservados. Como establecimos en nuestro marco teórico, la funcionalidad de los textos web es el principal motor de su evolución (Shepherd y Watters 1999), con lo que los niveles de convencionalización de las diferentes secciones podrían ser los de más rápida variación a corto plazo.

	Corpus original	Corpus localizado
<b>H. Interactividad con sitio web</b>		
1) Mapa web	20,34%	60%
2) Buscar	14,53%	67,36%
3) Preguntas o ayuda [faq]	9,30%	35,78%
4) Enlaces	11,62%	15,78%
5) Regístrese	9,30%	22,10%
6) Imprimir	6,97%	6,31%
7) Glosario	5,23%	6,31%
8) Descargas	4,06%	17,73%
9) Comunidad	2,90%	5,26%
10) Accesibilidad	2,32%	2,1%
11) Su opinión	2,32%	9,47%
12) Juegos-	1,16%	5,26%
13) Herramientas	0 %	9,47%

Tabla 20. Análisis contrastivo de secciones comunicativas en el bloque interactividad (H).

La **función textual** de este bloque es principalmente exhortativa con alternativa (Hatim y Mason 1990), con una función secundaria expositiva. El fin de estas secciones es que el usuario «actúe» y sea parte participante activa en el intercambio comunicativo y, consecuentemente, ciertas secciones como **buscar** (H2), **regístrese**

(5), o **juegos** (H11), poseen una función textual exclusivamente exhortativa. En tres secciones identificadas de carácter metacomunicativo, **mapa web** (H1), **preguntas** (H3) y **glosario** (H7), la función textual primaria sería expositiva con una función secundaria exhortativa sin alternativa, ya que se centran en informar al usuario sobre ciertos aspectos del funcionamiento interno del texto global. Otras secciones de carácter metacomunicativo son **buscar** (H2) y **accesibilidad** (H10).

La mayor parte de las secciones en este bloque son de carácter **único**, ya que incluyen una sola página web para su realización, como **mapa web** (H1), **enlaces** (H4), o **regístrese** (H5). Igualmente, existen ciertas secciones de carácter **virtual** debido a su creación instantánea mediante un servidor activo como la sección **buscar** (véase 5.5.6.1).

En análisis contrastivo de las variables arroja una diferencia muy significativa en las secciones de mayor nivel de convencionalización: la sección **buscar** (H2) revela el mayor diferencial y se presenta con carácter ocasional en el corpus original (14,53%) y de carácter recurrente en el localizado (67,37%), aunque se sitúa cerca del umbral del 70% para los altamente recurrentes. Igualmente, la sección **mapa web** (H1) es de carácter ocasional en el corpus original (20,3%), y de carácter recurrente en los textos localizados (60%). Esta gran diferencia podría responder a que las guías de desarrollo recomiendan incluir estas dos secciones en sitios web que traspasen el umbral de 100 páginas (Nielsen y Loranger 2006: 138). Consecuentemente, este hecho puede deberse a la diferencia en el número medio de páginas entre ambas secciones del corpus (véase 7.4).

- Los **mapas web** (H1) constituyen la representación superestructural del sitio web y poseen una función expositiva de carácter metacomunicativo con una función secundaria exhortativa. La estructura jerárquica en los mapas web es la única identificada en nuestro corpus (Engbresten 2000), aunque también es posible su estructuración en forma de índice alfabético (McGovern *et al.* 2002:

20). Este carácter de representación jerárquica visual de la estructura textual la comparte con los menús de navegación aunque, por economía de espacio, los menús no incluyen todas las secciones sino que se restringen a las de mayor importancia (Price y Price 2002).

- La sección de **búsqueda** (H2) suele aparecer tras la activación expresa por parte del usuario mediante la introducción de las palabras de búsqueda. Su carácter es metacomunicativo y defendemos que constituye una sección diferenciada. Nuestra posición se fundamenta en que su activación resulta en una página web de resultados que posee una estructura y funcionalidad altamente convencionalizadas y que, a su vez, forma parte del modelo mental genérico que poseen los usuarios. La ubicación del enlace se encuentra convencionalizada en la esquina superior derecha de las páginas web, e incluye un campo para cumplimentar junto con un icono que activa la búsqueda. En el análisis del corpus localizado se encontraron traducciones erróneas debidas a una deficiente internacionalización previa. En el siguiente ejemplo se observa el resultado de estas deficiencias con un texto que refleja la sintaxis original inglesa o la traducción de una forma verbal imperativa por un infinitivo (véase capítulo 10):

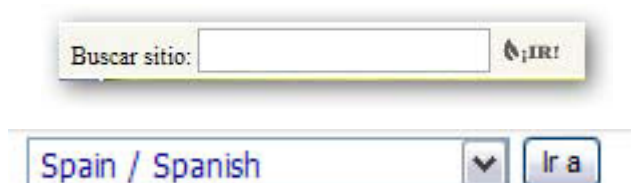


Figura 76. Errores de localización en la activación de la sección buscar (H2).

- Los sitios web constituyen una plataforma de asistencia al cliente que ha reducido los costes enormemente (Lommel 2007: 5). Las preguntas o **ayuda** (H3) constituyen una de las secciones más investigadas en el campo de los géneros digitales por el alto grado de convencionalización de las «Preguntas

más Frecuentes» o «FAQ» (Shepherd y Watters 1998; Crownston y Williams 1997; Santini 2005a). Éstas aparecen en la totalidad de las guías de estilo web como forma textual específica surgida en Internet (Price y Price 2002; Bly 2002; Jeney 2007). Las tres posibles alternativas, la atención al cliente, la ayuda y las FAQ se encuentran relacionadas, ya que suponen distintas posibilidades por las que asistir a usuarios que se enfrentan a problemas con los sitios web, o en menor medida, con los productos y servicios que ofrece la empresa. La estructura clara y concisa de las preguntas frecuentes tiene como objetivo el menor flujo comunicativo directo con la empresa (McAlpine 2001: 207), por lo que la alta calidad en su redacción y traducción es esencial. De no ser así, no se cumplirían sus objetivos de reducir los costes de asistencia y rapidez de acceso a la información que necesita el usuario. Hoy en día la evolución de esta sección queda reflejada en la presencia de un posible chat interactivo de asistencia (Mill 2005: 85).

- La sección **enlaces** (H4) refuerza la credibilidad del texto web, al reflejar la voluntad del emisor de informar convenientemente al usuario (Price y Price 2002). La existencia independiente de esta sección surge por la recomendación de no incluir enlaces externos en los sitios web corporativos, ya que entrañan el abandono del propio texto aunque, no obstante, esta sección puede asistir a aquellos usuarios cuyo objetivo sea encontrar información que no esté presente en el sitio en cuestión (Nielsen y Tahir 2002). El porcentaje de aparición es similar en ambos componentes del corpus, 11,62% en el original y 15,78% en el localizado.
- La sección **regístrese** (H5) cuenta con un formulario que permite, una vez cumplimentado, acceder a zonas reservadas dentro de los sitios web o a personalizar los contenidos o presentación. El objetivo final de esta sección se puede enmarcar dentro de los programas de fidelización de las empresas. Esta

sección se representa por medio de una sola página web interactiva, y por tanto es de carácter **único**.

- La sección **imprimir** (H6) consiste en la presentación de la misma página visitada en un formato texto y en la que se suelen excluir la mayoría de los gráficos. Esta sección se suele incluir en aquellas páginas o bloques comunicativos que hemos identificado como *e-textos*, tales como descripciones de **productos** (E1), **noticias** (D), o resultados de los formularios cumplimentados.

#### 9.1.2.8.1. Análisis contrastivo de títulos en el bloque «Interactividad»

1) Comenzaremos en primer lugar con la sección más frecuente en el corpus original, el **mapa web** (H1). El análisis contrastivo arroja una gran diferencia relativa a la unidad léxica más frecuente en ambos componentes del corpus. La unidad léxica más convencionalizada en el corpus original es el compuesto sintáctico, «mapa web», 54,3%, una forma que sería incorrecta en español desde un punto de vista sintagmático al no estar presente la preposición. A pesar de que se haya convencionalizado en esta sección, la frecuencia de «mapa del sitio», la forma correcta desde un punto de vista sintáctico, es 14,3%. Esta falta de la preposición en la formación del compuesto sintagmático podría deberse a la necesidad de brevedad en los menús de navegación, o quizás, a la falta de acuerdo relativa al género gramatical del sustantivo «web»<sup>408</sup>.

En los textos localizados la forma más frecuente es «mapa del sitio» (41,67%), que en este caso supondría una forma más correcta desde un punto de vista sintáctico y

---

<sup>408</sup> Belda (2003: 37) señala que lo más común es la asignación del género femenino si se refiere a una página web y a la WWW, la red mundial siguiendo el género del sustantivo principal en el compuesto nominal, y el masculino cuando se trate del sitio web, al asignarse el género de acuerdo con el sustantivo «sitio». En este caso, la forma correcta sería «mapa del web» puesto que esta sección constituye el mapa del sitio web. En el caso de usar la forma sustantiva femenina, haría referencia a una única página web o a al mapa global de la WWW, la red mundial.

terminológico que «mapa web». No obstante, esta unidad léxica refleja la forma más convencionalizada en inglés, «sitemap», que se presenta como segunda unidad léxica de mayor uso (18,33%). En el corpus localizado el uso de «mapa de la web», (5%), es superior a «mapa del web» (1,67%), que según Belda (2003) y Cabanillas (2007) sería la forma correcta de este compuesto sintagmático al asignarse el género gramatical según criterios semánticos en este caso. El uso terminológico en esta sección muestra uno de los menores índices de correspondencia entre ambas secciones, ya que la forma más convencionalizada en los textos originales aparece sólo en el 5% de los textos localizados. Estas unidades léxicas surgen de una identificación metafórica que en los textos originales puede representar los sustantivos «mapa» o «índice» y en los localizados aparecen además «guías» o «planos».

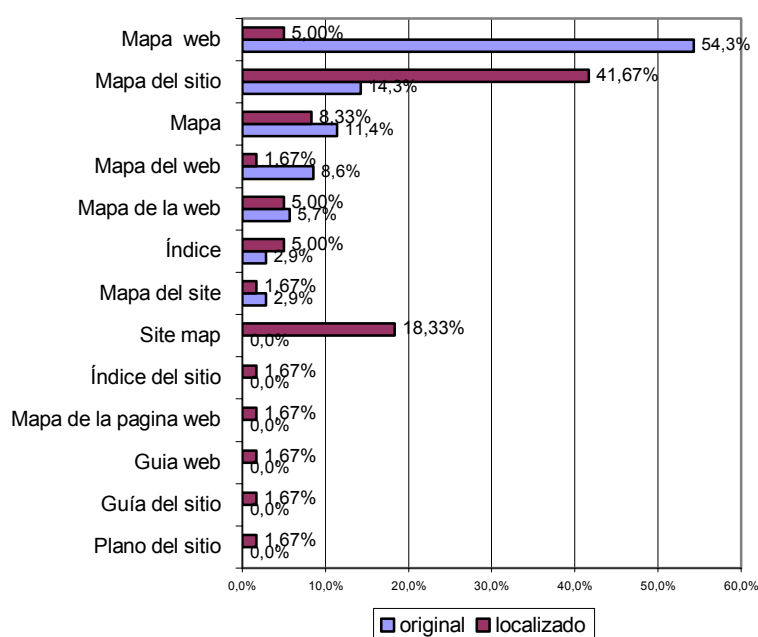


Figura 77. Análisis contrastivo de terminología en sección mapa web (H1).

Los textos localizados presentan ciertos anglicismos no adaptados como «site map» (18,33%) o anglicismos adaptados en compuestos sintagmáticos como «mapa del site» (2,9%). Se encontraron asimismo errores de traducción como «mapa de la página web» [Caterpillar], puesto que el mapa se refiere a la totalidad del sitio y



constituye un índice del conjunto de páginas web de las que se compone. Otras unidades léxicas localizadas son igualmente erróneas desde un punto de vista léxico como «guía web», con el 1,67% (n. 1), ya que apunta a ser una referencia de la totalidad de la WWW.

Las diferencias observadas se corresponden con los resultados de Nielsen y Tahir (2002), ya que la unidad léxica más usada en los sitios localizados, «mapa del sitio», se corresponde con «site map», presente en el 63% de sitios web corporativos estadounidenses. Las siguientes unidades léxicas por frecuencia según este estudio son «site index» (13%) y «site guide» (8%), cuya presencia se revela en la inclusión de la palabra «guía» en las unidades léxicas localizadas sin equivalentes en el corpus original.

2) La sección **buscar** (H2) muestra una frecuencia semejante de la unidad léxica de mayor frecuencia, «buscar», con un 69,23% en el original y un 54,10% en el localizado. La segunda unidad léxica más frecuente en el corpus original es la forma sustantiva «buscador», 11,54%.

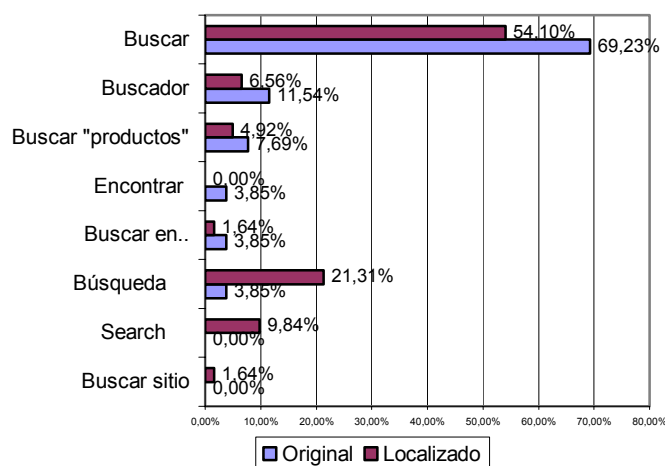


Figura 78. Análisis contrastivo de terminología en sección buscar (H2).

Debido a la interactividad que representa este bloque y a su función exhortativa, las unidades léxicas son mayoritariamente de carácter verbal, un 84,32% de originales y un 62,3% en los localizados. En estos porcentajes no incluimos el anglicismo no adaptado «search», al sernos imposible asignarle un carácter verbal o nominal en este idioma. La investigación de Nielsen y Tahir (2002) concluye que en los sitios web estadounidenses aparece *Search* en un 42% y *Go* en un 40%, recomendando finalmente *Search* por motivos de usabilidad. La mayor frecuencia de la unidad léxica «búsqueda» en los localizados, 21,31% frente a 3,85% en los originales, se puede deber a la conversión o derivación cero (Belda 2003: 177). Este fenómeno, característico de la lengua inglesa, responde a la posibilidad de que misma forma pueda pertenecer a distintas categorías gramaticales. En el caso de la unidad léxica más convencionalizada en inglés, *Search*, esta forma podría interpretarse como un verbo o sustantivo, y durante el proceso de localización el traductor ha de decidir si le asigna un carácter nominal o verbal. El análisis de los textos originales apuntaría a que, en aquellas unidades léxicas anglosajonas que presenten esta posibilidad de conversión, la traducción al español más apropiada sería una forma verbal en segmentos textuales de carácter interactivo.

Es de interés la aparición de la forma verbal «Encontrar» (3,85%), que supondría incluir el resultado del acto ilocutivo en la función realizada inexistente en el corpus localizado.

3) En tercer lugar, la sección **ayuda-preguntas** (H3) muestra la palabra «ayuda» como la unidad léxica más frecuente en ambas secciones. No obstante, es más significativa la diferencia en la segunda unidad léxica más frecuente: «FAQ», (27,76%) en el corpus original, y «preguntas frecuentes» (18,18%) en el corpus localizado, la traducción al español de *Frequently Asked Questions*. Otro calco de esta unidad léxica convencionalizada en inglés, «preguntas más frecuentes», aparece en el 9,08% de los textos localizados y no aparece en los textos originales.

Se observa, por tanto, que la traducción o explicitación completa del término «FAQ» es mayor en los textos localizados que en los originales.

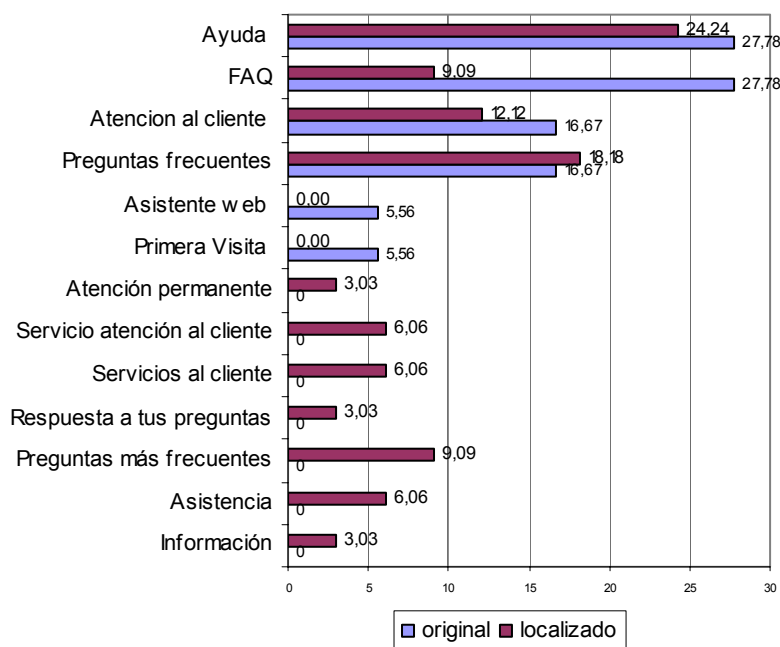


Figura 79. Análisis contrastivo de terminología en sección preguntas-ayuda (H3).

Al igual que en el caso de «Política de Privacidad», hemos constatado una influencia de la cultura anglosajona en la superestructura del género español, que en nuestra opinión se debe en parte al uso de guías de desarrollo y estilo web anglosajonas traducidas al español (Nielsen 2000; Kilian 1999; Price y Price 2002).

4) En cuarto lugar, la sección **enlaces** (H4) muestra directamente la palabra «enlaces» en ambas secciones como la unidad léxica de mayor frecuencia de uso. Queda patente un paralelismo con el caso anterior al aparecer anglicismo puro, «links» con una frecuencia mayor en los textos originales (35%) que en los localizados (6,67%). En éstos el término suele formar parte de un compuesto sintagmático nominal «de interés» o adjetival «relacionados / directos / rápidos».

Es asimismo de interés la aparición en los textos localizados del sinónimo «vínculo» (6,67%), inexistente en los textos originales, al igual que la aparición del

anglicismo léxico «enlaces rápidos» (20%), un compuesto sintagmático adjetival que se corresponde con «quick links» en inglés.

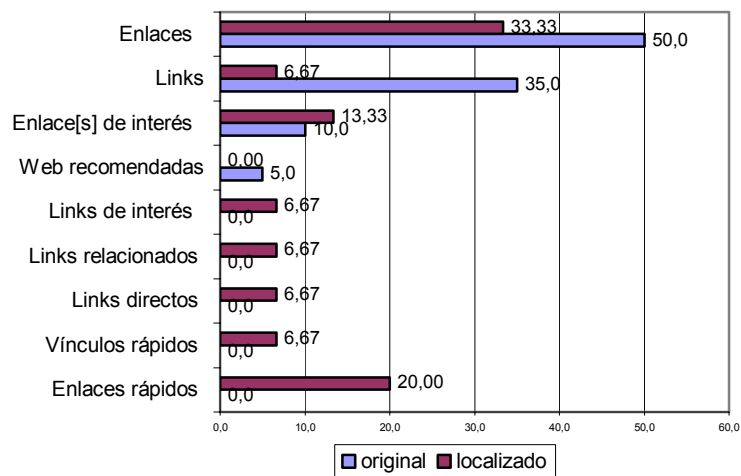


Figura 80. Análisis contrastivo de terminología en sección enlaces (H4).

5) El análisis de la terminología de la sección **regístrese** (H5) muestra dos posibles tipos de acciones que los usuarios pueden realizar: (1) un registro en el sitio web que posibilita la personalización o el acceso a funcionalidades y contenidos reservados o (2) una suscripción a otros contenidos que proporciona la empresa, como boletines de noticias. El primero se asocia al verbo «registrar» o «hacerse socio» y el segundo con la forma verbal «suscribirse». El corpus original sólo recoge la primera forma de registro en este caso, mientras que el corpus localizado muestra tres posibles lexemas «registro», «suscripción» y el sinónimo «inscripción», este último sin presencia en el corpus original. En los compuestos sintagmáticos nominales se refleja que el usuario al que se dirige la sección se denomina «cliente», «socio» o «afiliado» y el registro puede ser de «cuentas», «productos» o «datos».

El análisis muestra una preferencia por formas verbales que se corresponden con la función exhortativa propia de este bloque. En el corpus original las unidades léxicas con un componente verbal son las más frecuentes: «regístrese» (25,00%), «Hágase cliente» (13,33%), y «sea nuestro afiliado», «regístrate», «registrarse»,

«actívate» (6,67%). Globalmente el 81% de unidades léxicas incluyen una forma verbal tanto en el corpus original como en el localizado.

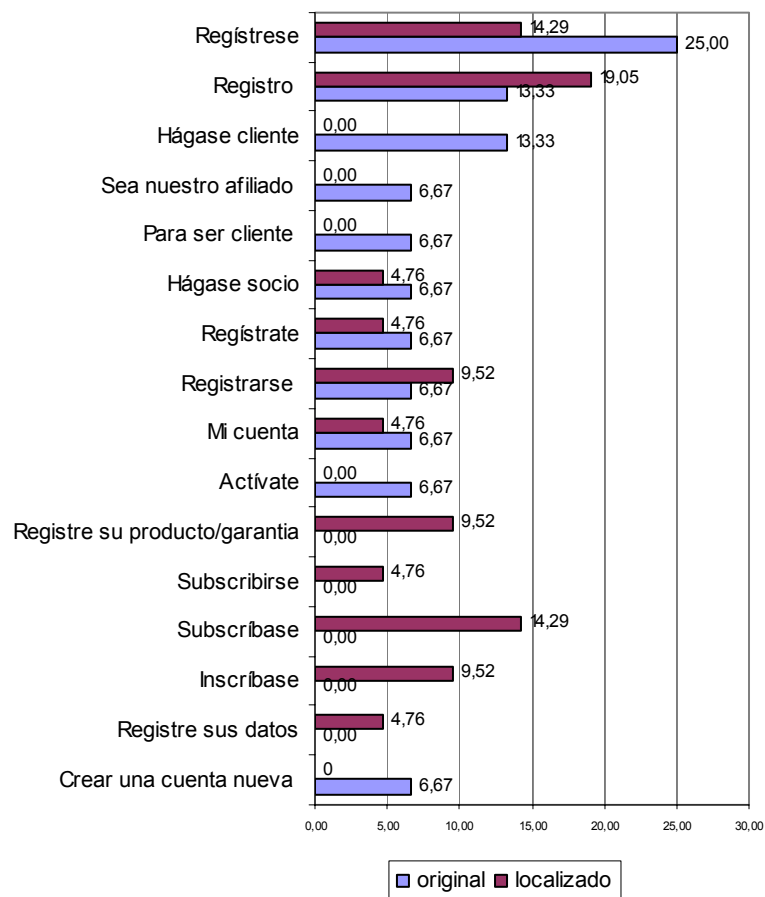


Figura 81. Análisis contrastivo de terminología en sección registro (H5).

A nuestro parecer, es significativa la frecuencia del sinónimo «inscribase» (9,52%), ya que este se presenta como una alternativa inexistente en el corpus original.

### 9.1.3. Discusión sobre el grado de convencionalización terminológica de terminología en menús de navegación

A continuación presentamos una tabla contrastiva en la que recogemos un resumen de las formas lingüísticas asociadas a cada bloque o sección y su frecuencia de uso. En ésta hemos incluido las dos primeras unidades léxicas de mayor frecuencia con el objetivo de contrastar las diferencias entre el nivel de

convencionalización de la terminología empleada en los textos originales con los niveles de frecuencia en los textos localizados. Con este análisis pretendemos ilustrar el grado de seguimiento de las convenciones establecidas en España en el caso específico de los menús de navegación, que constituyen uno de los elementos textuales más convencionalizados en este género (Nielsen y Tahir 2002). En la elaboración de la tabla se seleccionaron exclusivamente aquellos bloques y secciones con mayores índices de frecuencia, junto con las otras secciones representativas de mayor interés en nuestro estudio.

El análisis global de esta tabla arroja ciertos datos de interés con respecto a su contraste con las hipótesis de partida. En primer lugar, dentro del corpus original la media de frecuencia de la unidad léxica con mayor grado de convencionalización sería el 58,26%, mientras que la media de frecuencia de la unidad léxica de mayor grado de convencionalización en el corpus localizado es de 43,39%. Esta diferencia referente a la frecuencia apunta directamente a una mayor variación denominativa en el corpus localizado y principalmente, a una menor presencia de formas lingüísticas convencionales en este corpus, ya que es más pronunciada la distribución de las frecuencias entre las distintas posibles unidades léxicas. Si ahondamos en este aspecto, podemos repetir este mismo análisis con la media de frecuencia de la segunda unidad léxica en la frecuencia de uso. En el corpus original la media de frecuencia de la segunda unidad léxica más usada sería 19,05%, mientras que en el corpus localizado desciende al 13,26%. En su conjunto, la suma de los porcentajes de las dos formas más convencionalizadas en el corpus original sería 77,31%, mientras que en el corpus localizado esta misma suma arroja un porcentaje del 56,55%.

Bloques y Secciones	Corpus Original		Corpus Localizado	
	U. léxicas más frecuentes		U. léxicas más frecuentes	
A. Página de inicio	Inicio Home	60% 30,36%	Inicio Home	44,64% 24,04%
B. Contacto [interacción]	Contacto Contactar	49,42% 14,18%	Contáctenos Contacte con nosotros	21,11% 13,33%
C. Información sobre la empresa.	Empresa Quienes somos	37,2% 16,8%	Acerca de... Empresa	30,48% 13,04%
1)Localización	Localización Oficinas	20,41% 20,41%	Oficinas Donde comprar	21,82% 10,91%
2)Experiencia de la empresa	Clientes Trabajos-realizados	36,66% 30%	Casos de éxito Premios	18,51% 18,51%
3)Misión	Misión Filosofía	27,07% 10,6%	Responsabilidad Compromiso social	21,05% 18,42%
4)Sistema de calidad	Calidad Acreditaciones Certificados de calidad	37,5% 12,5%	Calidad Certificaciones	45,45% 18,18%
5)Historia	Historia	100%	Historia Historial	84,84% 3,03%
6)Instalaciones	Instalaciones	61,53%	Fábrica	50%
8)Medio Ambiente	Medio ambiente	100%	Medio ambiente Programa medioambiental	92,30% 7,70%
11)I+D	I+D Innovación	57,14% 28,54%	Tecnología I+D	35,29% 17,64%
D. Noticias/Actualidad	Noticias	40,63%	Noticias	34,11%
D2) Novedades	Novedades	14,58%	Sala de prensa	17,02%
D3)Notas de prensa	Notas de prensa	9,57%	Novedades	9,57%
3)Eventos	Eventos	100%	Eventos Actos	83,33% 8,33%
6)RSS	RSS	100%	RSS	100%
E. Productos -Servicios				
1)Productos	Productos Catálogo	68,47% 19,56%	Productos [nombre producto]	56,25% 35%
2)Servicios	Servicios Soluciones	88,31% 49,6%	Servicios Soluciones	53,97% 25,40%
3)Ofertas	Oferta Promoción	70,58% 23,52%	Promoción Oferta	50,% 28,57%
F. Legal				
1)Aviso-Nota Legal	Aviso legal Nota legal	69,75% 6,82%	Aviso legal Información legal	22,73% 15,91%
2)Declaración de privacidad	Política de privacidad Privacidad	52,17% 21,74%	Política de privacidad Privacidad	33,80% 26,76%
3)Condiciones Generales	Condiciones de uso  Normas de uso/ Condiciones generales/Términos de uso/Términos y condiciones	37,50%  12,50%	Condiciones de uso  Términos y condiciones	29,73%  24,32%
4)Derechos de propiedad	Copyright	100%	Copyright Derechos de propiedad Titularidad	75%  6,25%
5)Marcas registradas		0%	Marcas registradas Marcas	36,36% 18,18%
G. Zona de Clientes	Acceso clientes Clientes	24,32% 10,81%	Log in Clientes	10,81% 8,10%

Bloques y Secciones	Corpus Original		Corpus Localizado	
	U. léxicas más frecuentes		U. léxicas más frecuentes	
1)Trabajo	Bolsa de empleo Empleo	12,50% 8,33%	Empleo Carreras	14,55% 12,73%
8)Inversores	Inversores  Información financiera Información para accionistas Relaciones con los inversores Investor relations	33,23%  16,67%	Inversores  Relaciones con los inversores Investor Relations	20,59%  14,71%
H. Interactividad con sitio web				
1)Mapa web	Mapa web Mapa del web	54,21% 14,3%	Mapa del sitio Site map	41,64 18,33%
2)Buscar	Buscar Buscador	69,23% 11,54%	Buscar Búsqueda	54,10% 21,13%
3)Preguntas o ayuda[faq]	Ayuda FAQ  Atención al cliente Preguntas frecuentes	27,78%  16,67%	Ayuda  Preguntas frecuentes	24,24%  18,18%
4)Enlaces	Enlaces Links	50% 35%	Enlaces Enlaces rápidos	33,33% 20%
5)Regístrese	Regístrese  Hagase cliente Registro	14,29%  13,33%	Registro  Suscribase Regístrese	19,05%  14,29%
7)Glosario	Glosario	100%	Glosario  Lexicón Diccionario	66,66%  16,66%

**Tabla 21. Cuadro resumen contrastivo del grado de convencionalización de la terminología en menús de navegación.**

Una de nuestras hipótesis de partida consiste en que las localizaciones son el resultado de un proceso de producción textual sujeto a condiciones específicas, y este contexto llevará a textos con ciertos aspectos cuantificables (Baker 1995, 1996, 1998). En nuestro caso, si atendemos a las diferencias en el uso de las formas lingüísticas convencionales, el porcentaje de uso en el corpus localizado de aquellas unidades léxicas más frecuentes en el corpus original se reduce al 34,38%, una diferencia porcentual de 23,88%. Estos datos, junto con los ofrecidos anteriormente, confirmarían en parte nuestra hipótesis inicial.

El impacto del proceso de localización se puede reflejar asimismo en errores de traducción. Con este objetivo analizamos la frecuencia de uso de unidades léxicas



erróneas en los menús de navegación localizados. Esta frecuencia se obtuvo mediante el análisis de las formas localizadas en las veintidós tablas de frecuencia presentadas anteriormente. En el recuento de errores no se incluyeron las inadecuaciones relacionadas con las formas lingüísticas convencionales en los textos originales (véase 11.2). Los resultados de este análisis arrojan un porcentaje del 6,82% en el uso de unidades léxicas no adecuadas o erróneas. Este porcentaje significativo de inadecuaciones apunta a su carácter de los errores como posible marcador de los textos localizados de estos géneros, de ahí que por su importancia en la distinción entre los sitios originales y localizados los revisaremos en el capítulo 11.

<b>Convenciones terminológicas</b>	<b>Sección original del corpus</b>	<b>Sección localizada del corpus</b>
% de frecuencia de la unidad léxica más recurrente	58,26%	43,39%
% de frecuencia de la segunda unidad léxica más recurrente	<u>19,05%</u>	<u>13,26%</u>
% combinado de la primera y segunda unidades léxicas	77,31%	56,65%
% de la unidad léxica más frecuente en la <i>locale</i> ES-es	58,26%	<b>34,36%</b>
<b>Errores</b>		<b>6,82%</b>

**Tabla 22.** Tabla contrastiva relativa a frecuencia de uso de unidades léxicas convencionalizadas en menús de navegación.

Finalmente, a lo largo de nuestro análisis comparativo sobre la terminología hemos constatado una menor convencionalización de un bloque o sección en el género de estudio con un impacto observable en la mayor variación denominativa en los menús de navegación. Hemos constatado este impacto en la terminología de la sección **privacidad** (F2), en donde hemos identificado catorce unidades léxicas inexistentes en el corpus original. El corpus original muestra seis posibles alternativas en este caso frente a diecisiete en total en el corpus localizado. Con el objetivo de profundizar en este aspecto, hemos recurrido a la terminología de

aquellos bloques que presentan la mayor diferencia en su frecuencia y que presentamos en la Figura 52. Seleccionamos exclusivamente aquellas secciones que presentan una diferencia superior al 30%: F2, H2, H1, F3, C3, G9, E1. A continuación computamos el porcentaje de uso de la unidad léxica más frecuente en cada caso. En los textos localizados la media de frecuencia sería el 35,23%, relativamente menor al 43,39% de media en el conjunto presentada anteriormente. Por su parte, la frecuencia de uso de la forma más frecuente en estos mismos bloques de menor nivel de convencionalización en los textos originales es 49,71%. Estos datos confirmarían esta mayor dispersión y la asociación existente entre los distintos niveles de convencionalización de un bloque determinado en un género y la variación denominativa presente en sus traducciones.

## 9.2. Análisis intratextual de elementos en la estructura interna de las páginas web

Los sitios web corporativos presentan un doble eje superestructural paralelo: por un lado se encuentra la superestructura textual global compuesta por los distintos bloques textuales que hemos estudiado previamente, y por otro, ésta coexiste con otro nivel superestructural interno de cada página web (Bolaños 2002). En la presente tesis hemos decidido denominar a esta estructura interna **estructura intranodal**<sup>409</sup>, una denominación que se corresponde con los distintos niveles de coherencia textual en hipertextos<sup>410</sup> (véase 3.2): la coherencia **intranodal** se produce entre los elementos de una misma página web, mientras que la coherencia **internodal** es la que se produce entre el conjunto de elementos textuales globales de un sitio web (véase 5.3). Esta estructura intranodal o interna de cada página web se encuentra convencionalizada por las normas internacionales de base, es decir, los criterios mínimos para la integración de toda página web en la hiperweb

<sup>409</sup> Siguiendo el estudio sobre coherencia hipertextual de Engebresten (2000).

<sup>410</sup> La coherencia intranodal, internodal y hiperestructural en términos de Engebresten (2000). En la teoría del hipertexto la unidad la constituye el nodo hiperestructural como unidad (Landow 1992).

global (véase 3.3). Parte de estos elementos se encuentran normalmente ocultos al usuario final, es decir, forman parte del *back-end* o estructura profunda de la página. Estas normas de base, tales como la estructura básica html, son de obligado cumplimiento para conseguir que la página sea funcional en el contexto tecnológico de la web. No obstante, desde nuestra perspectiva traductológica son de interés aquellos elementos que no son de carácter recurrentes sino optativos, puesto que éstos incluyen elementos textuales que se procesan durante la tarea de la traducción y forman parte de este proceso.

Los elementos estructurales significativos en el proceso de localización se encuentran cuantificados contrastivamente en el análisis que realizamos con la herramienta informática Webbudget (véase Tabla 4, p. 273). Desde nuestra óptica, los elementos de mayor interés por constituir subtextos sobre los que opera el traductor-localizador son: (1) los títulos de las páginas web, (2) los textos en el atributo «Alt», (2) los textos en la sección <Meta>, (3) los textos en las etiquetas «OnMouseOver», «Select» e «Input», (4) el texto en scripts y (5) las imágenes y presentaciones. Pasamos a continuación a analizar estos elementos internos de toda página web que conforman la estructura intranodal.

### **9.2.1. Los títulos de las páginas html**

Los títulos de las páginas aparecen en la barra superior del navegador y podemos resumir sus objetivos en: (1) dar a conocer al usuario dónde se encuentra, (2) identificar la página en las listas de favoritos del usuario y en listados en las páginas de resultados en buscadores y, (3) describir la página para su categorización por parte de los buscadores e indexadores de páginas web (Garrand 2001). Los títulos son un componente esencial en toda página a pesar de que el usuario raramente lo lea (Brinck *et al.* 2002: 273) y se identificar por la etiqueta html <title> [...] </title> en cuyo interior se sitúa el texto a traducir.

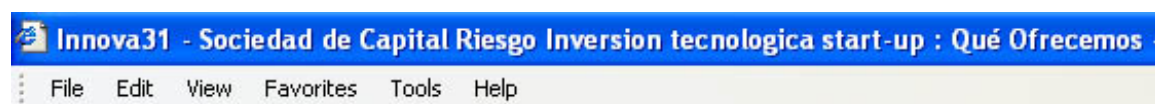
Con el objetivo de analizar los títulos de las páginas web recopilamos dos subcorpus del conjunto de títulos de las páginas web originales y localizadas. A continuación los analizamos léxicamente con la herramienta informática Wordsmith Tools, con lo que obtuvimos la lista de palabras lematizada por orden de frecuencia. En segundo lugar comparamos ambas listas con la herramienta *Keywords* para obtener las palabras clave de mayor representación en el corpus localizado frente al original, que usamos como corpus de referencia. La siguiente tabla muestra las palabras de mayor frecuencia en los títulos. En esta lista queda patente la alta presencia de los nombre propios que se corresponden con las empresas. Hemos subrayado estos nombres propios para su fácil identificación.

Lista original		Frec.	Lista localizado		Frec.
1	com	2.304	1	lexmark	4.844
2	bosch	1.988	2	ford	3.244
3	electrodomésticos	1.972	3	spain	2.737
4	ono	1.533	4	españa	2.265
5	particulares	1.115	5	es	1.836
6	publiestilo	1.051	6	hp	1.734
7	televisión	975	7	introducción	1.436
8	croexsa	841	8	motorola	1.058
9	tormo	809	9	emc	715
10	franquicias	784	10	air	620
11	jamones	759	11	productos	619
12	portal	757	12	sealed	599
13	madrid	750	13	europa	550
14	internacionales	728	14	medio	546
15	spain	723	15	medtronic	526
16	normas	721	16	oriente	513
17	contabilidad	720	17	artes	511
18	iasplus	720	18	gama	508
19	sl	680	19	gráficas	498
20	consulta	604	20	áfrica	437
21	programación	603	21	lauder	430
22	untitled	525	22	apple	406
23	caja	514	23	estée	392
24	web	508	24	ea	385
25	grupo	500	25	soluciones	382
26	bellota	470	26	boeing	363
27	españa	463	27	accesorios	351
28	tiendape	462	28	nuevo	348
29	productos	438	29	battlefield	343
30	franquicia	416	30	dell	335
31	ibéricos	395	31	cartridge	327
32	seguridad	389	32	online	313
33	domecq	388	33	software	305
34	industrial	384	34	portal	303
35	menú	368	35	print	298
36	datos	336	36	servicios	293
37	libre	336	37	mfp	285
38	instalación	334	38	server	283
39	jazztel	326	39	ups	283
40	cm	321	40	carrocerías	282

Tabla 23. Palabras de mayor frecuencia en subcorpus de segmentos textuales en <title>.

Estadísticas obtenidas con Wordlist de Wordsmith Tools.

El análisis de ambos subcorpus muestra el carácter descriptivo de los títulos, que se corresponde con una **función** expositiva en descripción. En los mismos ocupan un lugar prominente los nombres de productos y empresas, con un 25% de nombres de empresas entre las cien primeras palabras en el original y 26% en el localizado. Es necesario recordar el diferente número de páginas que puede contener cada sitio web, de ahí que podamos concluir que en nuestro corpus original el sitio web con mayor número de páginas sea posiblemente el de «Bosch», y en el localizado el de las empresas «Lexmark» y «Ford».



**Figura 82.** Título de la página web de [www.innova31.es](http://www.innova31.es) con indicación del título general del sitio web y de la página específica del mismo.

Este carácter descriptivo se muestra por la presencia casi exclusiva de sustantivos frente a las formas verbales, adjetivales o adverbiales. Entre las 100 primeras palabras en el corpus original, el 85% los constituyen sustantivos, el 12% adjetivos y el 0% verbos. La única forma verbal identificada, «prevenir», forma parte del nombre de una empresa, «Grupo Prevenir», y constituye por tanto un nombre propio. En el corpus localizado el 92% de palabras se corresponden con formas sustantivas, el 5% adjetivas y el 2% verbales.

Nuestro análisis contrastivo muestra una ligera mayor presencia de sustantivos en el corpus localizado y una menor presencia de adjetivos, el 5% frente al 12%, que se explica en parte por la mayor inclusión de adjetivos en los nombres de las empresas españolas, como «EURISA. Bilbao Euro-Industrial. S. A.» o «Saniber Grifería Temporizada. S. L.», y una presencia mayor explicitación de la labor de la empresa: «Grupo Prevenir Consulting: Prevención de riesgos *laborales*». Los

adjetivos más frecuentes en el corpus original serían: «libre», «internacional», «industrial», «ibéricos», «virtuales», «temporizada», «electrónicos», «privado», «corporativa», «laborales». En el corpus localizado son: «gráficas», «nuevo», «pequeña», «mediana» y *high*. Este último adjetivo en inglés se corresponde con la descripción de productos de la empresa HP mantenido en inglés.

La frecuencia prominente del nombre del país, *Spain* y «España» refleja tanto la vertiente internacional de la comunicación web de estas empresas como el carácter descriptivo contextualizador de estos títulos. Estas palabras aparecen respectivamente en las posiciones tercera y cuarta en la lista del corpus localizado, y la decimoquinta y vigésimo séptima del original. La presencia del nombre del país es más importante en la sección localizada del corpus, y este hecho se debe a una mayor necesidad de contextualizar el sitio específico para la *locale* de destino, España, dentro del carácter multilingüe de los sitios localizados.

El uso de formas verbales es insignificante en estos subtextos: el 2% en el localizado y 0% en el natural entre las 100 primeras palabras<sup>411</sup>. Al igual que la palabra «prevenir», encontramos «alquiler» en el puesto 84 pero se corresponde con el servicio ofrecido por una empresa inmobiliaria y, por tanto, una forma nominal. Por su parte, las formas verbales encontradas en el localizado son «ver» en el puesto 72 y «vive» en el 92. La segunda forma verbal en modo imperativo que se correspondería con una exhortación en el corpus original es «adquiera» en el puesto 245. Es lógico concluir que la presencia casi inexistente de formas verbales confirma su carácter eminentemente descriptivo.

La aparición de la abreviación «com» como primera palabra en el corpus original es indicativo de la presencia en los títulos del dominio del sitio web, como

---

<sup>411</sup> Como se indicó anteriormente, la aparición de «prevenir» se corresponde con el nombre propio de una empresa.

«www.azud.com»<sup>412</sup>. Esta abreviación normalizada en los dominios de Internet constituye un truncamiento regular de *commercial business* (Belda 2003: 188). En el corpus localizado aparece «es» como quinta forma, que tras un análisis de concordancia se corresponde casi exclusivamente con la inclusión del dominio de la empresa con el dominio «.es», en lugar de la forma verbal del verbo «ser». Esta abreviación constituiría asimismo un truncamiento regular de la palabra «España»<sup>413</sup>. Este aspecto es indicativo de la preferencia por el dominio «.com» en los sitios web corporativos españoles frente al dominio específico de cada país, «.es», en los localizados. En los sitios web localizados de multinacionales el dominio «.com» se suele corresponder con el sitio web corporativo en inglés.

Por último, en el corpus original aparece la palabra *untitled* en el puesto 22, que revela la falta de título en el 2,75% de las páginas web frente al 0,1% de las localizadas.

Desde una perspectiva prescriptiva, podemos observar de nuevo en las recomendaciones de las guías de usabilidad las características textuales básicas de estos segmentos textuales (McAlpine 2001: 152-156; Mill 2005: 135; Garrand 2001). Las podemos resumir en:

- Nunca usar más de 10 palabras o un máximo de 80 caracteres, siendo esenciales los primeros 65 caracteres, puesto que es la media que mostrarán los motores de búsqueda. En el caso de incluir más de cinco palabras, puede que las últimas no aparezcan en las listas de favoritos. Esta recomendación se cumple en nuestro corpus: la media de palabras en los títulos originales es 6,49 y 4,69 en los localizados

---

<sup>412</sup> Las herramientas informáticas léxicas consideran el punto como un separador entre palabras, de ahí que «.com» o «.es» se recuente con una palabra independiente.

<sup>413</sup> Los códigos de países se encuentran normalizados por la norma <sup>413</sup> ISO 3166 *Codes for the representation of names of countries*. Entre los códigos de países normalizados se encuentran asimismo truncamientos irregulares como el caso de México, «mx», o Chile, «cl».

- Dar un título diferente a cada página para aumentar las posibilidades de que se encuentre el sitio web.
- Resumir el contenido de la página de igual modo que los títulos en periódicos.
- Incluir un porcentaje alto de palabras clave que resuman el contenido.
- Variar la longitud.
- No usar caracteres extraños como &, %, \$, #, @.
- La claridad absoluta y la brevedad son primordiales, sin incluir palabras no relevantes.
- Su objetivo descriptivo desaconseja el uso de la ironía, juegos de palabras, humor, etc.
- No comenzar el título con determinantes o pronombres. Esta recomendación se cumple por igual en ambas secciones.

### **9.2.2. Las descripciones en el atributo ALT y la accesibilidad**

Los segmentos textuales en el atributo <alt> describen las imágenes en las páginas web y son de utilidad para las personas con discapacidad visual<sup>414</sup>. Su presencia creció de forma exponencial con la aparición de leyes en diferentes países que regulan la accesibilidad web (McAlpine 2001). En el contexto estadounidense, las enmiendas a la Sección 508 de la Rehabilitation Act estipulan que toda la información electrónica que se desarrolle o mantenga con fondos federales ha de ser accesible para personas con discapacidades. A pesar de ser de aplicación exclusiva en sitios institucionales, constituye la norma que se demanda a las grandes empresas con presencia en la WWW (Jeney 2007).

---

<sup>414</sup> Originalmente, estos subtítulos eran además de utilidad para aquellos usuarios que preferían no descargar las imágenes por defecto, sea por que cuentan con conexiones lentas, por que usan navegadores de solo texto o por que visualizan las páginas web en dispositivos con pantallas reducidas como teléfonos móviles o PDAs.



En el desarrollo de directrices relativas a la accesibilidad desempeñan un papel clave la Iniciativa de Accesibilidad Web (WAI)<sup>415</sup> y el Educación & Outreach Working Group (EOWG)<sup>416</sup> ambos del Wide Web Consortium (W3C). Ocupa una importancia primordial la descripción por medio de este atributo de imágenes, animaciones, mapas visuales, gráficas, tablas y demás elementos multimedia. La descripción textual de estos elementos se sitúa como la primera norma de obligado cumplimiento por parte de los desarrolladores (W3C 2000):

1.1 Provide a text equivalent for every non-text element (e.g., via "alt", "longdesc", or in element content). This includes: images, graphical representations of text (including symbols), image map regions, animations (e.g., animated GIFs), applets and programmatic objects, ASCII art, frames, scripts, images used as list bullets, spacers, graphical buttons, sounds (played with or without user interaction), stand-alone audio files, audio tracks of video, and video. (W3C 2000)

A pesar de la importancia de estos segmentos textuales, hemos constatado que no suelen incluirse con regularidad. En nuestro corpus original este atributo está presente en el 61,71%, frente al 89,21% en el corpus localizado. En la publicación de Nielsen y Tahir (2002) sobre sitios corporativos estadounidenses se recoge que solamente el 42% de los sitios web incluyen este atributo.

Variable	Original	Localizado
Páginas con Alt.	61,71%	89,21%
Total atributos Alt con texto descriptivo	44,72%	47,88%
Alt Vacíos o con un carácter	55,28%	52,18%
Palabras por Alt con texto	3,51	1,764
Media de palabras en Alt por pág.	15,18	18,53

**Tabla 7. Estadísticas globales sobre textos en el atributo «Alt».**

<sup>415</sup> <http://www.w3.org/WAI>.

<sup>416</sup> La labor del grupo EOWG se encamina a la educación del conjunto de participantes de la comunidad web en lo relativo a la accesibilidad, [www.w3.org/WAI/EO/](http://www.w3.org/WAI/EO/).

Sin embargo, la presencia del atributo Alt no es en sí misma indicativa de su uso efectivo, ya que nuestro análisis demostró que el 55,28% de éstas en el corpus original y el 52,18% en el localizado no contenían texto alguno. En este porcentaje hemos incorporado aquellos segmentos que presentan un solo carácter, como \*, [, -, \*, /, espacios o números. Este alto índice se debe a que las guías de estilo recomiendan incluir un Alt vacío para aquellas imágenes que organizan la presentación sin una función descriptiva, como las que sirven de separadores entre texto. En caso contrario, las herramientas informáticas que leen las páginas web para personas ciegas unirán las palabras de segmentos textuales distintos<sup>417</sup>. En nuestro análisis incluimos los segmentos *Alt* con dos caracteres, puesto que éstos contenían palabras descriptivas con contenido proposicional como *GO, IR, OK, EA* (Electronic Arts), o L1, L2 para la selección de lengua.

La media de palabras es 3,51 por Alt en el corpus original y 1,76 en el localizado. En lo que respecta al contenido textual, el predominio de formas sustantivas es casi absoluto, con una aparición insignificante de formas verbales en analogía con los segmentos en el atributo *Title* (véase 9.2.1). La presencia de palabras es mayor en el corpus original que en el localizado, en un porcentaje cercano al 50%. Esta diferencia se repite en el análisis de los segmentos textuales en el atributo *Keywords* (véase 9.2.3), otro segmento invisible en la estructura de presentación de la página web. Estos resultados apuntan a una menor importancia en textos localizados de aquellos segmentos textuales que requieren traducción pero que no se visualizan directamente en esta estructura de presentación.

El mayor porcentaje de palabras que incluyen la etiqueta Alt en los localizados, 89,21% frente a 61,71% en originales, apunta a un uso más extendido de herramientas informáticas de desarrollo web y al uso de programas de gestión de

<sup>417</sup> En caso de no incluir el atributo Alt vacío los lectores de páginas web para personas con discapacidad visual leerán repetidamente o «imagen» o el nombre del archivo, como «SRC='spg/getImageed36.gif?imageName=SPG\_13\_29\_0\_21539.gif&filename=594x243.jpg'», ejemplo de la página web de [www.ford.es](http://www.ford.es).

contenidos multilingües GMS<sup>418</sup> (Global Management Systems), que insertan automáticamente el atributo Alt en el código de la página con cada imagen.

En nuestro análisis hemos dejado patente la escasa atención a las cuestiones de accesibilidad, que se refleja igualmente en la escasa presencia de la sección comunicativa sobre las directrices de accesibilidad, la sección H10. En el corpus original existen cuatro sitios que contienen una sección comunicativa sobre accesibilidad, el 2,32%, mientras que en el localizado solo existe un sitio web, [www.adobe.es](http://www.adobe.es), el 1,4%. Otros dos sitios web localizados incluyen un enlace hacia este bloque, pero el contenido se encuentra sin localizar<sup>419</sup>. Este hecho sería indicativo de la menor importancia que otorgan las empresas a esta sección, en virtud de su decisión de mantenerla en inglés.

Las recomendaciones de las guías de estilo sobre los segmentos textuales en Alt se dirigen a la redacción de descripción clara y sucinta de los mismos (Garrand 2001). En el proceso de redacción McAlpine (2001: 101-113) señala:

- Encontrar la función que desempeña la imagen en el contexto de la página.
- No incluir nada cuando se trate de imágenes decorativas sin función alguna que organizar o decorar el texto.
- En caso de que una imagen se use como fin de párrafo o sección, incluir esta descripción de su función «fin de la sección», «fin del párrafo», etc.
- Terminar el alt text con un punto con el objetivo de que haya una pausa en la lectura de la página entre imagen o imagen, Ej. «alt="Logotipo de la empresa."».

---

<sup>418</sup> Ciertos sitios web muestran sistema GMS que utilizan en sus páginas web como el caso de Oracle, «Powered by Oracle Application Server Portal»

<sup>419</sup> Los sitios web con el bloque de accesibilidad en inglés y, por tanto, no localizado son [www.apple.es](http://www.apple.es) y [www.hp.es](http://www.hp.es).

- Siempre describir la imagen si su función es parte del contenido del texto.
- Los robots indexadores de páginas inspeccionan este elemento textual para categorizar las páginas, ya que describen una parte integrante del texto. No obstante, se recomienda no usarlo para incluir texto *keywords* o incluir demasiadas palabras (Mill 2005: 139).
- En caso de que la descripción de una gráfica o presentación que incluya grandes cantidades de texto, incluir una descripción extensa por medio del atributo *longdesc*, que enlaza un elemento visual o multimedia a describir con una página web independiente para incluir su descripción. En el corpus localizado sólo dos sitios web emplearon este atributo, [www.ford.es](http://www.ford.es) y [www.hp.es](http://www.hp.es), el 2,1%, frente al 2,9% de las originales, un porcentaje similar.

### 9.2.2.1. Elementos visuales y multimedia

Los elementos multimedia como sonido, imágenes con texto, animaciones flash, vídeos, clips de sonido, etc., son elementos esenciales en la textualidad web y esenciales en la labor localizadora (Tercedor 2005). Todos estos elementos integran el texto global y forman parte intrínseca del mismo (Göpferich 1995). En el corpus original se contabilizaron 35,40 imágenes por página y 46,58 imágenes por página en el corpus localizado. Este aparente alto número de imágenes se explica porque a menudo las imágenes se usan para decorar los sitios web u organizar la presentación como separadores etc., así como para incluir elementos de navegación, menús o anuncios web. Como ejemplo, la siguiente imagen del menú de navegación de [www.pandasoftware.es](http://www.pandasoftware.es) incluye un menú de navegación compuesto por 31 imágenes en formato .gif; cada opción cuenta con otra imagen que se superpone al pasar el cursor con el atributo *Onmouseover*.



Figura 83. Menú de navegación con texto codificado en imágenes. [www.pandasoftware.es](http://www.pandasoftware.es).

Al igual que los segmentos textuales incluidos en Flash, no hemos podido incluir el texto en imágenes en el cómputo total estadístico del corpus por ser imposible realizarlo con las herramientas utilizadas. La investigación de Bolaños *et al.* (2005) señala que el 50,32% de las imágenes analizadas en las páginas de inicio de sitios web corporativos incluyen texto al ser mayoritariamente menús de navegación o anuncios web (*banner ads*)<sup>420</sup>.

A pesar de que el texto desempeña el papel fundamental en los textos web (Nielsen y Tahir 2002; McAlpine 2001; Price y Price 2002), la imagen constituye un elemento esencial en la configuración textual y durante el proceso de localización el traductor ha de identificar la relación funcional existente imagen-texto y como influye la misma en la comprensión del texto global. En este doble eje cognitivo y pragmático, el traductor de estos textos se enfrenta a cuestiones como la identificación de la función de la imagen en el texto, su valor denotativo, su valor connotativo y cómo potenciar la accesibilidad a la imagen en entornos web. En este contexto se puede requerir que el traductor realice tareas de redacción técnica al crear o adaptar este texto alternativo para imágenes (Tercedor y Jiménez 2007).

Como reflexión final en este epígrafe, somos conscientes de una posible deficiencia en nuestro análisis del género al no haber analizado directamente las imágenes

---

<sup>420</sup> La investigación señala que el 47,74% de las imágenes son fotografías, el 15,13% son anuncios o banners, y el 5,26% son iconos.

recopiladas en el corpus. Otras investigaciones previas las incorporan dado su importancia (Bolaños 2002a; Bolaños *et al.* 2007 2005; Renau 2004). En nuestro caso, hemos incorporado los segmentos textuales en el atributo <Alt> que describen las imágenes y, en parte, defendemos que este aspecto quedaría compensado. En cualquier caso, ya hemos iniciado el análisis del conjunto de imágenes en el corpus desde una perspectiva traductológica (Tercedor y Jiménez 2007), y constituirá una de las futuras vías de investigación.

### 9.2.3. Segmentos textuales en etiquetas <meta>

Las etiquetas META constituyen una norma de base en toda página web, y contienen información metatextual sobre cada página web dirigida a (1) su descripción y clasificación por parte de motores de búsqueda e indexadores web de diversa naturaleza, (2) describir la codificación empleada, la lengua del texto, su creador, etc. Esta información permanece oculta al usuario final en la estructura de presentación. De entre los diferentes atributos posibles, nos es de interés desde la óptica del proceso de traducción el contenido de la etiqueta <meta NAME="DESCRIPTION" CONTENT="[texto a localizar]"> o <meta NAME="KEYWORDS" CONTENT= "[texto a localizar]">. En ellas se encuentra la descripción de la página web y la lista de palabras clave que resumen o describen su contenido. Desde el punto de vista terminológico, en los atributos meta *Keywords* se pueden observar relaciones ontológicas de interés con hiperónimos, hipónimos o sinónimos que facilitan la identificación y clasificación correcta. Dado su uso por los motores de búsqueda en la clasificación de las páginas, la calidad en su traducción es de vital importancia.

En los segmentos textuales en el atributo *Keywords* se recomienda el uso de 10 palabras básicas seguidas de otras 10 opcionales, con un máximo de 1000 caracteres, sin repetir muchas palabras o incluir una cantidad excesiva de palabras

puesto que en ambos casos los robots indexadores penalizan a la página por intentar manipularlos. Las palabras seleccionadas han de variar de página en página en un mismo sitio web para mejorar las posibilidades de que el sitio se encuentre (McAlpine 2001: 131), y han de colocarse en orden de relevancia dentro del sitio web. Todas las palabras han de separarse por comas y estar en minúscula, puesto que ciertos indexadores ignoran las palabras en mayúscula (Garrand 2001).

La importancia de la localización de las palabras clave reside en que asisten para que la versión localizada sea encontrada en la región o país de destino (Yunker 2003: 428). La labor del localizador puede incluir su creación, la selección de sinónimos, variaciones de cada palabra o incluso las palabras con las faltas de ortografía más prominentes ya que ciertos usuarios pueden intentar encontrar la palabra a partir de las faltas de ortografía específicas de cada lengua (ibid).

El valor del atributo *Description* se usa por parte de los motores de búsqueda como resumen del contenido de las páginas en los resultados, con lo que es importante tener una oración de calidad breve y concisa. Se recomienda que no supere los 200 caracteres.

Tipo	Corpus original 172 webs españoles		Corpus traducido 95 webs estadounidenses	
	Total	Media	Total	Media
Páginas web	<b>19.102</b>	<b>111,5</b> <b>Páginas/sitio</b>	<b>21.322</b>	<b>224,93</b> <b>Páginas/sitio</b>
Palabras en <Meta>	<b>932.255</b>	<b>48,80</b> <b>palabras</b>	<b>742.112</b>	<b>34,80</b> <b>Palabras/pag.</b>

**Tabla 8.** Estadísticas sobre palabras en segmentos textuales a traducir en META.

La comparación de las palabras incluidas en <meta> entre ambas secciones del corpus muestra una aparición de 48,80 palabras por página en el corpus original y 34,80 palabras por página en el localizado.

Puesto que estas palabras son invisibles al usuario final, esta diferencia podría apuntar a la menor consideración de estos elementos textuales debido a los costes adicionales por palabra en su localización. Hemos identificado que ciertos sitios web no localizan el contenido de estos atributos, como Dell, HP, Starbucks, Lexmark, o Sun<sup>421</sup>. Este hecho se debe en parte a que estos elementos textuales no aparecen en la presentación de la página web. En el caso del sitio web de Starbucks se encontraron dentro del contenido de los *Keywords* comentarios originales de los desarrolladores dirigidos a los localizadores:

```
<meta name="keywords" content="Starbucks, coffee, social responsibility, commitment to origins, conservation international., [enter market name], espresso, cafe', tea, food, beverages, barista, roasting"> [Starbucks].
```

Se observa además la posibilidad de que estos atributos no se localicen en unas páginas de un mismo sitio web<sup>422</sup>, como ocurre en el caso de American Express o EMC. Como ilustración presentamos el contenido de estos atributos en dos páginas de un mismo sitio, www.HP.es:

**Localizado:** <meta name="description" content="Recursos de impresión para artes gráficas de HP para diseño gráfico, fotografía y artes gráficas - Impresoras de diseño gráfico de HP"

**Sin localizar:** <meta name="description" content="The HP Graphic Arts Portal offers prepress professionals resources and guides to prepress and proofing printing product selection, technical support, and training and education resources.">

Por último, el atributo *Keywords* sólo está presente en el 33,99% de las páginas web originales y en el 63,25% de las páginas web localizadas, aunque como indicamos, cierto número de ellas se mantienen sin localizar. El atributo *Description* está presente en el 31,95% de las páginas web originales y en el

<sup>421</sup> Como ejemplo de Keywords sin localizar: de Lexmark ,<meta name="keywords" content="C770, C772 Intermediate Transfer Unit Maintenance Kit 40X1680 Lexmark">, de Sun, <meta name="keywords" content="sun microsystems, sun, java, java computing, solaris, sparc, unix, jini, computer systems, server, mission critical, RAS, high availability, cluster, workgroup server, desktop, workstation, storage, backup solutions">.

<sup>422</sup> Este aspecto entronca con la posición de Pym relativa a la diferencia entre traducción y localización: «software is localized but not entirely translated» (Pym 2004: 3).



50,66% de las localizadas. Se observa por tanto una presencia mayor en el corpus localizado, que asciende a casi el doble que en los originales.

#### **9.2.4. Los segmentos textuales en eventos *OnMouseover***

Los segmentos en *OnMouse* se basan, entre otros, en la posible aparición de un segmento textual en una ventana emergente cuando se pasa el puntero del ratón sobre una imagen o parte de la página web. El foco contextual de este tipo de textos suele ser expositivo-descriptivo, ya que describen el contenido de ciertas áreas de la página.

La inclusión de este elemento textual es facultativa. La media por página en el corpus original es de 35,51 palabras por página y de 50,43 palabras en el localizado. Esta diferencia puede apuntar a una mayor interactividad en el conjunto de los sitios web localizados.

#### **9.2.5. Los segmentos textuales en scripts, *<Select>* y *<Input>***

Los segmentos textuales que aparecen en *scripts* o eventos html se basan en la funcionalidad que presenta el nuevo medio. Estos segmentos textuales aparecen en determinadas situaciones que surgen por la interactividad del usuario con el texto, sobre todo como respuesta o clarificación a una acción que ha realizado el usuario. A menudo el traductor necesita recurrir a las posibles opciones que puedan aparecer en un momento determinado para determinar el contexto y la coherencia del segmento textual en un *script*.



Figura 84. Página protegida que indica un área reservada de [www.diode.es](http://www.diode.es).

La etiqueta `<Select>` constituye un evento html usado para la inclusión de menús de navegación o de opciones que se despliegan una vez que se activa un área específica. En general, se desaconseja su uso como método de navegación, puesto que ciertas opciones pueden quedar ocultas a priori al usuario. Así, estas permanecen ocultas a no ser que active este menú y, por lo tanto, éste puede no llegar a encontrar la información que necesite (Nielsen 2000). La brevedad en estos segmentos textuales es esencial.

**CONTACTO**

Asunto:

Nombre:

Compañía:

e-mail:

Teléfono:

Fax:

Comentarios:

Figura 85. Segmentos textuales en evento «*Select*».

La etiqueta `<Input>` se aconseja principalmente para su uso en formularios. Como se observa en la siguiente figura, el texto a traducir aparece en los atributos «name» y «value». Se emplea en selecciones con botones radio, en cajones de envío reinicio del formulario, así como en el proceso de acceso a áreas restringidas con nombre de usuario y clave. En el siguiente ejemplo se observa tanto la presentación como el código html de una etiqueta `<input>`:

Código html en <input>	Presentación en pantalla
<pre>&lt;input type="radio" name="sexo" value="hombre" checked id="sexoh"&gt;&amp;nbsp;&amp;nbsp; &lt;input type="radio" name="sexo" value="mujer" id="sexom"&gt;</pre>	<p>Sexo Hombre <input checked="" type="radio"/> Mujer <input type="radio"/></p>

Figura 86. Ejemplo de texto a localizar en etiqueta <input>.

El porcentaje de páginas web que incluyen la etiqueta **input** es similar en ambas secciones (32,72% de originales y 33,15% de localizadas). No obstante, las páginas originales muestran un mayor contenido global por página de palabras en las etiquetas <input> y <select>, 81,66 palabras por página frente a 42,11 palabras por página en las localizadas.

En lo que respecta a las palabras en *Scripts*, las páginas localizadas incluyen un mayor número de palabras, 27,23 por página de media frente a las 13,30 por página de las originales. La mayor inclusión de texto en scripts sería indicativa de una mayor complejidad e interactividad de las páginas web localizadas, ya que la inclusión de mayores niveles de interactividad requiere competencias en programación más avanzadas.

## 10. ANÁLISIS CONTRASTIVO DE FORMAS LINGÜÍSTICAS CONVENCIONALES EN BLOQUES CONTACTO (B) Y LEGAL (D)

Siguiendo la investigación de Silvia Gamero, definimos las formas lingüísticas convencionales como las unidades fraseológicas cuya función es verbalizar unidades complejas de sentido recurrentes y que el redactor textual tiene a su disposición en función del género en cuestión (Gläser 1997: 90, cf. Gamero 2001: 167). Desde la perspectiva de la didáctica de la traducción, Hurtado Albir (2001: 68) apunta que la descripción de estas formas convencionales constituye uno de los principales objetivos de las investigaciones contrastivas sobre géneros. En el capítulo anterior realizamos un análisis contrastivo de la terminología en menús de navegación, que se centró en estas formas lingüísticas convencionales asociadas a la superestructura textual. No obstante, la presencia de formas convencionales recurrentes no se circunscribe a estos subtextos, sino que éstas se encuentran asociadas al conjunto de la superestructura y predominan en bloques o secciones determinados (Gamero 2001: 167). El grado de lexicalización o rigidez de las formas lingüísticas convencionales que nos ocupan varía en función de la unidad de la que se trate<sup>423</sup>. Dentro de la fraseología existen unidades fraseológicas de mayor fijación como dichos, refranes y modismos, junto con un segundo grupo que engloba a unidades con menor grado de fijación y en las que se refleja el significado original de los elementos integrantes, como las colocaciones<sup>424</sup> (Tercedor 1999). Nuestro análisis de formas lingüísticas convencionales propias de

---

<sup>423</sup> Este grado de variación se observa en la definición del objeto de estudio de la fraseología ofrecido por Tercedor (1999: 83): «una expresión de más de una palabra, con mayor o menor grado de lexicalización».

<sup>424</sup> Las colocaciones se definen como: «the tendency of certain words to co-occur regularly in a given language» (Baker 1992: 47). Tercedor (1999: 88-96) ofrece una detallada revisión bibliográfica de las distintas definiciones de este concepto.

los sitios web corporativos las sitúa dentro del segundo grupo de diverso grado de variación. Por consiguiente, pueden admitir una amplia gama de variantes, si bien existe la posibilidad de que identifiquemos ciertas formas altamente convencionalizadas.

En este sentido, el análisis de las formas lingüísticas convencionales en menús de navegación ya mostró que, al tratarse de textos semiespecializados, el grado de variación fraseológica en los textos originales es medio<sup>425</sup>. El grado de fijación fraseológica dentro de los sitios web corporativos depende igualmente del bloque o sección comunicativa del que se trate, ya que ciertos bloques como **noticias** (D) o **productos y servicios** (E), pueden mostrar un menor nivel de fijación intra e inter lingüística al ser ejemplos de *e-textos*.

### **10.1. Metodología de análisis de formas lingüísticas convencionales**

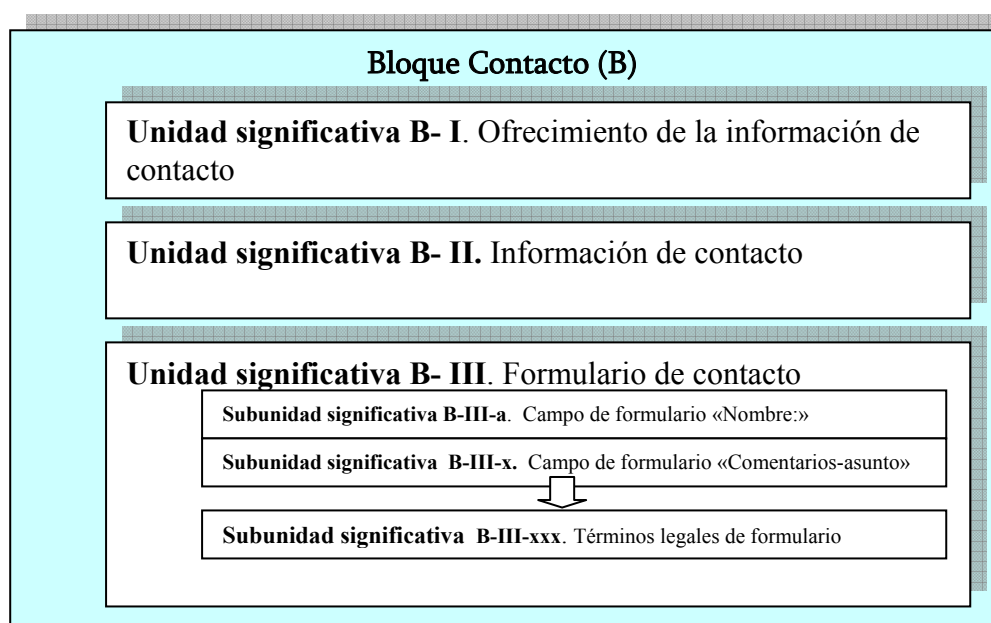
Dada la gran extensión de nuestro corpus comparable, el análisis adicional de formas lingüísticas convencionales se circunscribió a dos secciones que seleccionamos por su relevancia contrastiva. En primer lugar estudiaremos a fondo el bloque **contacto** (B), por presentar la mayor frecuencia de aparición en ambas secciones del corpus. En segundo lugar hemos seleccionado el bloque comunicativo **legal** (F), que presenta el mayor grado de divergencia en su frecuencia de aparición entre ambos perfiles textuales (véase 9.1). Además, asumimos que este bloque mostrará un alto grado de convencionalización en sus estructuras lingüísticas al tratarse de textos jurídicos<sup>426</sup> (Borja Albí 1998, 2000, 2003). Ambos bloques se eligieron en virtud de su carácter convencionalizado, por lo que incluirán segmentos lógicos y recurrentes de información y presentarán formas lingüísticas convencionales asociadas a los mismos. Siguiendo a Gamero

---

<sup>425</sup> No obstante, defendemos que contrastivamente los textos anglosajones presentan una mayor fijación terminológica en los menús de navegación (Nielsen y Tahir 2002; Nielsen 2006).

<sup>426</sup> De hecho, Borja Albí argumenta que en los textos jurídicos es donde mejor se puede apreciar la importancia de las convenciones de género (Borja Albí 1998).

(2001: 169), denominaremos **unidades significativas** a estos segmentos lógicos de información constitutivos de unidades de sentido. Estas unidades significativas se colocan en un *continuum* estructural tras los bloques y secciones y poseen, en igual medida, una función secundaria propia que contribuye al propósito retórico global del texto en el que se imbrica. Las unidades significativas presentan una configuración modular interna que se formula mediante una estructura modular y flexible. Su orden jerárquico interno puede regularse por una organización lógica de presentación; por ejemplo, y como se verá posteriormente, el bloque de contacto (B) presenta una unidad significativa cuyo objetivo es ofrecer al usuario la información de contacto (B-I), y ésta lógicamente ha de anteceder a su propia exposición (B-II).



**Figura 87.** Relación entre sección, unidades significativas y subunidades significativas en bloque contacto (B).

En un último plano jerárquico, algunas unidades significativas pueden subdividirse en **subunidades significativas** (ibid). En nuestro caso, los diferentes campos en los formularios de contacto constituyen subunidades significativas integrantes de la unidad significativa «formulario de contacto» (B-III). La Figura 87 ilustra esta

estructura jerárquica del bloque **contacto** (B) que, tras nuestro análisis, podemos concluir que se compone de secciones, unidades significativas y subunidades significativas.

## **10.2. Análisis adicional del bloque «contacto» (B)**

Con el fin de analizar las formas lingüísticas convencionales asociadas a este bloque comunicativo recopilamos, a partir del corpus primario, un subcorpus comparable compuesto por las páginas de contacto. Al igual que el análisis por textos globales, ambas secciones del subcorpus se convirtieron a formato texto y a codificación ASCII mediante la herramienta informática Detagger. Como resultado, en esta fase elaboramos dos subcorpus comparables similares pero en distinta codificación, (1) html y (2) texto. En este último excluimos los elementos textuales en el encabezamiento de las páginas web, <Head> </Head>, que en virtud de su distribución irregular y su carácter metatextual podrían sesgar nuestro análisis léxico posterior.

Tras analizar manualmente la estructura de este bloque en el subcorpus comparable, constatamos que el perfil textual de este bloque se compone de tres unidades significativas. Con afán identificativo hemos asociado cada unidad significativa a un número romano consecutivo para distinguirlas de las secciones comunicativas. En primer lugar, el perfil textual presenta una unidad por la que se ofrece la información para contactar con la empresa (B-I), y que posee una función subsidiaria exhortativa con alternativa. En un segundo lugar figura la unidad significativa que presenta la información de contacto en sí, a través de de la dirección, teléfono, fax, correo electrónico (B-II), cuya función subsidiaria es expositiva en descripción. En esta unidad pueden aparecer diferentes subunidades según las personas o departamentos responsables de cada petición. En un tercer lugar se encuentra la unidad significativa «formulario de contacto interactivo» (B-

III). Esta unidad resulta menos frecuente que las anteriores, con una frecuencia de aparición en los textos originales de 31.42% y 26.31% en localizados. A su vez, esta unidad significativa presenta un conjunto de subunidades significativas que hemos identificado con los distintos campos que articulan un formulario, como el campo que requiere la identificación del usuario por su nombre (B-III-a), o el campo por el que se incluyen los comentarios o asuntos en el formulario (B-III-g).

Como paso posterior a este análisis superestructural, realizamos un análisis contrastivo léxico dirigido a identificar formas lingüísticas convencionales con la herramienta informática WordSmith Tools. Procedimos a obtener los listados de frecuencia por palabras con el objetivo de caracterizar el contenido proposicional de estos subtextos, para así identificar posibles divergencias que apunten a diferencias de contenido:

Listado de palabras en subcorpus original			Listado de palabras en subcorpus localizado		
1	Contacto	1.05%	1	Información	0,74%
2	Fax	0,69%	2	Contacto	0,66%
3	Información	0,65%	3	HP	0,61%
4	es	0,59%	4	Productos	0,59%
5	Teléfono	0,55%	5	España	0,49%
6	Empresa	0,54%	6	Soporte	0,45%
7	Formulario	0,53%	7	Servicios	0,39%
8	Datos	0,50%	8	Teléfono	0,38%
9	Mail	0,50%	9	es	0,34%
10	Dirección	0,48%	10	Correo	0,33%
11	Servicios	0,47%	11	Fax	0,33%
12	Nombre	0,42%	12	Oracle	0,30%

**Tabla 24. Listado de palabras en subcorpus de sección contacto (B).**

El análisis del listado revela que las palabras más frecuentes destacadas en ambas secciones del corpus, y por tanto indicativas de los contenidos proposicionales de estos subtextos, son «contacto», «información» «productos», «teléfono», «fax» y «servicios». Las tres primeras se encuentran relacionadas con la unidad significativa B-I, mientras que «teléfono» o «fax» se encontrarían asociadas con la dos siguiente unidades significativas de información de contacto (B-II), o formulario de contacto (B-III). El contexto geográfico se encuentra representado en ambas secciones del



corpus tanto por el truncamiento «es», representativo del dominio de sitios web o direcciones de correo electrónico, como por «España», que con una mayor frecuencia en el corpus localizado atestigua el carácter de localización de estos subtextos. Nos es de interés la alta frecuencia del término «datos» en el subcorpus original (0,58%), un sinónimo de «información» en este caso que no figura con tal frecuencia en el subcorpus localizado (0,14%). Por último, la unidad léxica para referirse a la dirección de correo electrónico, asociada con las unidades significativas B-II y B-III, se encuentra representada por «mail» en el corpus original y por «correo» en el localizado. Este hecho puede apuntar al carácter de traducción de los textos localizados, puesto que la forma más convencional en los textos originales es el anglicismo puro «e-mail»<sup>427</sup>, si bien es incorrecta a la luz del DRAE (edición XXII).

### 10.2.1. Análisis de la unidad significativa de ofrecimiento de información de contacto (B-I)

Las dos palabras con mayor frecuencia en ambas secciones del corpus, «información» y «contacto», se encuentran mayoritariamente relacionadas con la primera de las unidades significativa identificadas: «el ofrecimiento de la información de contacto» (B-I). En primer lugar, y con el objetivo de analizar el radio colocacional de la palabra «información», realizamos un análisis fraseológico usándola como base de sus colocaciones. A continuación presentamos un extracto de las concordancias obtenidas en el subcorpus original de contacto:

1. cernos más a fondo y recibir más **información** remitanos un e-mail
2. ed quiere más **información** acerca de nuestros pr
3. te@rianxeira.com \*Solicitud de **información** / Sugerencias\* Esc
4. tro: Mensaje: Deseo recibir **información** de sus actividades ¿
5. " "servicios de la sociedad de la **información** y de comercio electró
6. onsuming (Worldsites network) **Información** Legal - Protección de
7. este formulario si desea obtener **información** sobre nuestra oferta d

**Figura 88. Extracto de concordancias de «información» en subcorpus de contacto original.**

---

<sup>427</sup> El lexema «mail» muestra una frecuencia de 0,22% en el subcorpus localizado, inferior a la mitad que la frecuencia de este mismo lexema en el subcorpus original (0,55%).

El siguiente extracto pertenece al análisis del subcorpus de contacto localizado:

1. io: \*Formulario de solicitud de **información**\* Los campos de l
2. Servicios de la Sociedad de la **Información** y de Comercio Elect
- 3 lectrónico.\* \*Incluya más **información** a continuación.\* V
4. 0800 / 962 116 Si desea más **información** comercial o técnica,
5. versity \*Soporte \* Para más **información** sobre Oracle Suppor
6. o Por e-mail No quiero recibir **información** de Oracle \* \* \*
7. cto con CA\* Si desea solicitar **información** sobre los productos
8. y servicios. Si desea recibir tal **información**, por favor, marque e

**Figura 89. Extracto de concordancias de la palabra «información» en subcorpus de contacto localizado.**

Dada la gran cantidad de líneas de concordancia obtenidas, hemos resumido en la siguiente tabla las formas más recurrentes a la derecha y a la izquierda con ordenadas por su frecuencia, que hemos indicado entre paréntesis. Esta síntesis en su presentación facilitará nuestro estudio contrastivo:

Subcorpus original			Subcorpus localizado		
Izquierda	Pal. clave	Derecha	Izquierda	Pal. clave	Derecha
más [6] solicitud de [3] deseo recibir [2] [nombre de ley] [1] incluya más [1]	<b>información</b>	sobre nuestro- a...[5] comercial [4] general [2] financiera [1] corporativa [1]	más [5] [no] Deseo recibir [3] [nombre de ley] [3] Si desea más (2) solicitud de (2) recibir más... (2) no quiero recibir (2) para encontrar (2) para (2)	<b>información</b>	sobre nuestro-a (8) acerca de...(5) general (5) de contacto (4) importante (4) personal (3) adicional (2) comercial (2) financiera (1) corporativa (1)

**Tabla 25 . Colocaciones de la palabra «información» en subcorpus de bloque contacto.**

En el análisis excluimos aquellos elementos que concuerdan en una sola ocasión en uno u otro subcorpus al no representar grado de convencionalización alguno<sup>428</sup>. El análisis revela que el radio colocacional de la palabra «información» es mayor en los textos localizados, con nueve posibles colocadores o elementos colocacionales a la derecha y diez elementos a la izquierda de la palabra clave. De los colocadores

<sup>428</sup> En línea con la propuesta de Sinclair, defendemos que cualquier colocación repetida es significativa aunque su frecuencia sea muy baja « [...] many typical collocations are of such low frequency compared with the overall length of a text. Because of the low frequency of the vast majority of words, almost any repeated collocation is a most unlikely event. » (Sinclair 1991: 116) .

encontrados a la derecha, seis están formados por una forma verbal con preposición, dos están formados por una preposición, y uno constituye una unidad léxica fija al ser el nombre de una ley española<sup>429</sup>. De los elementos colocacionales a la derecha del núcleo, seis son adjetivos y tres lo constituyen compuestos sintagmáticos con las preposiciones «sobre», «acerca de» y «de». De entre estas preposiciones, en el corpus original no apareció ningún caso en el que «acerca» se combinara con «información», ya que en éste se usa la preposición «sobre». Esta asimetría se mostró igualmente en el análisis de la terminología de los menús de navegación de la sección **empresa** (C) (véase 9.1.2.3.1).

Los adjetivos que se presentan a la derecha de la base son «importante», «adicional» y «personal». Este último se presenta exclusivamente en el subcorpus original en conjunción con el sustantivo «datos» y, en este sentido, en el subcorpus original es inexistente la unidad fraseológica «información personal»<sup>430</sup> que aparece en los textos localizados. Por el contrario, los textos originales presentan un espectro colocacional menor; cuatro elementos colocacionales a la derecha del núcleo y cinco a la izquierda, estando todos ellos presentes en el subcorpus localizado.

En segundo lugar, la unidad significativa que representa el ofrecimiento de la información de contacto se podría analizar a través de las posibles formas nominales y verbales del lema [contact\*]. Con este objetivo realizamos un análisis de co-ocurrencia léxica organizado por grupos de dos palabras, cuyos resultados hemos incluido en el Apéndice 14.4 (p. 638). Los grupos de mayor frecuencia son (1) «de contacto», (2) «en contacto», y (3) «contactar con». El primer caso constituye el grupo de mayor frecuencia y revela la combinación habitual de unidades sintagmáticas nominales mediante los sustantivos «datos», «información» y «formulario».

---

<sup>429</sup> Ley 34/2002, de 11 de julio, de Servicios de la Sociedad de la Información y de Comercio Electrónico.

<sup>430</sup> Apuntaría a su carácter como calco léxico del término inglés *personal information*.

EN CONTACTO	
Subcorpus original	Subcorpus localizado
<p>-Si desea <b>ponerse en contacto</b> con nosotros [3]                      -No <b>duden en ponerse en contacto</b> con nosotros para...[2]                      -Rellenar los datos de este formulario y <b>nuestro departamento comercial se pondrá en contacto</b> con usted                      -Escribenos [...] y <b>nos pondremos en contacto</b> en seguida.                      -<b>Nos pondremos en contacto</b> con usted a la mayor brevedad posible                      -<b>Se pueden poner en contacto</b> con nosotros a través del teléfono                      -<b>Puede ponerse en contacto</b> con nosotros llamando al número de teléfono...                      -Rellene el formulario <b>para ponerse en contacto</b> con nosotros                      -Si lo desea puede rellenar este formulario <b>para ponerse en contacto</b> con nosotros                      -<b>Para ponerse en contacto</b> con nosotros,...                      -Cubra este formulario <b>si desea ponerse en contacto</b> con nosotros                      -<b>Si quieres ponerte en contacto</b> con nosotros rellena el siguiente                      -<b>Indique cómo quiere que nos pongamos en contacto</b> con usted                      -<b>Póngase en contacto</b> con nosotros para realizar...                      -<b>Ponte en contacto</b> con nosotros a través del correo electrónico...                      -Para cualquier tipo de consulta o duda <b>póngase en contacto</b> con nosotros a través de este formulario...                      -Si desea más información [...] <b>póngase en contacto</b> con la dirección de correo electrónico</p>	<p>-<b>Póngase en contacto</b> con nosotros [13]                      -Por favor, <b>póngase en contacto</b> con las siguientes personas:                      -<b>Póngase en contacto</b> con el servicio de atención al cliente                      -<b>Ponte en contacto</b> con nosotros a través del formulario                      -Un representante [...] y <b>se pondrá en contacto</b> con usted                      -Díganos en qué podemos ayudarle y en breve, <b>nos pondremos en contacto</b> con usted                      -Información sobre <b>cómo ponerse en contacto</b> por teléfono o correo electrónico                      -<b>No dude en ponerse en contacto</b> con el equipo de Ernst &amp; Young Online</p>

Tabla 26. Formas lingüísticas convencionales asociadas con la colocación «en contacto».

El siguiente grupo de interés seleccionado para analizar las formas lingüísticas convencionales fue «en contacto». Este grupo aparece en el séptimo lugar del listado del subcorpus original y el vigésimo en el localizado. La unidad fraseológica «en contacto» se encuentra asociada exclusivamente a la más amplia «ponerse en contacto» en nuestro subcorpus. Esta forma lingüística es por tanto convencional y se encuentra asociada a la unidad significativa de ofrecimiento de la información de contacto (B-I). Al igual que en el caso anterior, y con el objetivo de facilitar su estudio contrastivo, hemos recogido en la Tabla 26 las posibles patrones de colocación identificados. En la misma hemos identificado con un número las posibles colocaciones repetidas.

El análisis contrastivo revela una diferencia primordial en la forma verbal por la que se expresa este acto de habla directivo. Como demuestran los datos presentados en la tabla anterior, los formularios originales españoles poseen una mayor cantidad de recursos léxico-gramaticales disponibles para formular este acto directivo que los localizados. Este aspecto está en consonancia con los resultados en el estudio contrastivo español-alemán del género formulario recogidos por Sánchez Nieto (2006: 127). En el corpus localizado es mayoritario el uso de formas verbales en modo imperativo para la expresión de la función exhortativa o apelativa de esta unidad significativa. Así, dieciséis de las líneas de concordancias obtenidas en el corpus localizado figuran en modo imperativo, mayoritariamente «póngase en contacto», frente a sólo cuatro en el original. Este predominio del modo imperativo en las formas verbales queda patente en el siguiente extracto de colocaciones localizadas:

N	Concordance
5	e expresa más arriba, me <b>pondré</b> en contacto con Alcoa para infor
6	tos autorizados deberían <b>ponerse</b> en contacto con los equipos de
7	plataforma, no dude en <b>ponerse</b> en contacto con el equipo de Ern
8	Información sobre cómo <b>ponerse</b> en contacto por teléfono o correo
9	Información de Interés <b>Ponte</b> en contacto con nosotros a través
10	Información de Interés <b>Ponte</b> en contacto con nosotros a través
11	los negocios, por favor <b>póngase</b> en contacto con las siguientes pe
12	<b>póngase</b> en contacto con nosotros idio
13	Conecte con Avaya <b>Póngase</b> en contacto con Avaya: 91 387 6
14	rt Services Directory <b>*Póngase</b> en contacto con CA* Si desea
15	y cuestiones técnicas. <b>Póngase</b> en contacto con DELL =====
16	istema que necesita? <b>Póngase</b> en contacto con DELL =====
17	de compra y entregas. <b>Póngase</b> en contacto con el servicio de ate
18	itio España - español » <b>Póngase</b> en contacto con HP
19	ol » <b>Póngase</b> en contacto con HP Inicio > PÓN
20	r Página imprimible <b>Póngase</b> en contacto con Network Applianc
21	ficinas de la empresa* <b>Póngase</b> en contacto con Network Applianc
22	ntacto con HP Inicio > <b>Póngase</b> en contacto con nosotros PÓng
23	ontacto con nosotros <b>Póngase</b> en contacto con nosotros *****
24	HP Inicio > <b>Póngase</b> en contacto con nosotros PÓng
25	ontacto con nosotros <b>Póngase</b> en contacto con nosotros *****
26	----- <b>Póngase</b> en contacto con nosotros -----

**Figura 90. Extracto de concordancias de «en contacto» en subcorpus localizado de contacto con predominio del modo imperativo.**

En el subcorpus original la forma verbal más frecuente asociada con esta colocación es el infinitivo con función de imperativo, doce casos frente a dos en el localizado, que se corresponden con una mayor modulación por la que se expresa

esta función exhortativa con opción mediante perífrasis verbales: «si desea ponerse en contacto», «puede ponerse en contacto» o «no dude en ponerse en contacto». Aparecen igualmente tres formas verbales en el tiempo futuro en el original y dos en el localizado, tanto en la tercera persona del singular como la primera del plural. Éstas se corresponden con la futura acción por parte de la empresa contingente con el inicio del intercambio comunicativo por parte del usuario. Finalmente se encontró una forma verbal en subjuntivo «indique cómo quiere que **nos pongamos en contacto** con usted». Las siguientes líneas de concordancia pertenecen al subcorpus original:

N	Concordance
5	ulario. "Para cualquier otro tema, se pueden <b>poner en contacto</b> con nosotros a través" del teléfono *902
6	na impresora laser Hp o un plotter Hp puede <b>ponerse en contacto</b> con nosotros llamando al número de tel
7	lo desea puede rellenar este formulario para <b>ponerse en contacto</b> con nosotros. En el plazo de tiempo
8	CR Class English Version "Si desea <b>ponerse en contacto</b> con nosotros o realizar algún pedido, pu
9	*CONTACTAR* Rellene el formulario para <b>ponerse en contacto</b> con nosotros! Nombre: Apellidos:
10	*Clientes* Inicio > Contacto > "Si desea <b>ponerse en contacto</b> con nosotros, para hacernos llegar cualq
11	uiera de "nuestras actividades, no dude en <b>ponerse en contacto</b> con nosotros en la" siguiente dirección
12	para la agricultura. Información Si desea <b>ponerse en contacto</b> con nosotros puede rellenar el siguiente
13	on alguna Delegación Comercial?... "Para <b>ponerse en contacto</b> con nosotros, puede utilizar nuestro dire
14	del site   Política de privacidad No dude en <b>ponerse en contacto</b> con nosotros para pedir presupuesto o "
15	sa.com *»* Cubra este formulario si desea <b>ponerse en contacto</b> con Reflexion Arts : Nombre: Cargo:
16	05 Jazztel - Jazz Telecom, S.A." Si desea <b>ponerse en contacto</b> conmigo o solicitar un presupuesto sin
17	ORES _CONTACTO ***** Si quieres <b>ponerte en contacto</b> con nosotros rellena el siguiente formula
18	ORES _CONTACTO ***** Si quieres <b>ponerte en contacto</b> con nosotros rellena el siguiente formula
19	Apellidos:* *Indique cómo quiere que nos <b>pongamos en contacto</b> con usted:* *Por teléfono * *Por e-ma
20	edicaciones Investigación Referencias <b>Pongase en contacto</b> con nosotros para realizar cualquier con
21	Preguntas más frecuentes o en el sitio Web, <b>ponte en contacto</b> con nosotros por correo electrónico en l
22	ioj Para cualquier tipo de consulta o duda <b>póngase en contacto</b> con nosotros a través de este formulario
23	re nuestra empresa o nuestros productos, <b>póngase en contacto</b> con la dirección de correo electrónico "

**Figura 91. Extracto de concordancias de «en contacto» en subcorpus original de contacto con predominio de formas de cortesía.**

Como resumen, hemos observado que el contraste entre ambos subcorpus se centra en el predominio casi exclusivo de formas verbales en modo imperativo en la formulación de esta unidad compleja de sentido en el corpus localizado frente a una mayor variación y modulación en el ruego en los originales.

Si pasamos al tercer caso, «contactar con», esta colocación aparece en el puesto décimo en la lista realizada de co-ocurrencia en el subcorpus original y el decimosegundo en el localizado. El análisis de esta colocación, siempre dentro de

la unidad significativa que nos ocupa, también revela claras diferencias entre ambas secciones del corpus. Al igual que la anterior, esta colocación está asociada al acto directivo por el que la empresa comunica las instrucciones necesarias al receptor para que actúe.

CONTACTAR CON	
Subcorpus original	Subcorpus localizado
<ul style="list-style-type: none"> <li>-¿<b>Necesita</b> contactar con alguna Delegación comercial?</li> <li>-<b>Cómo</b> contactar con Caditec Testing:</li> <li>-<b>Desea</b> contactar con:</li> <li>-Si <b>desea</b> contactar con nosotros puede hacerlo rellenando el siguiente formulario</li> <li>-Si <b>desea</b> contactar con nuestro departamento de Recursos Humanos pulse aquí</li> <li>-Si <b>desea</b> contactar con nosotros, rellene el siguiente formulario</li> <li>-<b>Para</b> contactar con nosotros por Email...</li> <li>-Rellena el siguiente formulario <b>para</b> contactar con nosotros</li> <li>-<b>Puede</b> contactar con nosotros llamando al teléfono...</li> <li>-Con este formulario <b>puede</b> contactar con nuestro equipo</li> <li>-¿Cómo <b>puedes</b> contactar con nosotros?</li> <li>-Si <b>quieres</b> contactar con nosotros, rellena y envía el siguiente formulario</li> <li>-Si <b>quiere</b> más información [...] o <b>quiere</b> contactar con nosotros, por favor rellene el siguiente formulario</li> </ul>	<ul style="list-style-type: none"> <li>-<b>Para</b> contactar con nuestro Servicio de Atención Técnica</li> <li>-<u>Complete</u> el siguiente formulario <b>para</b> contactar con nosotros</li> <li>-Por favor, rellene <u>los campos de abajo</u> <b>para</b> contactar con STANLEY ESPAÑA</li> <li>-<b>Para</b> contactar con <u>Comercial de Soporte</u></li> <li>-<b>Puede</b> contactar con nosotros <u>con los siguientes</u> medios de comunicación</li> </ul>

**Figura 92. Formas lingüísticas convencionales asociadas con la colocación «contactar con».**

El subcorpus localizado muestra una asociación casi exclusiva de esta colocación con la preposición «para» seguida de una forma verbal infinitiva, «para contactar con [...]», toda vez que las formas observadas en el subcorpus original presentan una mayor variación. No obstante, es mayoritaria la modulación del acto directivo a través de expresiones condicionales como «si desea» o «si quiere», o a través de perífrasis verbales con los verbos «poder» o «necesitar». También identificamos preguntas dirigidas al receptor, un hecho que entronca con el carácter dialógico que caracteriza los distintos subgéneros de formularios y que se expresa interactivamente entre dos interlocutores que van alterando los turnos de

preguntas y respuestas (Gülich 1981; Sánchez Nieto 2006). Desde un punto de vista estilístico, la expresión de este acto de habla como pregunta o como ofrecimiento en forma condicional la dota de un mayor grado de formalidad, y entroncaría con la recomendación de redactar el texto web desde la perspectiva del usuario (Price y Price 2002; Bly 2002).

Como última reseña, en el análisis de concordancias observamos una gran cantidad de errores de traducción en los segmentos que incluyen la colocación de «contactar con», cuatro de los cinco segmentos presentados y que hemos subrayado en la tabla anterior. Estas inadecuaciones son sobre todo de carácter léxico al usar anglicismos como «complete el formulario», «rellene los campos de abajo» o «contactar con Comercial de Soporte», así como inadecuaciones sintácticas en la formulación como «contactar con nosotros **con** los siguientes medios».

### **10.2.2. Análisis de la unidad significativa formulario de contacto (B-III)**

Esta unidad significativa es de gran interés en nuestro estudio, ya que junto con los menús de navegación, los formularios suelen presentar los mayores niveles de convencionalización y se prestan, por ende, al análisis contrastivo de las formas lingüísticas convencionales asociadas a la superestructura. Otra razón por la que los consideramos de interés recae en que este género textual ha sido objeto de investigaciones desde la Traductología y la Lingüística, especialmente en el ámbito germánico. Primordialmente, los estudios se han centrado en subgéneros como formularios de inscripción o formularios que sirven como medio de comunicación entre ciudadanos y autoridades o instituciones (Grosse y Mentrup 1980; Gülich 1981; Nord 2001: 22-29; Fernández Sánchez 2004: 409-429; Sánchez Nieto 2006a, 2006b). La existencia de estas investigaciones contrastivas previas supone una base para realizar el análisis que desarrollaremos a continuación.



El carácter altamente convencional de los formularios se observa en la definición del género formulario de Gülich (1981: 329), que los define como «un instrumento comunicativo altamente estandarizado propio del trato entre la administración y el ciudadano»<sup>431</sup>. En el caso de esta unidad significativa, la comunicación se entabla entre la empresa y cualquier usuario con independencia de la relación entre ambos. Éstos pueden poseer diversos motivos para entablar la comunicación, si bien los posibles motivos están limitados a la posible relación lógica entre ambos.

La equiparación de dicho género con la unidad significativa B-III que nos ocupa es plausible ya que ésta última participa de la totalidad de los rasgos básicos del primero: ambos constituyen un instrumento comunicativo por el que el emisor obtiene una serie de datos determinados sobre el receptor con el objetivo de asistirle con la mayor brevedad posible, que se lleva a cabo mediante una serie de preguntas previamente establecidas en la elaboración de la macroestructura textual (Gülich 1981; Sánchez Nieto 2006). Este carácter dialógico propio de los formularios es compartido por la unidad significativa «formulario de contacto», en cuanto que el receptor se identifica y lleva a cabo las acciones comunicativas por medio de las preguntas e instrucciones emitidas por el emisor. No obstante, éstas no se presentan por medio de frases interrogativas sino que formulan por medio de palabras clave seguidas de espacios en blanco adjuntos (Gülich 1981: 335-336). En el caso de formularios web, este intercambio comunicativo finaliza con una subunidad significativa de interacción que valida el contenido y le confiere el carácter de texto completo como tal, ya que una vez cumplimentado está en disposición de cumplir sus objetivos comunicativos (Sánchez Nieto 2006).

Hemos denominado a esta unidad significativa que nos ocupa «formulario de contacto», puesto que este término constituye el compuesto sintagmático de mayor

---

<sup>431</sup> Traducción de Fernández Sánchez (2004: 410).

frecuencia en el subcorpus original. No obstante, aparecen otras siete posibles denominaciones para tal formulario, entre las que incluimos una aparición de la forma sinonímica «cuestionario»<sup>432</sup>. En la siguiente tabla recogemos las unidades léxicas que incorporan esta palabra recogidas en nuestro subcorpus:

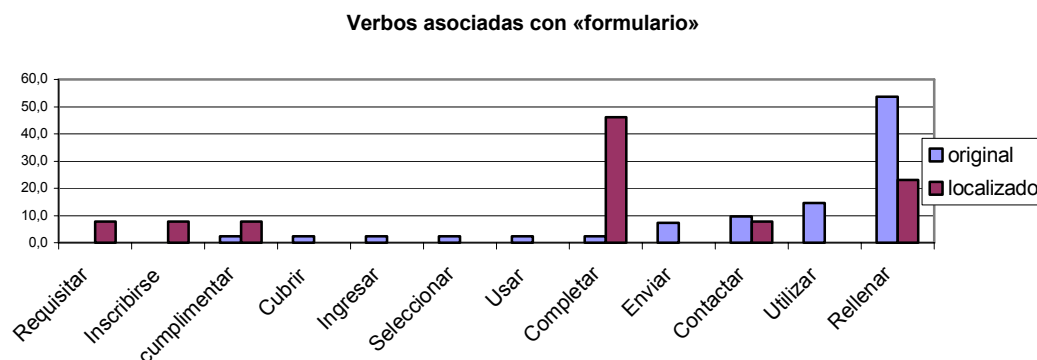
Subcorpus original		Subcorpus localizado	
Pal. clave	Derecha	Pal. clave	Derecha
Formulario	de contacto [7] de alta [3] de consultas técnicas [2] de información comercial [2] de consultas [1] de marketing-comercial [1] de atención al cliente [1]	Formulario	de contacto [2] de solicitud de información [2] de solicitud [1] de comentarios [1] de atención al cliente [1]

**Figura 93. Colocaciones de la palabra clave «formulario».**

Es significativo el carácter de «consulta» de los formularios originales y de «solicitud de información» en los formularios localizados. En el subcorpus original aparece con una frecuencia significativa la unidad fraseológica «formulario de alta», ya que el motivo principal por el que inician la acción comunicativa los clientes de empresas proveedoras de servicios de telefonía o Internet es contratar estos servicios. Este caso se correspondería con un subgénero de formulario diferenciado, ya que su cumplimentación y envío supone la aceptación de unos términos legales y el establecimiento de un contrato jurídicamente vinculante entre ambas partes.

En segundo lugar, y puesto que la presentación del formulario supone la inclusión de actos de habla directivos o instructivos que incorporan esta palabra, analizamos las formas verbales asociadas con la palabra clave «formulario». Mostramos los resultados en la siguiente gráfica.

<sup>432</sup> Este sinónimo refleja el carácter dialógico de preguntas y respuestas. El sinónimo «cuestionario» es inexistente en el subcorpus localizado.



**Figura 94. Verbos asociados a la palabra «formulario» en subcorpus comparable de contacto.**

El análisis contrastivo de estas concordancias muestra una clara preferencia por el verbo «rellenar» en los textos originales y una frecuencia similar de «completar» en los textos localizados. Esta colocación está mínimamente representada en el corpus original, con lo que se confirma su carácter de anglicismo léxico. En el subcorpus original se encontraron seis verbos de distinta frecuencia inexistentes en el subcorpus localizado «utilizar», «enviar», «usar», «seleccionar», «ingresar» y «cubrir», aunque este último supondría un galleguismo léxico. De las formas verbales localizadas, constituirían errores de traducción «inscríbese en el formulario» y «requisite el formulario».

En lo que respecta a la macroestructura de esta unidad significativa (B-III), se han identificado nueve subunidades significativas que se corresponden con los diversos datos que la empresa recopila sobre el usuario: el nombre (B-III-a), la información personal adicional (B-III-b), la información de la empresa (B-III-c), dirección (B-III-d), medio de contacto (B-III-e), motivo del contacto (B-III-f), comentarios (B-III-g), cómo llegó al formulario (B-III-h), e interactividad del formulario (B-III-i). A su vez, dentro de las diferentes subunidades hemos identificado 27 segmentos textuales recurrentes que se corresponden con los distintos campos a rellenar por parte del usuario, y que se expresan mayoritariamente a través de formas

lingüísticas convencionales. Dado su valor convencional, éstos se convierten en marcadores funcionales del género en cuestión (Göpferich 1995: 305).

En la siguiente tabla están representadas (1) las distintas subunidades significativas que conforman esta unidad significativa, (2) su grado de convencionalización en ambos subcorpus, y finalmente, (3) las últimas tres columnas incluyen el análisis contrastivo de la frecuencia de uso de las unidades léxicas en las que se expresan estas subunidades y sus segmentos recurrentes.

Subunidad significativa en unidad significativa B-III, Formulario de contacto	Original	Localizado	Unidad léxica	% Original	% Localizado
A. Nombre	100%	100%	Nombre Nombre y apellidos Apellidos Apellido Primer Apellido Segundo Apellido	74,42% 16,28% 18,60% 0% 2,33% 2,33%	90,48% 9,52% 28,57% 33,33% 4,76% 4,76%
1. Iniciales	0%	9,52%	inicial segundo nombre iniciales	0% 0%	4,76% 4,76%
2. Tratamiento	0%	19,05%	Tratamiento Saludo Sr. Sra. Srta. Dr. Dra Título Sr. Sra. Srta.	0% 0% 0% 0%	4,76% 4,76% 4,76% 4,76%
B Información personal adicional					
1. DNI -CIF	2,33%	4,76%	DNI CIF	100% 0%	0% 100%
2. Edad	2,33%	0%	Edad	100%	0%
3. Sexo	2,33%	4,76%	Sexo	100%	100%
C. Información de empresa	41,65%	38,10%	Empresa Compañía	100% 0%	71,42% 28,58%
1. Cargo	11,63%	38,10%	Cargo Su papel	100% 0%	87,5% 12,5%
2. Departamento	0%	4,76%			
3. Área de interés	2,33%	0%			
D. Dirección	39,56%	61,90%	Dirección Domicilio	88,23% 11,76%	91,66% 8,33%
1. Ciudad	34,88%	61,90%	Población Ciudad Localidad	53,33% 33,33% 13,33%	0% 61,53% 38,46%
2. Código postal	32,56%	42,86%	Código Postal C.P.	85,71% 14,28%	100% 0%

Subunidad significativa en unidad significativa B-III, Formulario de contacto	Original	Localizado	Unidad léxica	% Original	% Localizado
3. Provincia	32,56%	38,10%	Provincia Estado/provincia/región	92,85% 7,14%	75% 25%
4. País	18,60%	33,33%	País	100%	100%
E. Medios de contacto	100%	100%			
1. Teléfono	81,40%	80,95%	Teléfono Número telefónico Número de teléfono Teléfono fijo Teléfono móvil  Teléfono del trabajo Teléfono directo	97,14% 0% 0% 2,85% 0%  0% 0%	62,05% 6,25% 18,76% 6,25% 12,50%  6,25% 6,25%
2. Fax	25,58%	23,80%	Fax	100%	100%
3. Correo electrónico	95,34%	100%	E-mail Correo electrónico Dirección de email Dirección electrónica Dirección de correo electrónico	85,36% 7,31% 7,31% 0% 0%	47,62% 19,05% 9,52% 14,29% 9,52%
4. Sitio web	4,66%	4,76%			
5. Persona de contacto	6,98%	0%			
F. Motivo de contacto	6,98%	9,52%	¿Qué desea? Tema Tipo de mensaje Motivo de contacto	25% 25% 25% 25%	33,3% 33,3% 0% 33,3%
G. Comentarios	83,72%	85,71%	Comentarios Comentario Asunto Defina el objeto de la consulta Comentarios o dudas Escriba sus comentarios o preguntas aquí Escriba el texto de consulta Consulta Texto [del mensaje] Mensaje Pregunta[s] Cuerpo del mensaje Información solicitada Comentarios o preguntas Observaciones	44,44% 5,56% 8,33% 2,78% 2,78% 0,00% 2,78% 8,33% 8,33% 11,11% 0,00% 0,00% 0,00% 0,00% 5,56	11,11% 0,00% 16,67% 0,00% 0,00% 11,11% 0,00% 11,11% 11,11% 11,11% 5,56% 5,56% 11,11% 0,00%
H. Cómo llegó	4,75%	9,52%	¿Cómo llegó? ¿Dónde nos encontró? ¿Cómo conoció...?	50% 50% 0%	0% 0% 100%
I. Interactividad	100%	100%			
1. Enviar	100%	100%	Enviar Añadir Send Aceptar	93,02% 2,32% 2,32% 2,32%	100% 0% 0% 0%
2. Borrar	44,18%	52,38%	Borrar Volver Reestablecer Borrar y empezar de nuevo Reset	84,21% 5,26% 10,52% 0% 0%	27,27% 0% 0% 9,52% 4,76%

Subunidad significativa en unidad significativa B-III, Formulario de contacto	Original	Localizado	Unidad léxica	% Original	%Localizado
			Poner formulario en blanco	0%	4,76%
			Cancelar	0%	4,76%
			Empezar de nuevo	0%	4,76%
			Reenviar [borrar]	0%	4,76%
			Actualizar [borrar]	0%	4,76%
3. Mensajes sobre el formulario tras enviar.	53,48%	52,38%			

**Tabla 27. Subunidades significativas identificadas en la unidad significativa «formulario de contacto», B-III, y análisis contrastivo de las unidades léxicas en las que se expresan.**

Las subunidades significativas se han organizado en virtud del orden lógico de presentación reflejado en una progresión relativamente fijada en lo que respecta a las subunidades de carácter altamente recurrentes (> 70%) o recurrentes (>70% y <50%). Éstas se suelen presentar según la siguiente progresión:

- I. **Nombre**, 100% en subcorpus original y 100% en subcorpus localizado.
- II. **Dirección**, 39,56% en subcorpus original y 61,90% en subcorpus localizado.
- III. **Medios de contacto**, 100% en subcorpus original y 100% en subcorpus localizado.
- IV. **Comentarios**, 83,72% en subcorpus original y 87,71% en subcorpus localizado.
- V. **Interactividad del formulario**, 100% en subcorpus original y 100% en subcorpus localizado.

El resto de subunidades significativas de carácter ocasional y, por tanto, de menor frecuencia no se insertan según un orden fijo, aunque ninguna de ellas puede aparecer tras la subunidad significativa interactividad (B-III-i), ya que ésta supone la validación y envío del formulario. En cualquier caso, las subunidad significativas de información personal adicional (B-III-b) e información sobre la empresa (B-III-

c) presentan una tendencia a figurar tras la subunidad nombre (B-III-a), aunque esta tendencia no se cumple en la totalidad de formularios.

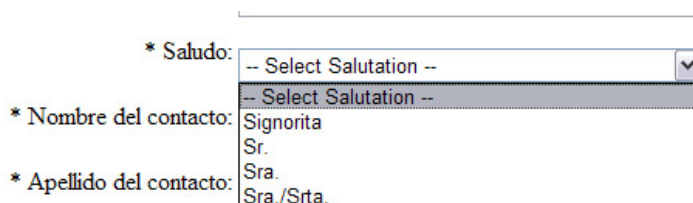
Desde una perspectiva contrastiva, la macroestructura de los formularios de contacto no presenta grandes diferencias en lo que respecta al grado de convencionalización de las distintas subunidades significativas, aunque sí que se observa una asimetría en relación con la subunidad dirección (B-III-d), que es más habitual en las subunidades localizadas (57,14%), que en las originales (39,56%). En general, esta subunidad contrasta con la frecuencia absoluta de la subunidad medios de contacto (B-III-e), que puede deberse a la inmediatez del establecimiento del contacto a través del correo electrónico o el teléfono en lugar de la comunicación por correo tradicional. Como resumen, las mayores diferencias se han observado en (1) el distinto uso de formas lingüísticas convencionales en las subunidades significativas y (2) la presencia de segmentos textuales recurrentes inexistentes en uno u otro subcorpus. Pasamos a continuación a analizar contrastivamente las distintas subunidades significativas y las formas lingüísticas en las que se expresan.

#### **10.2.2.1. Subunidad significativa «nombre» (B-III-a)**

La identificación mediante el nombre propio del usuario es imperativa en todo formulario, y este carácter queda patente al estar marcado este campo como obligatorio. En estas subunidades, el hecho de mayor relevancia relativo a la estructura es la aparición exclusiva en los subtextos localizados de dos segmentos adicionales: el tratamiento (19,05%) y las iniciales (9,52%). Ambos se corresponden con unas diferencias interlingüísticas, que en el caso del tratamiento, son referentes al distinto valor que se otorga en la interacción social al mismo en la cultura española y la estadounidense. Esta baja relevancia atribuida al tratamiento en la cultura española se ha observado igualmente en las investigaciones

contrastivas sobre formularios españoles y alemanes (Sánchez Nieto 2006). En el 50% de los casos este segmento se formula mediante abreviaciones preseleccionadas y, en un caso, constituye un error al incorporar una traducción directa del inglés «salutation» por «saludo».

En relación al segmento «iniciales», éste se corresponde con la abreviación del segundo nombre en el contexto sociocultural estadounidense que no ha lugar en formularios dirigidos a usuarios españoles. Los nombre compuestos españoles se consideran como una unidad y se incluirían en el campo «Nombre» como tal. Este hecho supondría una dificultad de comprensión para los usuarios españoles, que raramente separarían «Juan Carlos» como «Juan» y «C», o «J.» y «Carlos». En cualquier caso, en España la abreviación se puede igualmente producir en el primer elemento de los nombre compuestos, como «J. Carlos» o «M<sup>a</sup> Luisa».



\* Saludo: -- Select Salutation --

\* Nombre del contacto: -- Select Salutation --

\* Apellido del contacto: Signorita  
Sr.  
Sra.  
Sra./Srta.

**Figura 95. Subunidad significativa «tratamiento» (B-III-a-1) exclusiva de los textos localizados.**

En relación a los segmentos significativos propios de estas subunidad, la separación entre «nombre» y «apellidos» es más habitual en los textos localizados (90,48%) que en los originales (74,42%). Los textos originales muestran una mayor frecuencia de la combinación de ambos términos en un campo «nombre y apellidos»<sup>433</sup>. Consideramos que el hecho más significativo lo constituye la aparición de la forma sustantiva singular «apellido» en los corpus localizados (33,33%), que no figura en el corpus original ya que sólo está presente su forma plural. En este sentido, es preciso apuntar que este hecho responde al uso de un sólo apellido en el contexto

<sup>433</sup> Este hecho imposibilitaría la alfabetización posterior por apellidos en las bases de datos que se creen con estos formularios.



sociocultural anglosajón, con lo que esta forma singular reflejaría una realidad sociocultural ajena a la cultura meta y, por tanto, supondría un error de traducción al no ajustarse a la forma lingüística convencional asociada a este subtexto.

En resumen, la asimetría observada en esta subunidad se concreta en (1) la aparición exclusiva de los segmentos de tratamiento e iniciales en los textos localizados y (2) la alta frecuencia de la forma sustantiva singular «apellido». Estos hechos demuestran una relativa falta de adaptación a las convenciones socioculturales españoles en ciertos casos en la organización de estos datos personales y, por otro lado, la dependencia de este tipo de textos de la internacionalización al ser difícil al traductor realizar cambios a una superestructura prefijada.

#### **10.2.2.2. Subunidad significativa «información personal adicional» (B-III-b)**

Esta subunidad significativa se basa en la información personal adicional sobre los clientes o usuarios. Ésta se centra en los números de identificación personal, DNI, o el número de identificación fiscal para las empresas, CIF<sup>434</sup>, así como en el sexo y edad del usuario. Como demuestra su baja frecuencia, 2,33% en originales y 9,92% en los localizados, esta subunidad significativa revierte una escasa importancia en los formularios de contacto en comparación con otros tipos de formularios de carácter contractual (Fernández Sánchez 2004). Se hace evidente que la identificación legal completa del usuario no es obligatoria en el intercambio comunicativo que se entabla mediante estos formularios, aunque sí que sea necesaria la identificación plena del emisor del texto en el bloque **empresa** (B).

---

<sup>434</sup> Código de Identificación Fiscal.

### 10.2.2.3. Subunidad significativa «empresa» (B-III-c)

La presente subunidad significativa está dedicada a la información sobre aquellas empresas que se establezcan como receptoras del texto, puesto que las relaciones comerciales no sólo se producen entre empresas emisoras y clientes, sino que éstas también se entablan entre empresas, tal y como se observó en el bloque comunicativo **zonas de clientes** (G). En ambos subcorpus la frecuencia de aparición es similar, un 41,65% en textos originales y un 38,10% en localizados. Normalmente esta subunidad coexiste con la subunidad significativa nombre (B-III-a), ya que el intercambio comunicativo se establece entre individuos que actúan en representación de la empresa, tal como revela el hecho de que aparezca el segmento «cargo» en el 11,63% de textos originales y 38,10% de los originales. Este segmento no se da en la misma medida en ambos subcorpus y podría deberse una acción valorativa en la interacción social por parte del emisor sobre el receptor, o posiblemente a la mayor valoración social de los cargos de prestigio en las empresas en el contexto estadounidense. Ésta mayor importancia entroncaría en parte con la presencia exclusiva del segmento «tratamiento» en los textos localizados (véase 10.2.2.1).

En lo que respecta a la terminología, la unidad léxica «empresa» se presenta como forma lingüística convencional en esta sección, con una frecuencia del 100% en segmentos originales. Por contraste, en los textos localizados aparece además su sinónimo «compañía» (28,58%), que reflejaría el término original del inglés *company*. Esta influencia de las formas léxicas en inglés se refleja igualmente en el segmento «cargo» con una aparición del anglicismo léxico «su papel» (*your role*).

### 10.2.2.4. Subunidad significativa «dirección» (B-III-d)

En el contexto de la WWW, las direcciones de correo electrónico coexisten con las direcciones postales tradicionales. No obstante, éstas presentan una menor

importancia en contexto digital que las primeras, ya que el campo sobre la dirección postal figura en el 39,56% de textos originales y 61,90% de los localizados, mientras que el correo electrónico aparece en mayor media, el 95,34% de textos originales y 100% de los localizados. Queda así patente que la inmediatez de la interacción comunicativa que proporciona el medio ha relegado al correo postal a un segundo plano. Dentro de estas subunidad, los segmentos encontrados son dirección, población, código postal, provincia y país. Cabe señalar que todos estos segmentos presentan una mayor frecuencia en los textos localizados. Sobre todo es significativa la menor frecuencia del segmento país en textos originales (18,60%) frente a los localizados (33,33%), fruto del carácter internacional de estos últimos sitios globalizados.

De entre los datos más significativos del análisis contrastivo terminológico, hemos de resaltar la inexistencia en el subcorpus localizado de la forma lingüística convencional «población», que presenta en textos originales un alto nivel de convencionalización (53,33%) y es inexistente en los localizados (0%). En éstos aparece «ciudad» como la término más habitual (61,53%), seguido de «localidad», (38,46%). Es igualmente significativa la presencia de la abreviación «C.P.» en los textos originales (14,28%), inexistente en los textos localizados. Finalmente, resulta más habitual la inclusión de las distintas subdivisiones geográficas de los países «estado-provincia-región» en los localizados (25%) frente a los originales (7,14%), en virtud de su carácter de textos internacionalizados<sup>435</sup>.

#### **10.2.2.5. Subunidad significativa «medios de contacto» (B-III-e)**

La presente subunidad está dedicada a incorporar los medios de contacto por los que canalizar los futuros intercambios comunicativos entre ambos interlocutores.

---

<sup>435</sup> A pesar de que los textos recopilados en el corpus se dirigen exclusivamente a España y no a un usuario hispanohablante internacional, con lo que este campo debería incluir la subdivisión regional española convencional, la provincia.

A pesar de que el número de teléfono y fax suele preceder al correo electrónico, éste último resulta más frecuente: 95,34% de textos originales y 100% de localizados. De esta manera, el medio más significativo según la frecuencia de aparición sería el correo electrónico, seguido del teléfono (81,40% y 80,95%), el número de fax (25,58% y 23,80%) y el sitio web (4,66% y 4,76%). Todos estos segmentos presentan una frecuencia interlingüísticamente similar en virtud de un mismo carácter internacional de estos medios de contacto.

La unidad léxica convencional para el primer segmento de esta subunidad es el sustantivo «teléfono» (97,14%). En los textos localizados este segmento presenta una mayor variación, ya que junto con el sustantivo «teléfono» (62,05%), se encuentra el compuesto sintagmático «número de teléfono» (18,76%), y el compuesto adjetival «número telefónico» (6,25%), que reflejarían la unidad léxica original en inglés *phone number*. En inglés el uso del compuesto es obligatorio ya que el sustantivo en inglés, *phone*, se refiere al aparato en sí, con lo que su formulación se presenta obligatoriamente mediante el compuesto sintagmático *phone number*<sup>436</sup>. En los textos localizados suelen aparecer distintos tipos de teléfono como «teléfono fijo», «teléfono del trabajo» o «teléfono directo». Éstos dos últimos no aparecen en el corpus original.

Quizás el segmento de mayor interés desde un punto de vista contrastivo es el referente al correo electrónico. Contrario a lo que cabría esperar, la unidad léxica convencional en los formularios españoles es el anglicismo puro «E-mail»<sup>437</sup> (85,36%), que aparece asimismo en un anglicismo adaptado «dirección de email» (7,31%), de frecuencia similar al calco «correo electrónico» (7,31%). A pesar de su baja frecuencia, esta última unidad léxica sería la recogida en el Diccionario

---

<sup>436</sup> Es convencional su abreviación «Ph.», o el uso de la almohadilla para señalar el número, «phone #».

<sup>437</sup> El diccionario panhispánico de dudas de la RAE (2005) resulta tajante en su oposición a este anglicismo puro, «Su uso —así como el de su abreviación mail— es innecesario, por existir alternativas en español en todos estos casos.[...] Resulta inadmisibles la adopción del recurso inglés consistente en utilizar una e (abreviatura extrema de electronic) a modo de sufijo o prefijo en español» (RAE 2005: versión electrónica, <http://buscon.rae.es/dpdI/SrvltConsulta?lema=email>).

Panhispanico de Dudas de la RAE (2005). En el corpus localizado la unidad léxica más habitual es igualmente «email», aunque su frecuencia es menor (47,62%). En segundo lugar aparece «correo electrónico» (19,05%) y «dirección electrónica», (14,29%), ésta última inexistente en el corpus original, aunque también aparece recogida como alternativa en el mismo diccionario (RAE 2005). Se encontró igualmente el compuesto sintagmático «dirección de correo electrónico» (9,52%) que no aparece en el original, y «dirección de email» (9,52%). Contrastivamente, el 52,38% de unidades léxicas localizadas que se usan incorporan una traducción del término «email», mientras que éstas sólo constituyen el 14,62% de uso en los textos originales españoles. Esta asimetría podría comportar una mayor explicitación de los textos localizados, que ya fue observada en la mayor frecuencia de uso del anglicismo puro «links» en los menús de navegación originales en la sección H4.

#### **10.2.2.6. Subunidad significativa «comentarios» (B-III-g)**

La subunidad significativa «comentarios» representa el campo de mayor extensión, al permitir que el usuario se exprese libremente sin limitación alguna de espacio. El análisis muestra que este campo puede estar representado por los siguientes términos: «comentarios», «asuntos», «consultas», «dudas», «preguntas», «mensajes», u «observaciones». La presencia de este segmento es similar en ambos subcorpus, 83,72% y 85,71%. Hemos identificado asimismo la posibilidad de presentar opciones preseleccionadas en este segmento, aunque su frecuencia es relativamente baja.

La forma más convencionalizada en los textos originales es «comentarios» (44,44%), o «comentario» en singular (5,56%), que combinados supondrían el 50% de uso y, por tanto, supondrían una forma lingüística convencional. Las siguientes unidades léxicas serían «mensaje», 11,11% y «asunto» 8,33%. Este segmento incluye actos de habla directivos con la inclusión de formas verbales en modo

imperativo como «escriba sus comentarios», o «defina el objeto de su consulta». Por su parte, en los textos localizados no queda recogido el carácter convencional de «comentarios», ya que sólo está presente en el 11,11% de formularios. En éstos es más frecuente señalar este campo con las palabras «pregunta[s]» (16,55%), inexistente en el corpus original, o «asunto» (16,55%).

#### **10.2.2.7. Subunidad significativa «envío y validación de formulario» (B-III-i)**

Esta última subunidad significativa confiere a los formularios web su carácter distintivo frente a los impresos, sobre todo en virtud de la interactividad que supone su envío inmediato a través de Internet<sup>438</sup>. Su carácter distintivo surge por la posibilidad de comprobar la corrección en el formato de los datos y la consecuente interacción con el usuario<sup>439</sup>. Esta subunidad supone la validación y envío del formulario ya cumplimentado con lo que, tal como señala Sánchez Nieto (2006), éste adquiriría su carácter de texto completo funcional. Este segmento se encuentra tipográficamente marcado por medio de su inclusión en una caja o por presentar un color diferenciado, un hecho que indica al usuario su carácter interactivo<sup>440</sup>. Los formularios impresos pueden presentar una subunidad con instrucciones para su envío, aunque sólo en los formularios interactivos existe la posibilidad de borrar los datos del formulario o comprobar interactivamente la cumplimentación de campos obligatorios y la corrección de sus formatos. Este sería el caso de mensajes de error que aparecerían en pantalla si se escribiera un formato de correo electrónico erróneo, tal como «Compruebe si ha introducido correctamente el email» [Secúritas].

---

<sup>438</sup> La investigación sobre formularios de inscripción de Sánchez Nieto (2006) mezcla los formulario impresos y los formularios web, que según nuestra perspectiva presentan diferencias claras y suponen un género diferenciado como demuestra la inclusión de esta subunidad significativa por la que se envía, borra, o comprueba la validez de los datos del formulario.

<sup>439</sup> Que se realiza por medio de scripts, normalmente javascript (véase 9.2.5).

<sup>440</sup> En términos de usabilidad se presentan como una «affordance», un espacio en la pantalla que se puede activar y representa una interactividad.

El primer segmento se expresa a través una forma lingüística altamente convencionalizada en ambos subcorpus, «enviar», 93,02% y 100%. En los textos originales se identificó igualmente «añadir», «send» y «aceptar». Las mayores diferencias se encontraron en el segundo segmento, «borrar», que se encuentra altamente convencionalizado en español directamente por medio del verbo, «borrar» (84,21%). Este carácter de forma convencional no se refleja en los textos localizados, en donde identificamos ocho posibles alternativas, siete de ellas inexistentes en el subcorpus original. En este subcorpus «borrar» supone el 27,27% de uso, palabra que aparece igualmente en la segunda forma convencional más frecuente, «borrar y empezar de nuevo» (9,52%). Este carácter de reinicio en el rellenado se encuentra igualmente en «poner formulario en blanco» (4,76%), «empezar de nuevo» (4,76%), y dos unidades léxicas adicionales que incluyen el prefijo «re»; «reenviar» (4,76%) y el préstamo «reset» (4,76%). La traducción directa de este último, «reestablecer», aparece en el subcorpus original (10,52%).

Por último, el segmento textual que supone la posibilidad de validación de los datos suministrados es exclusivo de los formularios web y figura en el 53,48% de los textos originales y el 52,38% de localizados. La validación se realiza a través de una rutina en lenguaje javascript, que comprueba la cumplimentación de los campos obligatorios, la aceptación de avisos legales, y el correcto formato de correos electrónicos, números de teléfono y códigos postales. La media de oraciones por formulario interactivo que potencialmente se puedan presentar al usuario es similar, 7,26 oraciones en textos originales y 7,45 en los textos localizados.

La función de estos segmentos es exhortativa con opción y suponen actos de habla directivos. A continuación presentamos los recursos lingüísticos identificados en caso de no incluir un campo obligatorio:

- «Deber» + forma verbal infinitiva: «Debe rellenar [...]», «Debe introducir [...]», «Debe escribir [...]», «Debe revisar [...]», «debe seleccionar [...]». Su frecuencia es 50,70% de uso en textos originales y 19,51% en textos localizados.
- «Por favor» + forma verbal en modo imperativo: «Por favor, rellene [...]»  
Por favor, introduzca [...]». Su frecuencia es 27,46% de uso en textos originales y 31,7% en textos localizados.
- Forma verbal en modo imperativo: «compruebe si ha introducido [...]», «Introduzca [...]», «seleccione [...]», «rellene [...]», «limite la extensión [...]», «escriba [...]». Su frecuencia es 12,67% de uso en textos originales y 23,17% en textos localizados.
- Forma verbal en pretérito perfecto: «no [se] ha introducido [...]», «no ha especificado [...]», «no ha insertado [...]». Su frecuencia es 4,92% de uso en textos originales y 2,43% en textos localizados.

Contrastivamente, queda patente de nuevo la mayor frecuencia en el uso de formas verbales directas en modo imperativo en los textos localizados, 23,17% frente a 12,67% en originales, aunque en ambos es más frecuente modulado la expresión «por favor», el 31,7% en textos localizados y el 26,46% en los originales. En su conjunto, los textos localizados presentarían un 54,87% de uso de formas imperativas frente a 39,13%, aunque en los originales es más frecuente la inclusión de la fórmula de cortesía «por favor». No obstante, en los textos originales es mayoritario el recurso a la perífrasis verbal imperativa con el verbo «deber» (50,70%).

No se han identificado formas verbales impersonales en virtud del necesario carácter de instrucción directa al usuario. Es de interés que en la validación de formatos de correos electrónicos, números de teléfono o códigos postales, las



instrucciones se formulan de manera diferenciada entre ambos subcorpus: en el localizado es mayoritario el uso del adjetivo «válido» (14,63% del conjunto de segmentos identificados), sobre todo por medio de la estructura nombre + [verbo opcional] + adjetivo «no válido[a]». Por ejemplo «Formato de código postal no válido», «dirección de correo electrónico no válido». Esta estructura aparece en el 9,75% del global de segmentos identificados. Se presenta igualmente su forma adjetival prefijada «inválido-a» (2,43% de segmentos). Por su parte, en los textos originales los recursos para señalar la incorrección en el formato de algún campo se centran en las formas adjetivales «incorrecto» - «el email es incorrecto», «erróneo» - «correo erróneo» o «válido» - «indique un email válido». Tal como revela esta divergencia, la forma adjetival «no válido» representaría una forma no convencional en los textos originales y un anglicismo léxico que se corresponde con el adjetivo inglés «invalid».

### 10.3. Análisis adicional del bloque legal (F)

El bloque comunicativo **legal** (F) fue el segundo seleccionado de cara a un análisis adicional de su superestructura y formas lingüísticas convencionales. Su selección se debió a que presenta la mayor diferencia en frecuencia de aparición entre ambas superestructuras prototípicas (véase 5.3.1), de ahí su relevancia desde una perspectiva contrastiva. Asimismo, este bloque fue seleccionado dado el alto grado de convencionalización que presentan los textos jurídicos en general (Borja Albí 2000, 2003), con lo que asumimos que el mayor grado de convencionalización terminológica y fraseológica propia de los géneros jurídicos facilitará el análisis de sus formas lingüísticas convencionales<sup>441</sup>. En este sentido, el carácter altamente

---

<sup>441</sup> El alto grado de convencionalización de esta sección en los textos anglosajones queda reflejado en la existencia de generadores automáticos de políticas de privacidad en línea en los sitios web de organismos internacionales de prestigio como la OCDE (Organización para la Cooperación el Desarrollo Económico) en: [http://www.oecd.org/document/39/0,3343,en\\_2649\\_201185\\_28863271\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/document/39/0,3343,en_2649_201185_28863271_1_1_1_1,00.html)

convencional de estos subtextos legales se manifiesta en España a través de publicaciones con repertorios de contratos web tipo (ej. González Aguilar *et al.* 2004).

Para realizar este análisis adicional intralingüístico recopilamos un subcorpus comparable en el que incorporamos toda página web que constituyese una de las secciones legales en este bloque: las secciones **avisos legales** (F1), **política de privacidad** (F2) y **condiciones generales** (F3). El análisis se realizó combinando este conjunto de secciones, ya que los patrones de distribución de las unidades y subunidades significativas identificadas no es homogéneo entre ambos subcorpus. Su agrupación llega incluso a ser irregular entre sitios web de una misma sección del subcorpus, y este hecho constituiría una de las principales peculiaridades que hemos identificado en los subtextos legales.

Podemos ilustrar este punto con el siguiente ejemplo. En nuestro análisis identificamos una subunidad significativa en la que la empresa declina toda responsabilidad por la exactitud del contenido del sitio web (F-3-II-d). Esta subunidad forma parte de la sección legal de **condiciones generales** (F3), aunque dependiendo del sitio, ésta puede aparecer en una sección de **avisos legales** (F1), o incluso en **política de privacidad** (F-2). Esta unidad fue adjudicada dentro de la unidad significativa de limitación de responsabilidades (F-3-II), que a su vez se encuentra dentro de la sección de **condiciones generales** (F-3), y finalmente dentro del bloque **legal** (F). Por lo tanto, la adscripción de cada unidad y subunidad se asignó según su contenido proposicional y no exclusivamente según la sección a la que el emisor la hubiera incorporado.

Otro ejemplo sería la preferencia en los sitios web españoles por formular el bloque legal en una sola página denominada «avisos legales», que normalmente contiene cláusulas referidas a sección de **política de privacidad** (F2) y-o las **condiciones**

**generales** (F3). Defendemos que esta característica justifica nuestro análisis conjunto de estas tres secciones. En consecuencia, los porcentajes de frecuencia de cada unidad o subunidad que presentamos son representativos de su presencia global dentro del conjunto de sitios web con bloque legal, en lugar de su frecuencia en el conjunto de sitios del corpus general.

### **10.3.1. Análisis intratextual preliminar de subcorpus comparable legal**

El análisis estadístico de estos subtextos muestra una clara diferencia en cuanto a la media de palabras por página web. Las páginas originales dentro de esta sección muestran una media de 904,31 palabras, mientras que la media de palabras en las localizadas asciende a 1736,28, una cifra que supone un aumento cercano al 200%. Esta mayor extensión por sitio web se correspondería con la mayor explicitación de las distintas disposiciones legales en los sitios web estadounidenses<sup>442</sup>.

Subcorpus legal	Media de palabras por página en subcorpus	Media de palabras por página en corpus global	Media por sitio web con bloque legal (F)
Original	904,31	258,87	1074,94
Localizado	1736,28	416,07	2415,69

**Tabla 28. Tabla contrastiva de la media de palabras en secciones legales (F) por sitio web.**

Esta diferencia en la media de palabras es más significativa si computamos el conjunto de palabras en subtextos legales por sitio web analizado. En este caso, los sitios localizados presentan una media de 2415,69 palabras en secciones legales por sitio web, mientras que la media en textos originales sería de 1074,94 palabras. Esta mayor cantidad de palabras en los sitios web localizados se correspondería con una mayor división de este bloque legal en páginas independientes (véase 9.1.2.6). En

---

<sup>442</sup> Aunque como indican las guías de estilo web, el usuario raramente las leerá dada su mayor extensión (Price y Price 2002).

cualquier caso, si comparamos la media de palabras por página con contenido legal en ambos conjuntos textuales, ésta supera ampliamente la media global de palabras por página en general (véase 7.4) y, consecuentemente, se aleja de las recomendaciones de redacción sobre la máxima extensión por página o su posible disgregación en distintas páginas webs que se recogen en los manuales de estilo web (véase 5.9), o de usabilidad (Nielsen y Loranger 2006; Nielsen y Tahir 2002).

Dado que las páginas de contenido legal se presentan como los subtextos de mayor extensión en sitios web corporativos, este hecho conlleva la mayor frecuencia en el subcorpus localizado de la subunidad significativa que apela a leer todo el texto (F-1-I-b), el 40,52% frente a 4,32% en originales. Esta mayor extensión proyecta su reflejo en la presencia de segmentos textuales que mencionan expresamente esta mayor envergadura de los subtextos legales:

Aunque este documento pueda parecerle **demasiado extenso**, creemos que debe estar tan informado como sea posible y, por ello, hemos confeccionado una política de confidencialidad detallada [Autodesk].

### 10.3.2. La lengua de expresión en el bloque legal (F)

Dentro del bloque legal cobra una especial importancia la presencia de páginas en inglés o las menciones implícitas a los idiomas de redacción o traducción. En el subcorpus se eliminaron cuatro páginas en inglés dentro de los sitios web localizados (4,2%). Adicionalmente, en este subcorpus identificamos referencias al idioma de redacción o interpretación de estos subtextos legales, sobre todo referentes a la validez jurídica exclusiva de la versión en inglés. Los siguientes segmentos ilustrarán este punto:

En la medida máxima permitida por las disposiciones legales, el idioma que regirá las Condiciones de Uso aquí establecidas será el **inglés**. [...] Las partes declaran su exigencia de que estas Condiciones de Uso y todos los demás documentos, presentes o futuros, relativos a las mismas sean redactados **sólo en idioma inglés** [UPS].

Al proporcionar algún contenido en **idioma extranjero**, no asume obligación alguna para brindar **la traducción** de todo o parte de su Sitio web en ningún idioma, excepto por lo que Cargill determine [Cargill].

Además, en ambos casos los términos legales eximen a la empresa de cualquier responsabilidad en referencia a sus traducciones:

Las **traducciones** a otros idiomas, en caso de haberlas, se facilitarán únicamente a efectos prácticos [UPS].

Cargill no realiza declaración o manifestación alguna de que las **traducciones** sean fieles y precisas [Cargill].

En los textos localizados queda patente que, si se emprendieran acciones legales, el emisor no se limitaría a elegir el fuero o jurisdicción (F-1-IV), sino que también seleccionaría el idioma de interpretación del texto, relegando los textos traducidos a un segundo plano. En cualquier caso, estas menciones implícitas al idioma de validez jurídica en estos subtextos legales aparecen en sitios web con una misma versión de los subtextos legales para todas sus localizaciones, sin una adaptación adicional a la legislación vigente de cada país o región. De ahí que asignemos a estos sitios web un grado de localización medio por su falta de adaptación al contexto sociocultural de llegada (véase 2.7). En el polo opuesto, esta adaptación a la legislación española se recoge en el 15,78% de los sitios web localizados.

### **10.3.3. Análisis contrastivo de unidades y subunidades significativas en bloque legal (F)**

La metodología de recuento consistió en analizar manualmente la totalidad de textos que componen el subcorpus, procediendo a identificar las variables de análisis que asociamos a cada unidad y subunidad significativa. A continuación, procedimos a asignarlas a una sección determinada de las tres posibles en el bloque **legal** (F), aunque nuestra asignación no se correspondiera necesariamente con la sección a la que el emisor la incorporara. En el caso específico las **políticas de**

**privacidad** (F2), partimos de la investigación previa sobre sitios web corporativos estadounidenses de Liu y Arnett (2000), en la que se identificaron doce posibles variables dentro de esta sección. No obstante, en nuestro análisis identificamos unidades significativas específicas del contexto sociocultural español, con lo que nos fue preciso complementar esta propuesta previa. Éste fue el caso, por ejemplo, de la unidad referente a la ley española de protección de datos (F-2-I-a).

La tabla contrastiva global que sigue presenta las unidades significativas y subunidades significativas de mayor importancia. Como indicamos, el porcentaje de frecuencia es representativo de la presencia de tal unidad o subunidad dentro el conjunto de sitios web que incorporan textos legales. En este sentido, hemos de recordar la significativa diferencia en la frecuencia de bloques legales entre los textos originales y localizados (véase 5.3.1). En nuestra opinión, este análisis contrastivo realizado sobre la frecuencia de aparición de unidades y subunidades significativas por sitio web con secciones legales será más esclarecedor. Al igual que en el caso anterior, los bloques se identificarán con una letra en mayúscula, las secciones comunicativas con un número arábigo, las unidades significativas con un número romano, y las diferentes subunidades significativas con una letra consecutiva.

El cómputo de las posibles unidades significativas identificadas arroja una cantidad similar entre las tres secciones posibles: diez posibles unidades significativas en la sección **avisos legales** (F1), doce posibles unidades en la sección de **políticas de privacidad** (F2), y trece posibles unidades en la sección de **condiciones generales** (F3). No obstante, dada la mayor extensión de los textos localizados que supone un número superior de subunidades significativa, su identificación no es exhaustiva.

Sección y unidad significativa	Subunidad significativa	Original	Localizado
<b>F1. AVISOS LEGALES</b>		<b>27,90%</b> <sup>443</sup>	<b>47,36%</b>
F-1- I. Presentación			
	a. Bienvenida	0,00	17,39
	b. Apelación a leer texto	13,21	65,22
F-1- II. Aceptación de términos legales	a. Aceptar	58,49	93,48
	b. Abandonar el sitio	7,55	15,22
F-1- III. Registro legal	Inscripción de empresa	73,58	34,78
	a. Domicilio social	30,19	30,43
	b. Código de Identificación Fiscal	52,83	30,43
F-1- IV. Legislación Aplicable y Jurisdicción		18,87	65,22
F-1- V. Propiedad intelectual-uso del contenido		67,92	89,13
	a. Autorización por escrito	26,42	47,83
	b. Qué material está protegido	26,42	50,00
F-1- VI. Dónde almacenan		0,00	4,35
F-1- VII. Propiedad el sitio web		56,60	71,74
F-1- VIII. Destinatario del sitio web		5,66	23,91
F-1- IX. Uso de las marcas registradas		3,77	32,61
F-1-X. Fecha de entrada en vigor o revisión		0,00	23,91
<b>F2. PROTECCIÓN DE DATOS PERSONALES</b>		<b>13,37%</b>	<b>70,52%</b>
F-2-I. Cumplimiento de leyes españolas pertinentes			
	a. Ley española LOPD 1999	49,06	13,04
	b. Ley 34/2002, de 11 de julio, de Servicios de la Sociedad de la Inform. y de Comercio Electr.	22,64	6,52
F-2-II. Recogida		47,17	45,65
F-2-III. Derechos de acceso, rectificación, cancelación y oposición		47,17	28,26
F-2-IV. Enlaces hacia o desde el sitio web		43,40	86,96
F-2-V. Uso de datos personales		41,51	36,96
F-2-VI. Contacto		39,62	58,70
F-2-VII. Cookies		32,08	50,00
	a. Definición	22,64	19,91
F-2-VIII. Revelación		26,42	28,26
F-2-IX. Seguridad		26,42	36,96
	b. Riesgos	9,43	10,87
F-2-X. Notificación de revelación		13,21	4,35
F-2-XI. Protección de menores		7,55	39,16
F-2-XII. Direcciones IP		5,66	15,22
<b>F3. CONDICIONES GENERALES</b>		<b>4,65%</b>	<b>38,94%</b>
F-3-I. Limitación de responsabilidad		67,92	86,92
	a. Problemas por virus	41,51	58,70
	b. Errores del sitio – interrupción	35,85	56,52
	c. Por daños y perjuicios	24,53	52,17
	d. Exactitud de contenidos	20,75	23,91
	e. Por contenidos de usuarios	5,66	13,04

<sup>443</sup> Frecuencia de aparición según el análisis de títulos de páginas web (véase 9.1.1).

Sección y unidad significativa	Subunidad significativa	Original	Localizado
F-3-II. Modificaciones o cambios de contenidos		41,51	54,35
F-3-III. Modificación o cambio de términos		24,53	76,09
F-3-IV. Uso adecuado		22,64	17,39
F-3-V. Uso personal y privado		16,98	15,22
F-3-VI. Restringir el acceso		13,21	17,39
F-3-VII. Publicar materiales ilegales, polémicos, pornográficos, amenazantes		9,43	39,13
F-3-VIII. Uso del usuario con fines ilícitos, lesivos, negligentes o fraudulentos		7,55	2,17
F-3-IX. Exclusión de garantías		5,66	30,43
F-3-X. Prevalece el texto impreso sobre el texto web		5,66	4,35
F-3-X. Nielsen ratings		3,77	0,00
F-3-XI. Material enviado es público		1,89	26,09
F-3-XII. Expectativas futuras		0,00	10,87

**Tabla 29. Análisis contrastivo de unidades y subunidades significativas identificadas en bloque legal (F).**

A continuación elaboramos los perfiles textuales prototípicos a partir de estas frecuencias que presentamos en la Tabla 30. Estos perfiles se han obtenido mediante aquellas unidades y subunidades recurrentes cuya frecuencia fuera igual o superior al 50%. Los textos originales se caracterizan por incluir la aceptación de los términos legales (F-1-II), la inscripción de la empresa en el registro mercantil (F-1-III-a) y su código de identificación fiscal (F-1-III-b), la unidad significativa sobre los derechos de propiedad sobre los contenidos pertenecientes al sitio web (F-1-VI) y los detalles sobre la propiedad del mismo (F-1-VII). En lo que respecta a la sección **política de privacidad** (F2), sólo sería característica la referencia al cumplimiento de la ley española LOPD (F-2-I-a). Finalmente, en sección condiciones generales se incluiría la unidad significativa sobre la limitación de responsabilidades en general (F-3-I).

Por su parte, son quince las unidades y subunidades significativas características que configuran el perfil textual de los subtextos localizados, frente a sólo siete en los originales. Esta mayor presencia de unidades y subunidades se corresponde con su mayor extensión media presentada anteriormente. El perfil textual característico del bloque **legal** (F) en los textos localizados incluye la apelación a



leer el texto (F-1-I-b), la aceptación de los términos legales (F-1-II-a), la legislación aplicable y jurisdicción (F-1-IV), los derechos de propiedad intelectual del sitio web (F-1-VI), la especificación de los materiales que se encuentran protegidos (F-1-VI-b), y la unidad referente al titular del sitio web (F-1-VII). En la protección de datos personales son características la unidad significativa sobre los enlaces hacia o desde el sitio web (F-2-IV) y la información de contacto para ejercer el derecho de rectificar, modificar o cancelar los datos almacenados por la empresa (F-2-VI). Por último, dentro de la sección de condiciones generales son características la unidad significativa sobre la limitación de responsabilidad (F-3-I), las modificaciones o cambios de contenidos (F-3-II), la unidad significativa sobre la modificación o cambio de los términos legales (F-3-III), y las subunidades significativas de limitación de responsabilidades por virus informáticos (F-3-I-a), la limitación de responsabilidades por errores o interrupciones en el acceso al sitio (F-3-I-b), y la limitación de responsabilidad por daños y perjuicios (F-3-I-c).

De entre las diferencias más significativas entre ambos perfiles textuales prototípicos, destacan (1) la mayor frecuencia de las unidades comunicativas originales referentes al contexto jurídico español de los textos originales, tal como la inscripción de la empresa en el registro mercantil (F-1-III), 73,58% frente al 34,78% en los localizados, o la unidad que expresa el cumplimiento de la ley española de protección de datos personales (F-2-I-a), 49,06% frente a 13,04%. Por su parte, (2) en los textos localizados aparece con una mayor frecuencia la subunidad significativa en la que se apela a la lectura del texto legal (F-1-I-b), el 13,21% en originales frente a un 65,22% en localizados, que podría responder a una mayor apelación directa al usuario en los textos web anglosajones. De igual modo, (3) es superior la frecuencia de la subunidad significativa sobre la ley y jurisdicción aplicable (F-1-IV), 18,87% frente al 65,22%, que respondería a la necesidad de establecer la jurisdicción o fuero aplicable en caso de que se emprendieran acciones legales. En este sentido, los sitios españoles se regirían

implícitamente por la jurisdicción española y, por tanto, no se requiere la expresión explícita de esta subunidad en el texto.

<b>Textos originales</b>	<b>%</b>	<b>Textos localizados</b>	<b>%</b>
<b>F1. AVISOS LEGALES</b>		<b>F1. AVISOS LEGALES</b>	
		F-1- I-b. Apelación a leer texto	65,22
F-1- II. Aceptación de términos legales. a. Aceptar	58,49	F-1- II. Aceptación de términos legales. a. Aceptar	93,48
F-1- III. Registro legal a. Inscripción de empresa	73,58	F-1- III. Registro legal a. Inscripción de empresa	34,78
F-1- III-b. Código de Identificación Fiscal	52,83	F-1- III-b. Código de Identificación Fiscal	30,43
		F-1- IV. Legislación Aplicable y Jurisdicción	65,22
F-1- VI. Propiedad intelectual-uso del contenido	67,92	F-1- VI. Propiedad intelectual-uso del contenido	89,13
		F-1- VI-a. Qué material está protegido	50,00
F-1- VII. Titular del sitio web	56,60	F-1- VII. Titular del sitio web	71,74
<b>F2. PROTECCIÓN DE DATOS PERSONALES</b>		<b>F2. PROTECCIÓN DE DATOS PERSONALES</b>	
F-2-I-a. Ley española LOPD 1999	49,06		
		F-2-IV. Enlaces hacia o desde el sitio web	86,96
		F-2-VI. Contacto	58,70
<b>F3. CONDICIONES GENERALES</b>		<b>F3. CONDICIONES GENERALES</b>	
F-3-II. Limitación de responsabilidad	67,92	F-3-II. Limitación de responsabilidad	86,92
		F-3-II-a. Problemas por virus	58,70
		F-3-II-b. Errores del sitio – interrupción	56,52
		F-3-II-c. Por daños y perjuicios	52,17
		F-3-I. Modificaciones o cambios de contenidos	54,35
		F-3-III. Modificación o cambio de términos	76,09

Tabla 30. Perfiles textuales característicos en bloque legal (F) en ambos subcorpus.

### 10.3.4. Análisis léxico comparativo del subcorpus legal

Tras el análisis comparativo de la superestructura de los textos legales, el siguiente paso consistió en realizar un análisis léxico de ambas secciones del subcorpus en aras a dilucidar las formas lingüísticas asociadas a la misma. En primer lugar procedimos a realizar un listado de las palabras de ambos subcorpus por grupos de palabras o *clusters* de cuatro elementos. El objetivo fue observar la

correspondencia de las posibles formas lingüísticas convencionales y la fraseología específica de estos subtextos:

Lista de colocaciones léxicas recurrentes en subcorpus original (4)				Lista de colocaciones léxicas recurrentes en subcorpus localizado (4)			
N	Word	Freq.	%	N	Word	Freq.	%
1	datos de carácter personal	86	0.15	1	de este sitio web	87	0.08
2	de protección de datos	55	0.10	2	se reserva el derecho	76	0.07
3	de datos de carácter	48	0.08	3	en este sitio web	65	0.06
4	protección de datos de	42	0.07	4	en el sitio web	59	0.05
5	se reserva el derecho	42	0.07	5	derechos de propiedad intelectual	52	0.05
6	en el registro mercantil	39	0.07	6	estos términos y condiciones	47	0.04
7	inscrita en el registro	39	0.07	7	en el registro mercantil	44	0.04
8	el registro mercantil de	38	0.07	8	lmie y sus filiales	43	0.04
9	derechos de propiedad intelectual	33	0.06	9	términos y condiciones de	43	0.04
10	las presentes condiciones generales	32	0.06	10	inscrita en el registro	42	0.04
11	dirección de correo electrónico	31	0.05	11	los sitios web de	40	0.04
12	registro mercantil de Madrid	29	0.05	12	reserva el derecho a	40	0.03
13	los derechos de propiedad	28	0.05	13	la política de privacidad	39	0.03
14	la ley orgánica ##	25	0.04	14	dirección de correo electrónico	38	0.03
15	de # de diciembre	24	0.04	15	esta política de privacidad	38	0.03
16	los daños y perjuicios	24	0.04	16	política de privacidad y	38	0.03
17	de la página web	23	0.04	17	el registro mercantil de	37	0.03
18	de los derechos de	23	0.04	18	el sitio web de	36	0.03
19	de este sitio web	22	0.04	19	y tratamiento de datos	35	0.03
20	los datos de carácter	22	0.04	20	privacidad y tratamiento de	34	0.03
21	no se hace responsable	22	0.04	21	datos de carácter personal	33	0.03
22	a disposición de los	21	0.04	22	los términos y condiciones	33	0.03
23	el usuario se compromete	21	0.04	23	registro mercantil de Madrid	33	0.03
24	reserva el derecho de	21	0.04	24	con domicilio social en	32	0.03
25	en el sitio web	20	0.03	25	de privacidad y tratamiento	32	0.03

**Tabla 31. Listado de frecuencia contrastivo por grupos de cuatro palabras en subcorpus legal.**

Si nos centramos primero en los resultados del subcorpus original, esta lista de frecuencias muestra que las formas lingüísticas convencionales se asocian primordialmente a las unidades referentes a la **política de privacidad** (F2). De hecho, ocupan los cuatro primeros puestos de la lista original: «datos de carácter personal», «de protección de datos». Estas formas se corresponden con la terminología y fraseología empleada en la ley española de protección de datos del 1999, mientras que en los textos localizados se aprecia una relativa menor frecuencia de la unidad fraseológica «datos de carácter personal» [puesto 21]. Adicionalmente, los puestos catorce y quince en la lista original se corresponden con la denominación de esta ley española.

Contrastivamente, cabe mencionar que los subtextos localizados se decantan mayoritariamente por unidades fraseológicas asociadas con la palabra «privacidad», tales como «la política de privacidad» [13], «esta política de privacidad» [15], o «política de privacidad y» [16], «privacidad y tratamiento de» [20]. Este hecho demuestra la preferencia clara de los textos originales por emplear la terminología de la ley española, mientras que los textos localizados gravitan mayoritariamente alrededor del término «privacidad» en lo relativo a la sección F2.

La siguiente forma identificada, «se reserva el derecho», ocupa un puesto prominente en ambos subcorpus y se identifica con la sección de condiciones generales (F-3), especialmente con las unidades significativas de modificaciones o cambios de contenidos (F-3-II), modificaciones de las condiciones generales (F-3-III), o la restricción de acceso a posibles usuarios al sitio (F-3-VI).

En tercer lugar predominan las formas lingüísticas asociadas al registro legal de la empresa, «inscrita en el registro», «el registro mercantil de», o «registro mercantil de Madrid», asociadas a la unidad de registro legal (F-1-II), y más específicamente a la subunidad significativa de inscripción de la empresa (F-1-II-a). Finalmente, en cuarto lugar se encuentran las formas lingüísticas asociadas a la unidad significativa sobre los derechos de propiedad intelectual (F-1-VI), que se formula a través de «derechos de propiedad intelectual», recogida en la novena posición en el subcorpus original y la quinta en el localizado.

En una segunda fase analizamos las palabras clave o *keywords* de los textos localizados empleando el subcorpus original como corpus de referencia. Presentamos a continuación los resultados obtenidos de los que hemos eliminado los nombre propios, ya que se corresponden exclusivamente con los nombre de las empresas.

N	WORD	FREQ.	LOCTXT %	FREQ,NATLEG-1, TXT %	KEYNESS	
<b>2</b>	<b>usted</b>	<b>433</b>	<b>0,39</b>	<b>36</b>	<b>0,06</b>	<b>183,3</b>
3	sitio	1.212	1,08	283	0,49	167,3
8	información	1.195	1,06	327	0,57	114,8
11	materiales	145	0,13	10	0,02	67,8
12	filiales	102	0,09	3		63,9
<b>13</b>	<b>privacidad</b>	<b>317</b>	<b>0,28</b>	<b>61</b>	<b>0,11</b>	<b>60,7</b>
<b>17</b>	<b>términos</b>	<b>251</b>	<b>0,22</b>	<b>46</b>	<b>0,08</b>	<b>51,6</b>
20	afiliadas	59	0,05	0		48,9
22	línea	55	0,05	0		45,6
23	otros	284	0,25	61	0,11	45,5
31	sitios	265	0,24	56	0,10	43,7
<b>32</b>	<b>declaración</b>	<b>85</b>	<b>0,08</b>	<b>5</b>		<b>42,7</b>
37	estos	173	0,15	30	0,05	38,3
38	puede	315	0,28	80	0,14	36,1
41	marcas	202	0,18	41	0,07	35,6
42	utilizar	214	0,19	45	0,08	35,6
<b>43</b>	<b>leyes</b>	<b>103</b>	<b>0,09</b>	<b>12</b>	<b>0,02</b>	<b>34,4</b>
44	incluyendo	99	0,09	11	0,02	34,4
46	nuestros	144	0,13	25	0,04	31,8
48	otras	155	0,14	29	0,05	30,9

**Tabla 32. Palabras clave en el subcorpus legal localizado utilizando el subcorpus legal original como corpus de control.**

El dato más significativo de este análisis consiste en la mayor inclusión explícita del pronombre «usted» en los textos localizados. En su conjunto el pronombre «usted» aparece en 433 ocasiones a las que hemos de añadir 57 casos de la abreviación «Ud.», frente a 39 menciones en los textos originales, incluyendo tres formas abreviadas. Estos resultados se corresponderían con un proceso de explicitación en el proceso cognitivo de la traducción (Baker y Olohan 2001), uno de los controvertidos «universales de traducción» presentado por Mona Baker (1995). En este sentido, el carácter legal de estos textos puede conllevar un proceso de traducción más fiel y literal propio de traducciones documentales (Nord 1997a).

Dada la significativa sobrerrepresentación de este pronombre personal en los subtextos localizados legales, procedimos a analizar la frecuencia de otros pronombres en el conjunto global del corpus. Este análisis tuvo como objetivo discernir si este fenómeno es exclusivo de subtextos legales o si, por el contrario, se extiende al resto de subtextos. En nuestro análisis léxico del corpus global, el pronombre «nosotros» presenta una frecuencia de 0,0846 en la sección localizada y 0,0495 en la original, una diferencia del 170,9%. Por su parte, la frecuencia del

pronombre «tú» sería 0,00404 en los localizados y 0,002956 en los originales, de nuevo una cifra considerablemente superior. Por consiguiente, podríamos concluir que esta tendencia no es específica de los textos legales traducidos sino que, en general, los textos localizados presentan una tendencia a incorporar en mayor medida distintas formas pronominales que los originales, un posible fenómeno subconsciente de explicitación (Baker y Olohan, 2000).

Volviendo al análisis contrastivo terminológico recogido en la Tabla 32, es igualmente significativa la presencia mayoritaria de términos como «privacidad», «términos», «declaración», que como observamos anteriormente, son más propios de los textos localizados. Finalmente, hemos de señalar la presencia de la palabra «leyes» en plural, que podría apuntar a un carácter más generalista de estos textos localizados que se dirigen a un conjunto de países y marcos jurídicos.

Pasamos a continuación a revisar las distintas secciones identificadas y las formas lingüísticas convencionales asociadas a las mismas. En nuestro análisis trataremos sólo aquellas unidades y subunidades significativas que hayan presentado una frecuencia igual o superior al 50% en uno u otro componente del corpus (véase Tabla 30).

### **10.3.5. Análisis de unidades y subunidades significativas en la sección avisos legales (F1)**

Tal como presentamos en el análisis contrastivo superestructural de los sitios web corporativos (véase 5.3.1), la sección de avisos legales (F2) es la más frecuente en los textos originales en español (27,90%). Dado su carácter general, en la misma hemos incluido aquellas unidades y subunidades significativas que no se refieren a la protección de datos personales (F2) o las condiciones de uso (F3). No obstante, en los textos originales la sección «avisos legales» tiende a incorporar el conjunto

de unidades significativas básicas aunque pertenezcan a otras secciones. Pasamos a analizar y describir las distintas unidades y subunidades dentro de esta sección.

### **10.3.5.1. Unidad significativa de presentación de avisos legales (F-1-I)**

La unidad de presentación constituye una introducción al contenido del texto legal, en la que predominan la subunidad significativa en la que se apela a la lectura de la página (F-1-I-b), y la subunidad referente a la bienvenida al texto (F-1-I-a), cuya presencia es exclusiva de los textos localizados. La primera es asimismo más frecuente en los textos localizados, el 65,22% frente al 13,21% en originales, con lo que podemos concluir que constituye un reflejo de la estructura de los textos originales estadounidenses.

La subunidad de bienvenida al texto (F-1-I-a) es exclusiva de los textos localizados, y suele expresarse mayoritariamente mediante la locución «bienvenido[s] a [...]», o mediante la unidad léxica «dar la bienvenida». El siguiente sería un ejemplo: «Archer Daniels Midland Company ("ADM") **le da la bienvenida**». Hemos identificado igualmente otras fórmulas de escaso uso que no incorporan la palabra «bienvenida», tal como: «GLASER se alegra<sup>444</sup> de que Ud. Visite esta Página Web y de que muestre interés por sus productos» [Glaser].

En los textos localizados, la subunidad de apelación a la leer el texto se presenta principalmente mediante una forma verbal en modo imperativo: «lea» o «revise» en el 62,55% de los casos, en forma infinitiva el 25,92% de los casos y en forma subjuntiva junto con el verbo «rogar» o la locución «es importante [...]» (11,11%). En el 40,77% de los casos el imperativo se encuentra potenciado por medio de los adverbios «atentamente», «cuidadosamente» o «detenidamente».

---

<sup>444</sup> Inadecuación en la formulación en la lengua de llegada.

Por su parte, en los textos originales es mayoritario el uso de la forma infinitiva, (81,81%), junto con perífrasis verbales en forma imperativa que suavizan la exhortación «deber» y «recomendar» (9,09%), y la forma verbal subjuntiva «rogamos que lea» (9,09%). Los adverbios intensificadores «atentamente» y «detenidamente» complementan estas unidades fraseológicas en el 71% de los casos. En este sentido, el adverbio «cuidadosamente» que identificamos en el 11,11% de textos localizados es inexistente en los textos originales<sup>445</sup>. A modo de ilustración presentamos las líneas de concordancias originales de la forma infinitiva «leer» en la que se observa la presencia de estos adverbios:

N	Concordance
1	ón de dichos Servicios, el usuario ha de leer atentamente también las correspon
2	hos servicios, el Usuario también ha de leer atentamente las correspondientes
3	os servicios, el Usuario también ha de leer atentamente las correspondientes C
4	ontrato: El Usuario se compromete a leer detenidamente el contenido de est
5	. Por tanto el Usuario se compromete a leer atentamente las condiciones de co
6	llo, EUROPCAR recomienda al usuario leer el mismo atentamente cada vez q
7	mente, por lo que los usuarios deberán leer las mismas cada vez que accedan
8	O IBERICA S.A LE ROGAMOS QUE LEA CON ATENCIÓN LOS TÉRMINOS
9	okia. Cuando salga de este Sitio Web, lea la política de privacidad aplicable al
10	eb. En consecuencia, el Usuario debe leer atentamente el presente Aviso Leg
11	incluidas en este Aviso Legal, debiendo leer el Usuario atentamente el present

Figura 96. Extracto de colocaciones del verbo «leer» en subcorpus original legal que ilustra la presencia convencional de adverbios intensificadores.

### 10.3.5.2. Unidad significativa de aceptación de los términos legales (F-1-II)

Los términos legales en los sitios web constituyen textos exhortativos sin alternativa (Hatim y Mason 1990), ya que son vinculantes y el propio uso del sitio web implica su aceptación tácita por parte del usuario. De esta manera, esta subunidad significativa recoge cómo el mero uso del sitio implicaría su aceptación. Además, puesto que el mero uso del sitio web es contingente con esta premisa, en

<sup>445</sup> Constituirían por consiguiente un calco léxico del adverbio *carefully*.



caso contrario se convida al usuario a abandonarlo. Este último aspecto se formula mediante una subunidad significativa específica (F-1-II-b).

El análisis de las formas lingüísticas muestra que la aceptación de los términos (F-1-II-a) está ligada mayoritariamente a unidades fraseológicas que incorporan la palabra «aceptación», y en menor medida su forma verbal, «aceptar». Un ejemplo del corpus natural sería «El mero uso del Portal implica la aceptación sin reservas por parte del usuario de las presentes Condiciones Generales» [viajar.com]. En nuestra identificación de formas lingüísticas convencionales asociadas a esta subunidad hemos observado ciertas diferencias referentes a las posibles formas verbales en el rango colocacional de la palabra «aceptación». Éstas quedan recogidas en la siguiente tabla:

Verbo asociado. Original	Término	Verbo asociado. Localizado	Término
Implicar 52,38% Suponer 28,57% Expresar 4,76% Declarar 4,76% Significar 4,76% Presuponer 4,76%	ACEPTACIÓN	Implicar 30,76% Manifestar 19,23% Constituir 11,53% Presuponer 11,53% Suponer 7,79% Entenderse como 3,84% Quedar supeditada 3,84% Quedar sometida 3,84% Comportar 3,84% Representar 3,84% Expresar 3,84%	ACEPTACIÓN

**Tabla 33. Formas verbales asociadas a la palabra «aceptación» en la unidad significativa (F-1-II).**

Las formas verbales aparecen tanto en tiempo verbal presente como en futuro, con la preferencia en el subcorpus original por el verbo «implicar» (52,38% en originales y 30,76% en localizados). No obstante, el corpus localizado presenta siete verbos asociados a esta palabra que no están presentes en el corpus original, «manifestar», «constituir», «entenderse como», «quedar supeditada», «quedar sometida», «comportar», y «representar», cuya frecuencia combinada asciende al 49,96%.

Si nos centramos en el rango colocacional a la derecha de la base, los textos originales complementan esta unidad fraseológica de la siguiente manera: «aceptación plena y sin reservas», que supondría una forma lingüística convencional asociada a esta subunidad significativa. El siguiente extracto de concordancias demuestra este carácter convencional en los textos originales:

N	Concordance
1	usuario expresa su comprensión y aceptación plena y sin reservas
2	delante, el "Usuario") e implica la aceptación plena y sin reservas
3	El mero uso del Portal implica la aceptación sin reservas por parte
4	. El mero uso del Portal implica la aceptación sin reservas por parte
5	El acceso a dicho sitio implica su aceptación sin reservas. La utiliza
6	El acceso al mismo implica su aceptación sin reservas. • 3. S
7	et. El acceso al mismo implica su aceptación sin reservas. La utiliza
8	en el área de clientes implicará su aceptación expresa y sin reserva
9	n de la Tarjeta A.G.C. supondrá la aceptación plena y sin reservas,
10	n de la Tarjeta A.G.C. supondrá la aceptación plena y sin reservas,
11	. La visita del sitio web supone la aceptación plena y sin reservas

Figura 97. Extracto de concordancias de la palabra «aceptación» en subcorpus legal original que ilustra su fraseología convencionalizada.

Esta co-ocurrencia se observa igualmente en los textos localizados aunque en menor medida:

N	Concordance	Set Tag
1	inados servicios implicará asimismo la aceptación, sin reserva alguna, de las	
2	(en adelante, el "Usuario") e implica la aceptación plena y sin reservas de tod	
3	b, el visitante ("Usuario") manifiesta su aceptación sin reservas de las condi	
4	la página web el Usuario manifiesta su aceptación sin reservas de las present	
5	la aceptación por este sitio Web supone su aceptación sin limitaciones de estos t	
6	express.es" implica el conocimiento y la aceptación, plena y sin reservas, de lo	

Figura 98. Extracto de concordancias de la palabra «aceptación» en subcorpus legal localizado.

Adicionalmente, en el conjunto de colocaciones localizadas hemos identificado dos casos en los que la traducción no habría incorporado esta forma lingüística convencional, «sin reserva alguna» [línea 1] y «sin limitaciones» [línea 5]. De igual manera se ha identificado otro caso en el que se incluye «la total aceptación», que igualmente podría constituir otra inadecuación al no ajustarse a esta forma lingüística convencional.

Pasando a las formas lingüísticas convencionales asociadas a la subunidad de abandono del sitio (F-1-II-b), la formulación más recurrente en las subunidades originales se realiza mediante el verbo «rogar» (75%), y entraña una modulación en esta forma exhortativa. Los siguientes ejemplos ilustran esta forma convencional: «Si no está de acuerdo con ellos, **le rogamos** que no utilice este sitio web» [Jung Electro Ibérica], «si las consideraciones detalladas en el mismo no son de su conformidad, **rogamos** no haga uso del Portal». Esta forma se recoge sólo en el 42,85% de las subunidades localizadas: «Si no los acepta, **le rogamos** que no utilice ni acceda a nuestros servicios» [Autodesk].

El hecho más significativo que arroja el análisis contrastivo sería el gran número de inadecuaciones identificadas en los textos localizados, que asociamos a la falta de modulación en la apelación. De entre las posibles inadecuaciones de expresión en la lengua de llegada se encontrarían:

- «Si no está de acuerdo con los términos aquí descritos, **debería abandonar el site ahora mismo**» [Avaya].
- «**Si está a disgusto** con el Sitio web de Agribands o cualquier contenido del sitio o con estos términos o condiciones, **el único recurso exclusivo que usted tiene es dejar de usar** el Sitio web de Agribands» [Agribands].
- «Si no está de acuerdo con estos términos y condiciones, **debe abandonar este sitio web inmediatamente.**» [Philip Morris International Management S.A.].
- «si usted no acepta estos términos y condiciones, **no está autorizado a usar esta web.**» [Kimberly-Clark].
- «Si no acepta estos Términos y condiciones (que pueden enmendarse de vez en cuando) o está disconforme con este sitio Web, **su único y exclusivo remedio es interrumpir su uso**» [Hospira].

Las inadecuaciones identificadas en los segmentos traducidos no se circunscriben exclusivamente a la falta de modulación en la apelación o inadecuaciones estilísticas, sino que también incorporan errores semánticos, « su único y exclusivo remedio es interrumpir su uso», así como contrasentidos: «si no acepta alguno de estos términos o condiciones, no deberá acceder ni utilizar **de otra manera** este sitio web o la información del mismo» [Genworth Financial]. Este caso supondría

una traducción errónea de *in any way*, cuyo resultado sería un sentido opuesto al del texto original.

### 10.3.5.3. Unidad significativa de registro legal (F-1-III)

El registro legal de la empresa en España le confiere su estatus como entidad legal y, consecuentemente, posibilita el establecimiento de una relación jurídica vinculante entre ambas partes. Esta unidad significativa está presente en el 73,58% de sitios web originales con texto legal frente al 34,78% de localizados. La inclusión de esta unidad significativa, y en especial la subunidad significativa de inscripción de la empresa en el registro mercantil (F-1-III-a), sería uno de los mejores indicadores de un grado de localización máximo (véase 2.7). La presencia de esta unidad podría asociarse con un proceso de adaptación textual en el proceso de localización, aunque con mayor probabilidad, esta unidad sería indicativa de la posible redacción original de este subtexto en la lengua meta. Este hecho apunta a la presencia de subtextos redactados originalmente en español en un sitio web localizado, aun cuando en general formen parte de un sitio web localizado en su conjunto.

En el subcorpus original la subunidad significativa de inscripción de la empresa (F-1-III-a) se expresa mediante una forma lingüística altamente convencionalizada con la fórmula « inscrita en el Registro Mercantil de [...], tomo [...], folio [...], hoja [...], inscripción [...]». El último de los campos, la «inscripción», sería optativa según nuestro análisis. Esta fórmula aparece precedida en el 45% de los casos del adverbio «debidamente»: «**debidamente** inscrita en el Registro Mercantil de Madrid, al tomo 8171, folio 153, hoja M-131861, inscripción 1» [Adecco]. En el caso de los sitios web localizados, y dado que su presencia sería indicativa de la posible redacción original en español, esta forma lingüística aparece igualmente en la gran mayoría de los casos. Se encontró igualmente una alternativa en el uso de

formas convencionales que reza «Compañía inscrita en el Registro Mercantil de Madrid **en el Tomo** 18.606, Folio 191, Hoja M- 323639» [Verisign]. Este uso de «en el tomo» presenta una menor frecuencia de uso, ya que la referencia al tomo de inscripción viene precedida mayoritariamente de una coma, o como alternativa, en los textos originales hemos identificado el uso de la preposición «a» que se concreta en la forma «al tomo» (14,27%).

La siguiente subunidad significativa representativa contiene el Código de Identificación Fiscal (CIF) de la empresa, con un 52,83% de frecuencia en los textos originales y el 30,43% en localizados. Este código cobra una gran importancia en el establecimiento de la empresa como entidad jurídica, si bien en menor medida que la inscripción en el registro mercantil. De hecho, en los textos originales la frecuencia de esta subunidad es superior en un 20,75%. No obstante, su frecuencia es similar en los textos localizados, con un 4,35% de diferencia en la frecuencia de esta misma subunidad entre ambas secciones del subcorpus. Dada su menor importancia, esta subunidad siempre aparece precedida de la subunidad de inscripción en el registro.

En ambos subcorpus, la sigla usada en su formulación aparece mayoritariamente sin puntos en los textos originales, «CIF» (62,5%), frente a la forma que los incluye, «C.I.F.» (37,5%). La forma lingüística asociada a esta subunidad se presenta con la preposición «con» en la exposición de los datos identificativos tras la información de inscripción de la empresa «con C.I.F. A-83319400 y domicilio en Madrid» [Adecco]. Esta forma presenta una muy baja frecuencia en los textos localizados, en donde esta unidad suele aparecer tras los datos de inscripción precedido de una coma o punto, «Hoja M-122.623, inscripción 1ª. C.I.F.: A-80781727 Sociedad Unipersonal» [Genworth Financial].

#### **10.3.5.4. Unidad significativa sobre la legislación aplicable y jurisdicción (F-I-IV)**

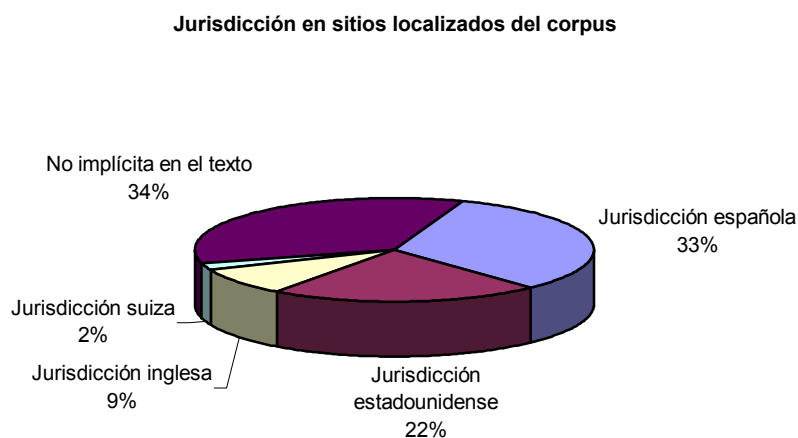
La localización de sitios web multilingües con un modelo centralizado entraña la producción de un sitio web único del que parten las localizaciones a distintas regiones lingüísticas o *locales* (Yunker 2003). No obstante, resulta evidente que existen diferencias considerables entre las legislaciones de las diversas áreas hacia las que se dirigen. Estas diferencias requerirían una adaptación a cada realidad jurídica de estos subtextos y, por ende, el establecimiento del fuero que gobierne cualquier relación jurídica entre ambas partes. Este particular no constituye una causa de conflicto en los sitios originales españoles, ya que la legislación aplicable sería implícitamente la española. Por el contrario, el fuero o jurisdicción aplicable en los sitios web multilingües localizados ha de ser consensuado, aunque hemos constatado que en su defecto tienden a ser gobernados por la legislación vigente en la región del TO. Esta posible discrepancia entre ambos subcorpus queda patente en la reveladora diferencia de frecuencia en esta unidad significativa, 18,87% en textos originales frente a 65,22% en localizados. Esta diferencia en cuanto a frecuencia constituye una de las mayores del subcorpus, 46,35%, fiel reflejo de este hecho.

Otra diferencia principal entre ambos subcorpus es la inclusión en los subtextos originales de los tribunales competentes dentro del marco geográfico español, mientras que en los localizados aparecen menciones a la jurisdicción de países o regiones<sup>446</sup> específicas. Al igual que la sección anterior, esta sección sería indicativa de un mayor grado de localización, especialmente en el caso de que la jurisdicción competente fuese la española. Este establecimiento de la jurisdicción conlleva la

---

<sup>446</sup> Como es el caso de la jurisdicción de los distintos estados en los Estados Unidos o a Québec en Canadá.

posible inclusión de una subunidad que justifique tanto esta selección como su consecuente aceptación por parte del usuario<sup>447</sup>.



**Figura 99. Jurisdicción en sitios web localizados del corpus.**

Si analizamos la elección de jurisdicción en los subtextos localizados, el 32,60% de sitios se rigen por la legislación española, el 21,73%<sup>448</sup> por la estadounidense, el 8,69% por la inglesa, y el 2,17% por la suiza. El 34,78% restante de textos no especifican la jurisdicción aplicable. En cualquier caso, se han identificado errores de traducción referentes a la falta de explicitación en el origen de ciertas leyes mencionadas, tal como «Es intención de Colgate la de adherirse al **Acta de Protección del Menor** y sus reglas cuando se recoja información personal sobre los menores» [Colgate].

En el seno de esta unidad se han identificado dos subunidades significativas de interés. Una de ellas se refiere al sometimiento de ambas partes a una jurisdicción

<sup>447</sup> «Reconocemos que es posible que usted acceda a este sitio desde cualquier jurisdicción del mundo pero no podemos evitar dicho acceso. Este sitio ha sido diseñado de conformidad con la legislación de Suiza. Si cualquier material de este sitio o el uso del sitio infringiera la legislación del país en el que se encuentra en el momento de acceso, el sitio no estará dirigido a usted y le rogamos que no lo utilice» [Phillip Morris International].

<sup>448</sup> Por orden de frecuencia aparece la legislación de los tribunales de los estados de California, Illinois, Delaware, Washington, Georgia, Minesota, y la federal de EEUU.

específica (F-I-IV-a), y una segunda subunidad significativa dedicada a la renuncia al fuero propio (F-I-IV-b).

Las formas lingüísticas convencionales asociadas a esta unidad que hemos identificado están relacionadas con ciertas unidades fraseológicas que incluyen el adjetivo «aplicable», que complementa los sustantivos «ley», «legislación», «leyes», o «jurisdicción». La frecuencia de este adjetivo es significativamente mayor en el subcorpus localizado (0.11%) que en el original (0.06%). A continuación presentamos un extracto de las distintas líneas de concordancias en textos originales que ilustran el rango colocacional de esta palabra.

1. varse de los mismos. 9. FUERO Y DERECHO **APLICABLE** Las Condiciones Básicas y las O  
2. 30 días. 19. LEGISLACIÓN Y JURISDICCIÓN **APLICABLE** Las presentes condiciones se rig  
3. das las menciones requeridas por la legislación **aplicable** en materia de protección de datos de  
4. ----- LEGISLACIÓN **APLICABLE** – TRIBUNAL COMPETENTE  
5. adora. 5. LEY **APLICABLE** La ley **aplicable** en caso de disputa o conflicto de inte  
6. o cualquier otro derecho de terceros. \*VII. Ley **aplicable** y jurisdicción\* Para cuantas cuestio  
7. egador a tal efecto. Jurisdicción y legislaciÓn **aplicables** - Los tÉRminos y condiciones que

**Figura 100. Extracto de concordancias de la palabra «aplicable\*» en subcorpus legal original.**

Para su comparación presentamos las posibles concordancias de esta misma palabra en el subcorpus localizado:

1. zar este Sitio Web de conformidad con la legislación **aplicable**, la presente Declaración Legal, el orden pú  
2. LA MEDIDA QUE LO PERMITA LA LEGISLACIÓN **APLICABLE**, NO SEREMOS RESPONSABLES, B  
3. tables a BMS. Ley aplicable y jurisdicción La ley **aplicable** en caso de disputa o conflicto de interpreta  
4. ER MEDIANTE ESTE SITIO. Volver a la lista Ley **aplicable** y jurisdicción Estos tÉRminos y condicio  
5. r a una orden judicial o cuando lo requieran las leyes **aplicables**, las sentencias de los tribunales o los regl  
6. dida por estos tÉRminos del Sitio y puede violar leyes **aplicables**, incluyendo las leyes de Propiedad Intelec  
7. cceso y/o utilización cumple con la legislación local **aplicable**. La presente Declaración Legal regula el  
8. á responsable del cumplimiento de la normativa local **aplicable**. Estos Términos se regirán e interpretarán  
9. tos, programas y servicios disponibles. La normativa **aplicable** podría no permitir la exclusión de garantías

**Figura 101. Extracto de concordancias de la palabra «aplicable\*» en subcorpus legal localizado.**

Podemos observar que en ambos subcorpus la colocación «legislación y jurisdicción aplicable» constituiría una forma lingüística convencional asociada a esta subunidad. En ambos subcorpus la frecuencia de colocación de la palabra



«aplicable» con «ley» y «legislación» sería similar entre ambas palabras e igualmente similar en entre ambos subcorpus.

La subunidad significativa identificada referente al sometimiento de ambas partes a una jurisdicción específica (F-I-IV-a) se encuentra relacionada con las unidades fraseológicas «someterse a la jurisdicción» y «regirse por la legislación». El análisis contrastivo de la primera de estas formas lingüísticas convencionales muestra una mayor frecuencia en los textos originales, así como preferencia absoluta por el tiempo verbal presente en su formulación.

N	Concordance
1	esa es consciente de ello, esto no significa que se <b>someta</b> a <b>las</b> jurisdicciones de países extranjeros;
2	vase del acceso o la utilización de este website se <b>somete</b> a <b>la</b> exclusiva jurisdicción de los tribunales
3	ualquier otro fuero que pudiera corresponderles, se <b>someten</b> a <b>la</b> Jurisdicción de los Juzgados y Tribun
4	ualquier otro fuero que pudiera corresponderles, se <b>someten</b> a <b>la</b> Jurisdicción de los Juzgados y Tribun
5	ualquier otro fuero que pudiera corresponderles, se <b>someten</b> a <b>la</b> Jurisdicción de los Juzgados y Tribun
6	MIGUEL se rigen por la legislación española y se <b>someten</b> a <b>la</b> jurisdicción de los Juzgados y Tribun
7	ualquier otro fuero que les pudiera corresponder, se <b>someten</b> a <b>la</b> jurisdicción y competencia exclusiva
8	ia expresa al fuero que pudiera corresponderles, se <b>someten</b> expresamente <b>a</b> la jurisdicción de los Juz
9	ia expresa al fuero que pudiera corresponderles, se <b>someten</b> expresamente <b>a</b> la jurisdicción de los Juz
10	rios, particularmente aquellos de tipo legal, pueden <b>someterse</b> a <b>distintos</b> criterios y no ser compartido
11	er otro fuero que pudiera corresponderles, acuerdan <b>someterse</b> a <b>la</b> jurisdicción de los Juzgados y Tribu
12	fuero que pudiera corresponder, el Usuario acuerda <b>someterse</b> a <b>la</b> Jurisdicción de los Juzgados y Trib

Figura 102. Extracto de concordancias del verbo «someter» en subcorpus legal original.

En el caso de los textos localizados aparecen formas verbales futuras en su expresión (37,5%). En la formulación de esta forma lingüística convencional aparece el adjetivo «exclusiva» y el adverbio «exclusivamente», siendo posible además la aparición de «expresa» o «exclusiva».

N	Concordance
1	ente usted irreversiblee incondicionalmenteconsiente y <b>somete</b> a <b>la</b> jurisdicción exclusiva de tales cortes toda
2	BLE Los usuarios de American Express de España se <b>someten</b> a <b>la</b> exclusiva jurisdicción de los tribunales
3	uerdo con las leyes de España. Stryker y el Usuario se <b>someten</b> expresamente <b>a</b> la jurisdicción exclusiva de lo
4	Johnson, no sea apropiado. *Varios*. Este acuerdo se <b>someterá</b> a <b>la</b> legislación vigente en España. Los usuar
5	no necesariamente limitadas a temas contractuales) se <b>someterán</b> a <b>la</b> jurisdicción exclusiva de los tribunales i
6	ción de España. Las disputas que pudieran resultar se <b>someterán</b> exclusivamente <b>a</b> la jurisdicción de los juzga
7	a legislación vigente en España. Los usuarios acuerdan <b>someterse</b> a <b>la</b> jurisdicción de los tribunales compet
8	ÑA, S.A. se sustanciará ante la jurisdicción española, <b>sometiéndose</b> las <b>partes</b> a los Juzgados y Tribunales

Figura 103. Extracto de concordancias del verbo «someter» en subcorpus legal localizado.

La segunda unidad fraseológica convencional en esta unidad significativa sería «regirse por la legislación». En ambos casos esta unidad suele aparecer tanto en tiempo verbal presente como en futuro, con una mayor preferencia por el tiempo verbal futuro en los textos localizados (68,42%) que en los originales (36,36%). El primero de los extractos de concordancias ilustra las distintas formas verbales que se usan en los textos originales:

1. ente al Usuario, que deberá **regirse** por los términos, cond
2. términos y condiciones que **rigen** el Acuerdo de Licencia
3. ios ofrecidos en este sitio **se regirá** además por las condi
4. es Condiciones Generales **se regirán** por la legislación espa
5. El presente Aviso Legal **se rige** en todos y cada uno de
6. e MAHOU- SAN MIGUEL **se rigen** por la legislación espa

**Figura 104. Extracto de concordancias del verbo «regir» en subcorpus legal original.**

La formulación más recurrente en los textos originales españoles sería por tanto «se regirán por la legislación española», siendo posible en menor medida la sustitución de «legislación» por «ley» o «normativa». El contraste de esta forma convencional muestra que es inexistente en los textos localizados, apareciendo la forma «se regirán por las Leyes española [de España]» en tres casos.

1. ===== **Rigen** las Condiciones Generales
2. cuestionarios en el Sitio web **se regirá** de conformidad con la Polí
3. ón de la Ley\* Este Acuerdo **se regirá** e interpretará según las le
4. ntes Términos y Condiciones **se regirán** e interpretarán de acuerdo
5. o a las leyes aplicables, que **se rigen** e interpretan de acuerdo co
6. ondiciones Generales de Uso **se rigen** por las Leyes españolas.
7. Estas Condiciones de Uso **rigen** su uso del Sitio Web y de t

**Figura 105. Extracto de concordancias del verbo «regir» en subcorpus legal localizado.**

En las concordancias obtenidas hemos observado una unidad fraseológica recurrente en los textos localizados, «se regirán e interpretarán», y que no aparece en nuestro subcorpus original. Su frecuencia asciende al 36,84%, con lo que se podría identificar como una unidad fraseológica calco de los TO en inglés. En lo que respecta a las colocaciones del verbo «regir», éste se encuentra seguido de la

preposición «por» en el 81,81% de los casos, mientras que la preposición empleada en los textos localizados presenta una mayor variación: el 42,1% incluye la preposición «por», 21,01% de «de acuerdo con [a]», 15,78% de «de conformidad con», y 10,52% de «según». Queda así patente una mayor variación en la formulación de esta unidad fraseológica entre ambos subcorpus, tanto en el tiempo verbal en que se expresa como en la menor fijación de sus elementos en los textos localizados.

Por último encontramos la subunidad significativa de renuncia al fuero propio (F-I-IV-b). Ésta presenta una mayor frecuencia en los textos originales y se encuentra asociada mayoritariamente a la forma lingüística convencional «con renuncia expresa a [cualquier] fuero que pudiera corresponder». En menor medida aparece la forma convencional «con renuncia a su propio fuero». Este carácter de convencional se puede observar en las líneas de concordancia que presentamos:

N	Concordance
1	resentes Condiciones Generales, las partes, con <b>renuncia</b> expresa al fuero que pudiera correspon
2	resentes Condiciones Generales, las partes, con <b>renuncia</b> expresa al fuero que pudiera correspon
3	os Juzgados y Tribunales de Madrid Capital, con <b>renuncia a</b> cualquier fuero que pudiera correspon
4	DE CONTROVERSIAS Las partes, con expresa <b>renuncia a</b> su propio fuero, aceptan como legisla
5	SLACIÓN APLICABLE. Las partes, con expresa <b>renuncia a</b> su propio fuero, aceptan como legisla
6	• 7. Xperience Consulting S.L. se someten, con <b>renuncia expresa</b> a cualquier otro fuero, a los
7	a, Telefónica S.A. y el usuario se someten, con <b>renuncia expresa</b> a cualquier otro fuero, a los ju
8	islación española. SECURITAS y el usuario, con <b>renuncia expresa</b> a cualquier otro fuero que pudi
9	entrada. 11. Telefónica S.A. y el usuario, con <b>renuncia expresa</b> a cualquier otro fuero, se som
10	fuera de España, SECURITAS y el Usuario, con <b>renuncia expresa</b> a cualquier otro fuero que pudie
11	dispuesto en las presentes disposiciones, y con <b>renuncia expresa</b> a cualquier otro fuero que pudie
12	es o de su visita al sitio web, ambas partes, con <b>renuncia expresa</b> a cualquier otro fuero que les p
13	pañola. Sistema AZUD, S.A. y el usuario, con <b>renuncia expresa</b> a cualquier otro fuero que pudi
14	da en que así lo permita la ley, las partes, con <b>renuncia expresa</b> a cualquier otro fuero que pudi
15	e entenderá por parte de YLOS que el cliente ha <b>renunciado a</b> su pedido. En el supuesto normal
16	e en relación con el Portal, las partes acuerdan, <b>renunciando</b> , en su caso, al fuero que pudiera c

**Figura 106. Extracto de concordancias de lema «renunc\*» en subcorpus legal original.**

Esta última unidad fraseológica, «con renuncia expresa a [cualquier] fuero que pudiera corresponder», presenta una escasa representación en los textos localizados, ya que por un lado la palabra «fuero» presenta una frecuencia significativamente menor, 0,002% frente a 0,3% en el subcorpus original.

N	Concordance
1	sus superiores jerárquicos, con expresa renuncia a otros fueros si lo tuvieren y f
2	O DE GEORGIA, ESTADOS UNIDOS, RENUNCIANDO A OBJETAR AHORA O
3	resente, siempre que dicha exclusión o renuncia sea permisible bajo la normativ
4	Tribunales de la Ciudad de Madrid con renuncia a su propio fuero, si éste fuese
5	de la *Firma PwC*, las partes acuerdan, renunciando a su fuero propio si lo tuvie
6	S PROVEEDORES Y LICENCIANTES RENUNCIAN, EN LA MEDIDA QUE LA

**Figura 107. Extracto de concordancias de lema «renunc\*» en subcorpus legal localizado.**

El análisis de colocaciones en el subcorpus localizado del lema [renunc\*], muestra además que dos de los casos encontrados incorporan a la unidad fraseológica: «con expresa renuncia a otros fueros» o «con renuncia a su propio fuero». A pesar de poder clasificarlas como convencionales en los textos originales, las mismas no incluyen la parte final de la unidad fraseológica que hemos identificado en los textos españoles, «que les pudiera corresponder», e incorporan otras formas no convencionales como «si lo tuvieren» o «si éste fuese otro».

### 10.3.5.5. Unidad significativa sobre los derechos de propiedad intelectual (F-I-V)

Los derechos de propiedad intelectual sobre los contenidos web se encuentran debidamente amparados por leyes nacionales e internacionales, de ahí que esté altamente convencionalizada la presencia de una unidad significativa que recoja tal información. De hecho, esta subunidad presenta la mayor frecuencia en los textos legales localizados (89,13%), y la tercera en los textos originales (67,92%). Esta alta frecuencia de aparición contrasta con su escasa presencia en forma de página independiente en los textos originales (2,90%) (Véase 9.1). Este particular nos lleva a concluir que los sitios web españoles muestran una clara tendencia por incorporar las consideraciones sobre los derechos de propiedad intelectual en el seno de los **avisos legales** (F1). Por su parte, los textos localizados presentan en el

17,89% de los casos una sección diferenciada, confirmándose la mayor importancia de esta unidad presente en estos textos localizados.

Original	Localizado
<b>Protegido por los correspondientes derechos</b>	<b>Protegido por los derechos correspondientes</b>
-los derechos de propiedad intelectual [3] -los derechos de autor [2] -los derechos -los correspondientes derechos de propiedad intelectual e industrial	-por los [correspondientes] derechos de propiedad intelectual [4] - <u>por</u> derechos de autor [3] -por los correspondientes derechos de reproducción, patente, etc. - <u>como derechos de autor</u> por la legislación en materia propiedad intelectual - <u>tiene derechos protegidos</u>
<b>Protegido por ley-leyes</b>	<b>Protegido por ley-leyes</b>
-por las leyes de Propiedad Intelectual [2] -por las leyes españolas e internacionales sobre propiedad intelectual -por ley -conforme a lo dispuesto en la Ley de Propiedad Intelectual	-por la ley de protección de la imagen comercial -por las leyes de copyright -por las leyes de derechos de autor de EEUU y otros países -por las disposiciones de ley y tratados sobre derechos de autor -por las diversa leyes y los tratados internacionales en materia de propiedad intelectual -por las leyes en materia de derechos de autor
<b>Protegido por la legislación</b>	<b>Protegido por la legislación</b>
-la legislación nacional e internacional en función de la legislación vigente sobre propiedad Intelectual nacional e internacional -por la legislación vigente en materia de marcas -en virtud de la legislación aplicable en materia de -propiedad Intelectual	-la legislación <u>aplicable a la materia</u> -la legislación aplicable en materia de marcas -la legislación internacional -la legislación de Estados Unidos
<b>Protegido por [otros]</b>	<b>Protegido por [otros]</b>
-por las normas nacionales e internacionales de propiedad intelectual	-por la normativa y jurisprudencia en materia de propiedad intelectual - <u>contra</u> su reproducción, imitación, dilución, ... - <u>por copyright</u>

**Tabla 34. Tabla contrastiva de las colocaciones de la palabra «protegido» en subcorpus legal.**

El análisis de las formas lingüísticas convencionales muestra que su contenido gravita en torno a la protección jurídica del conjunto de contenidos, la prohibición de su reproducción, modificación, o uso sin la necesaria autorización por escrito (F-I-IV-a), así como la especificación de una lista no exhaustiva de los materiales que se encuentran protegidos (F-I-IV-b). Esta última subunidad presenta una

frecuencia del 26,42% en los textos originales y 50% en los textos localizados, por lo que esta subunidad significativa es recurrente en éstos últimos.

Hemos identificado dos formas lingüísticas convencionales asociadas a esta unidad significativa. En un primer lugar, encontramos «está protegido por [...]», que expresa la protección de los contenidos del sitio web en virtud de la legislación vigente en materia de propiedad intelectual. El análisis contrastivo muestra una formulación similar de esta forma lingüística, que presenta cuatro posibles sustantivos o compuestos sustantivos que la completan: «derechos de propiedad intelectual», «ley-leyes en materia de propiedad intelectual», «legislación», y «normas-normativa». En ambos corpus la forma lingüística de mayor frecuencia de uso es «protegido por los derechos de propiedad intelectual». En segundo y tercer lugar se encuentran las unidades fraseológicas «protegido por los derechos de autor» y «protegido por las leyes de propiedad intelectual».

En los textos originales se manifiesta la influencia de la terminología de la legislación vigente. En España este aspecto está regulado por la Ley de Propiedad Intelectual del 22 de junio del 2006, y esta unidad fraseológica se refleja en los textos originales: el 52,38% de las formas identificadas incluyen el compuesto nominal «propiedad intelectual», frente al 31,81% de los textos localizados. El contexto de producción de los textos localizados se refleja tanto en la mención expresa de leyes estadounidenses en la materia como por la mención de tratados internacionales sobre derechos de propiedad intelectual. Este carácter de localización se manifiesta además por la aparición de ciertos errores de traducción subrayados en la Tabla 34, como la inclusión del préstamo «derechos de copyright». Otro hecho diferenciador entre las formas lingüísticas convencionales de ambos subcorpus es la mayor variación estilística en el subcorpus original, siempre dentro de las posibilidades disponibles en la fraseología jurídica española como: «en virtud de», «conforme a la legislación», o «en función de la legislación».

Estas tres formas que denotan esta posible variación no aparecen en el corpus localizado.

La segunda de las formas lingüísticas convencionales que identificamos está asociada a la prohibición del uso o reproducción de los contenidos a menos que se cuente con la autorización expresa y por escrito de la empresa propietaria. Ésta se formula por medio de «queda [...] prohibido-a», con la posible incorporación recurrente de los siguientes adverbios intensificadores: «expresamente», «estrictamente», «totalmente», y «terminantemente». La presencia de estos adverbios en las unidades fraseológicas queda patente en el extracto de colocaciones originales:

1 nal vigente. Queda absolutamente **prohibido** el uso del contenido de  
2 macional. Queda estrictamente **prohibido** la utilización de todos l  
3 ndustrial. Queda expresamente **prohibida** la reproducción total o p  
4 en la nota de copyright. Queda **prohibida** la reproducción, retrans  
5 titulares. Queda terminantemente **prohibida** la utilización de tales e  
6 quedando por lo tanto, totalmente **prohibida** su utilización con objeti

**Figura 108. Extracto de concordancias del lema «prohib\*» en subcorpus legal original.**

Su presencia puede observarse también en el siguiente extracto de colocaciones perteneciente al subcorpus localizado:

1 ual e Industrial"). Queda estrictamente **prohibido** la utilización de todos los elem  
2 presente Acuerdo, queda expresamente **prohibido** y nada de lo especificado o imp  
3 con fines colectivos o lucrativos. Queda **prohibido** cualquier uso de estos material  
4 ontrato de Licencia de Software. QUEDA **PROHIBIDA** LA REPRODUCCIÓN,  
5 nal, sin intenciones comerciales. Queda **prohibido** copiar, reproducir, publicar,  
6 demás disposiciones legales, quedando **prohibido** copiarlos o imitarlos en todo o  
7 i la Información. Queda terminantemente **prohibido** cualquier otro acceso o uso co

**Figura 109. Extracto de concordancias del lema «prohib\*» en subcorpus legal localizado.**

En los textos originales predomina el adverbio «expresamente», mientras que en los textos localizados es más frecuente el uso del adverbio «estrictamente». Nuestro análisis de las líneas concordancias muestra que los textos originales predomina la prohibición sobre la «utilización», «uso», o «reproducción». Éste último término se

presenta una baja frecuencia en el subcorpus localizado con sólo tres casos, siendo más común la aparición del sustantivo «copia», o su forma verbal «copiar». Es de interés el predominio de formas sustantivas en los textos originales tras el núcleo de la unidad fraseológica, tal como «queda prohibida la reproducción», mientras que en los textos localizados son mayoritarias las formas verbales como «copiar» o «utilizar». Estas formas verbales son inexistentes en los textos originales.

#### **10.3.5.5.1. Análisis de subunidad significativa sobre materiales protegidos por los derechos de propiedad intelectual (F-I-VI-a)**

Esta subunidad significativa es la única de carácter recurrente dentro de la unidad significativa sobre los derechos de propiedad intelectual, y contiene una lista no exhaustiva de los contenidos web protegidos según la legislación vigente. Esta subunidad presenta una frecuencia del 50% en los textos localizados y 26,42% en originales y, por consiguiente, constituiría una subunidad significativa recurrente en los primeros y ocasional en los segundos.

En los textos originales la forma lingüística convencional de mayor frecuencia es: «Todos los contenidos [material] de este sitio web [portal]» (64,28%). Su formulación en plural es mayoritaria, aunque se recoge además y con una menor frecuencia la forma en singular «todo el contenido». Esta forma lingüística convencional se encuentra presente en el 21,73% del conjunto de subunidades identificadas en los textos localizados, con una mayor incidencia de la forma singular de esta unidad fraseológica al contrario de lo que ocurre en las formas originales. Asimismo, la formulación de esta forma convencional en los textos localizados muestra una tendencia a una mayor variación como «Todo el contenido disponible en el Sitio Web, incluyendo, a título meramente enunciativo» [Estee Lauder], o a su formulación inadecuada como «Todo el contenido que vea, escuche



o de otra forma experimente en el Sitio web de Agribrands [...] incluyendo de manera enunciativa mas no limitativa [...]» [Agribrands].

Los contenidos de los listados que incluyen los materiales protegidos se corresponden entre ambos subcorpus. A modo ilustrativo presentamos una lista de los contenidos a proteger que hemos recopilado de esta subunidad organizados de mayor a menor frecuencia: (1) imágenes, (2) contenidos, (3) gráficos, (4) textos, (5) sonido, (6) diseños, (7) fotografías, (8) informaciones, (9) logotipos, (10) archivos, (11) códigos fuente, (12) datos, (13) dibujos, (14) iconos, (15) software, (16) fuentes, (17) música, o (18) videos. En los textos localizados se identificaron además los siguientes componentes adicionales: audio clips, combinación de colores, disposición contenidos, la selección de elementos o colores, ilustraciones, interfaz y melodías. Contrastivamente, nos es de interés la presencia mayoritaria de la forma singular de los sustantivos «texto» y «contenido» en los textos localizados, mientras que esta tendencia es la inversa en los textos originales con un uso mayoritario de la forma en plural.

#### **10.3.5.6. Unidad significativa sobre el titular del sitio web (F-I-VII)**

La identificación del titular del sitio web constituye un imperativo legal dado el carácter contractual de estos textos. Este aspecto se encuentra debidamente recogido en el artículo 10 de la Ley 34/2002, del 11 de julio de Servicios de la Sociedad de la Información y Comercio Electrónico. En los textos originales esta unidad suele expresarse por medio de los datos de registro legal y, por tanto, se encuentra relacionada con la unidad significativa de inscripción (F-I-III). De hecho, según la legislación española la presencia de esta información sobre el registro legal cumple el requerimiento relativo a la identificación del emisor.

En los textos originales las formas lingüísticas convencionales asociadas a esta unidad significativa gravitan en torno a las siguientes unidades fraseológicas: «es titular de» (40%), « [la empresa] pone a disposición de los usuarios de Internet [el sitio]» (25%), «en cumplimiento de la Ley 34/2002» (15%), «es propiedad de» (15%), y «fue creado por» (5%). A excepción de la forma recurrente «pone a disposición de los usuarios de Internet», el resto se encuentran presentes en los textos localizados si bien con una menor frecuencia. De hecho, si sumamos las frecuencias de las dos formas que consideramos convencionales en los textos originales, «es titular de» y «pone a disposición de los usuario de Internet», ésta ascendería al 65%, mientras que las mismas forman parte sólo del 4,34% de unidades localizadas. Ambas se suelen encontrar junto con los datos de registro legal de la empresa: «**El titular de este sitio web es Ingenieros Asesores S.A. C.I.F. A-33062407, con domicilio en el Parque tecnológico de Asturias**» [Ingenieros Asesores], y «Estas condiciones generales regulan el acceso y utilización del sitio web telefonica.es que Telefónica **pone gratuitamente a disposición de los usuarios de Internet.**» [Telefónica].

Unidad fraseológica	% Original	% Localizado
...es titular de – son titularidad de...	40%	4,34%
[empresa] pone a disposición de los usuarios de Internet	25%	0%
...en cumplimiento de la Ley	15%	4,34%
...propiedad de...	15%	26,08%
...fue creado por ...	5%	4,34%
...es operado ...	0%	17,39%
...gestiona-es gestionado...	0%	13,04%
...mantiene...	0%	13,04%
...es de...	0%	8,69%
...es administrado...	0%	4,34%
[Empresa] dirige el sitio...	0%	4,34%

Tabla 35. Tabla contrastiva de unidades fraseológicas asociadas a la unidad significativa sobre el titular del sitio web (F-I-IV).

La forma de mayor frecuencia en los textos localizados la constituye « [...] es propiedad de [...]», (26%) seguido de «es operado por» (17,39%). Esta última unidad fraseológica no figura en el subcorpus original y podría considerarse un

préstamo léxico y sintáctico. Como se demuestra en la tabla contrastiva, los textos localizados muestran una mayor variación en la formulación de las formas lingüísticas asociadas a esta unidad significativa: el 60,84% de frecuencia se concreta en formas que no se encontraron en los textos originales, tales como «es operado por», «es gestionado», o « [el sitio] es de [nombre de empresa]».

De entre los errores de traducción identificados en la expresión de esta unidad significativa, destacan errores sintácticos como «Este sitio **está operado por** Philip Morris International Management S.A.» [Philip Morris International]. En estos textos es igualmente significativa la falta de homogeneidad, al referirse al sitio web indistintamente como «sitio» o « página web»<sup>449</sup>. Este último uso sería inadecuado ya que indicaría que la titularidad se refiere exclusivamente a la página web presente en pantalla, en lugar del sitio web en su conjunto: «Esta **página Web** es de Mattel, Inc.», «**Esta website (el "Sitio")** es administrada por Chevrolet España S.A.». Otro error sintáctico de expresión relacionado con la propiedad del sitio web sería: «Esta Web **está bajo la propiedad**, el control y la gestión de...» [Kimberly-Clark].

### **10.3.6. Análisis de unidades y subunidades significativas en la sección de privacidad (F2)**

Como apuntamos anteriormente, el análisis de las unidades y subunidades significativas en esta sección partió de la investigación previa sobre las políticas de privacidad de sitios web estadounidenses de Liu y Arnet (2000), en la que se identificaron doce posibles unidades. A éstas les añadimos una unidad adicional referente al cumplimiento de las leyes españolas pertinentes (F-2-I). Del conjunto de unidades y subunidades identificadas, dos unidades significativas y una subunidad significativa presentan una frecuencia superior al 50% y, por consiguiente, son de carácter recurrente. Por un lado se encuentra la unidad

---

<sup>449</sup> El término «página web» se emplea en el 17% de las unidades significativas localizadas sobre titularidad identificadas.

referente a la información legal sobre enlaces que surgen o se dirigen hacia el sitio web (F-2-IV) y, por otro, la unidad significativa que ofrece la información de contacto para ejercer los derechos de oposición, modificación o cancelación (F-2-IV). Ambas constituirían unidades significativas recurrentes sólo en los textos localizados. La única subunidad significativa recurrente en los textos originales la constituiría la mención expresa al cumplimiento de lo estipulado en la Ley de Protección de Datos de Carácter Personal (Ley orgánica de 15/1999) que pasamos a revisar.

#### **10.3.6.1. Unidad significativa sobre el cumplimiento de la ley española, LOPD, del 15/1999 (F-2-I-a)**

La subunidad significativa referente al cumplimiento de esta ley presenta una frecuencia del 49,06% en los textos originales y 13,04% en los textos localizados. Como ya apuntamos, la presencia de esta subunidad suele constituir un indicador del alto grado de localización del sitio web en su conjunto. La formulación de esta subunidad podría considerarse en sí misma una creación original en español, de ahí que los segmentos textuales asociados a la ésta supondrían subtextos redactados originalmente en español y no el producto de un proceso de traducción en sí. Consecuentemente, dado este mismo carácter de producción original en español en ambos subcorpus, hemos de argumentar que no sería pertinente su análisis lingüístico contrastivo para nuestros objetivos.

En cualquier caso, la unidad fraseológica de mayor frecuencia en los textos originales según el análisis de las formas lingüísticas es: « [...] de acuerdo con/de conformidad con lo establecido en la Ley Orgánica 15/1999, [de 13 de diciembre], de Protección de Datos de Carácter Personal». Existen asimismo gran variación en la posibilidad de expresión de esta subunidad como «a tenor de lo establecido en», «en cumplimiento de lo dispuesto en», «a efectos de lo previsto en». El siguiente

extracto de concordancias concordancia extraídas del subcorpus original muestra la posible variación fraseológica en su expresión:

1 do ello de conformidad con lo establecido en la Ley Orgánica 15/1999, de 13 de diciembre. 8  
2 y cumpliendo todos los requisitos que obliga la Ley Orgánica de Protección de Datos Persona  
3 Privacidad Renfe-Operadora cumple con la Ley Orgánica 15/99, de protección de datos d  
4 stán ajustados a la normativa establecida en la Ley Orgánica 15/1999,13 de diciembre, de Pro  
5 or medios electrónicos, son conformes a la Ley Orgánica 15/1999, de 13 de diciembre, de  
6 s: Europastry, S.A. De conformidad con la Ley Orgánica 15/1999 de Protección de Datos  
7 ntos sitios en Internet. En cumplimiento de la Ley Orgánica 15/1999 de 13 de diciembre de P  
8 Privacidad, de acuerdo con los principios de la Ley Orgánica 15/1999, de 13 de diciembre, de  
9 S se regirá por las estipulaciones fijadas en la Ley Orgánica 15/1999, de 13 de diciembre, de  
10 personales, de acuerdo a lo establecido en la Ley Orgánica 15/1999, de 13 de Diciembre, de  
11 ersonal\* De acuerdo con lo establecido en la Ley Orgánica 15/1999, de 13 de diciembre, us  
12 TOS\* En cumplimiento en lo dispuesto en la Ley Orgánica 15/1999 de 13 de Diciembre, so  
13 nal: En cumplimiento de lo establecido en la Ley Orgánica 15/1999, de 13 de diciembre, d  
14 De conformidad con lo establecido por la Ley Orgánica 15/1999, de 13 de diciembre,  
15 ales.\* En cumplimiento de lo dispuesto en la ley orgánica de protección de datos de caráct  
16 DE PRIVACIDAD A efectos de lo previsto en la Ley Orgánica de Protección de Datos de cará

**Figura 110. Extracto de concordancias del término «ley orgánica» en el subcorpus original legal.**

El análisis de las formas lingüísticas en el corpus localizado muestra el carácter de producción original en español de estos segmentos textuales, ya que todas las formas identificadas se encuentran recogidas en el subcorpus original. Las siguientes líneas de concordancia del subcorpus localizado ilustrarán este aspecto:

1 de dichos datos. De acuerdo a la Ley Orgánica 15/1999 de Protección  
2 os confidencialmente conforme a la Ley Orgánica 15/1999, de 13 de dici  
3 te sitio Web, de conformidad con la Ley Orgánica 15/1999, de 13 de dic  
4 XEROX, de conformidad con la Ley Orgánica 15/1999 de 13 de dici  
5. Privacidad BMS cumple con la Ley Orgánica 15/1999, de protección  
6 de las disposiciones recogidas en la Ley Orgánica 15/1999, de 13 de Di  
7 cumplimiento de lo dispuesto en la Ley Orgánica 15/1999 de 13 de dici

**Figura 111. Extracto de concordancias del término «ley orgánica» en el subcorpus legal localizado.**

### **10.3.6.2. Unidad significativa sobre enlaces hacia y desde el sitio web (F-2-IV)**

Esta unidad regula el uso de enlaces o hipervínculos, tanto de los que parten del sitio web en cuestión como los que se dirigen al mismo. Esta subunidad surge por

la necesidad de excluir toda responsabilidad por el contenido de sitios hacia los que se dirijan los enlaces propios. Por otro lado, esta unidad supone un intento de ejercer los derechos sobre la propiedad intelectual ya que, en el contexto de la WWW, cualquier sitio web puede potencialmente incorporar un enlace hacia el sitio de la empresa. Esta unidad significativa es de carácter recurrente en los textos localizados (86,96%) y optativa en los textos originales (43,40%).

En primer lugar, las posibles subunidades significativas que articulan esta unidad se centran en las estipulaciones referidas a los sitios web hacia los que se dirigen los enlaces. Por un lado, se declina la responsabilidad por el contenido de cualquier sitio web hiperenlazado, ya que este sitio no se encontraría bajo el control o responsabilidad de la empresa. El siguiente extracto será clarificador en este sentido: «la empresa no asume responsabilidad alguna derivada de la concesión o contenidos de los enlaces de terceros a los que se hace referencia en la web [www.unionfenosa.es](http://www.unionfenosa.es) [Unión Fenosa]. Aparece además una subunidad significativa sobre la inexistencia de acuerdo o relación alguna entre la empresa emisora del sitio y los responsables del sitio al que apunta el enlace. De nuevo presentamos un extracto con el fin de aclarar este aspecto: «el mero establecimiento del enlace no supondrá la existencia de ningún tipo de acuerdo, contrato, patrocinio ni recomendación» [Renfe.es]. En la formulación de esta subunidad en textos localizados hemos identificado el uso recurrente de los sustantivos «respaldo» o «endorso», inexistentes en los textos originales y, consecuentemente, constituirían posibles préstamos léxicos<sup>450</sup>.

Puesto que la empresa declina toda responsabilidad sobre cualquier enlace a terceros, hemos identificado otra subunidad significativa referente al carácter informativo y de referencia de los enlaces que se ofrecen. En los textos originales

---

<sup>450</sup> Como ejemplo «La inclusión de cualquier enlace no implica **su respaldo** por parte de Deere & Company» [John Deere].

esta subunidad significativa suele formularse mediante una referencia a este propio carácter informativo:

La función de los links que aparecen es exclusivamente **la de informar** al usuario sobre la existencia de otras fuentes de información sobre la materia [Xtremesolutions].

Por el contrario, en los sitios web localizados se han identificado ciertas unidades fraseológicas recurrentes como «para su comodidad», y «para-por conveniencia», ambas inexistentes en el corpus original y calco de la unidad fraseológica inglesa *for your convenience*.

Incluimos los enlaces **para su comodidad**, pero no revisamos, controlamos ni vigilamos los materiales de los sitios web que no sean de Medtronic. [Medtronic]

Genworth Financial y sus Afiliadas pueden ofrecerle, **para su conveniencia**, vínculos a sitios Web operados por otras entidades. [Genworth Financial].

Por último, en los textos localizados se ha identificado una subunidad significativa referente al abandono del sitio web corporativo una vez activado un enlace hacia otro sitio web corporativo, «En caso de utilizar dichos vínculos, se abandonará este Sitio» [HP] (véase Figura 16, p. 104).

En un segundo lugar, esta unidad significativa se compone de ciertas subunidades que detallan las condiciones a cumplir por aquellos sitios que deseen incluir un enlace hacia el propio sitio web. La primera de las subunidades identificadas apela al cumplimiento de ciertas condiciones estipuladas por la empresa para que se establezca este tipo de enlaces: «La creación por terceros de links o enlaces con el portal dia.es, deberá ajustarse a las siguientes normas:» [DIA]. De entre las posibles condiciones, la subunidad significativa de mayor frecuencia designa a la página de inicio como la única página de acceso al sitio web. Esta se encuentra altamente representada en ambos subcorpus, tanto en el original: «Cualquier "hiperenlace" al Web de SECURITAS se efectuará a su página principal» [Secúritas], como en el

localizado: «el Usuario no establecerá Vínculos con páginas distintas de la Página de Inicio del Sitio Web» [Stryker].

Otras subunidades significativas en esta unidad son la prohibición del uso de marcos o *frames* al activar el enlace en el sitio web<sup>451</sup>, «no se podrán utilizar marcos, ni cualesquiera otros medios que puedan crear la presunción de que existe una relación entre ATENTO» [Atento], o la obtención por escrito de la autorización para incorporar un enlace hacia el sitio web corporativo. Un ejemplo ilustrativo en el subcorpus original sería: «se prohíbe establecer una conexión, insertar un link o frames de cualquier tipo desde otra website a la página web de HayGroup **sin el consentimiento expreso y por escrito de HayGroup.**» [HayGroup]. Un ejemplo en el subcorpus localizado sería: «Los vínculos a este Sitio Web podrán establecerse únicamente con el previo consentimiento **por escrito** de Stryker» [Stryker].

Tras un análisis contrastivo no hemos identificado formas lingüísticas asociadas a esta unidad significativa y sus subunidades que podríamos considerar convencionales. No obstante, a través de esta unidad significativa podemos analizar el grado de convencionalización de los diferentes términos empleados para denominar el concepto «enlace». La siguiente tabla contrastiva muestra los porcentajes relativos de uso de los diversos términos que se corresponden con este concepto en ambos subcorpus:

Término	% Original	% Localizado
Enlace	60%	70,71%
Link	21,38%	0%
Hiperenlace	11,03%	2,92%
Hipervínculo	4,83%	7,11%
Vínculo	3,45%	18,41%

**Tabla 36. Análisis contrastivo de la terminología empleada para el concepto «enlace».**

<sup>451</sup> El uso de marcos consiste en la inclusión de una página web como parte de otra página global. En caso de incluir en un marco el título y dirección http de la página final.



En ambos subcorpus el término más frecuente sería «enlace». En los textos originales el segundo término más empleado sería el préstamo no adaptado «links» (21,38%), que no aparece en el subcorpus localizado. Este uso exclusivo de un préstamo no adaptado en los textos originales es congruente con los resultados obtenidos anteriormente en referencia al préstamo «email» (véase 10.2.2.5), y constituye una de los rasgos distintivos entre los textos originales y los localizados: la mayor frecuencia de uso de ciertos préstamos no adaptados en los textos originales.

### **10.3.6.3. Unidad significativa de contacto por cuestiones de protección de datos (F-2-VI)**

Esta unidad de contacto en el contexto de la sección de protección de datos personales surge por la gran complejidad de esta sección legal, en la que el emisor se ve obligado a asistir al usuario si fuese preciso. Este ofrecimiento surge por imperativo legal, dada la posibilidad de ejercer los derechos de acceso, oposición, acceso e información, rectificación, o cancelación de los datos personales proporcionados por el usuario y mantenidos por la empresa<sup>452</sup>. Estos derechos comportan el necesario ofrecimiento de la información de contacto en caso de que el usuario desee ejercerlos, aspecto que se cristaliza en la aparición de esta subunidad significativa. Además, la mayor complejidad en general de los textos legales entraña asistir a los usuarios que lo requieran en caso de duda o pregunta. Unos ejemplos ilustrativos en la formulación de esta última subunidad en los subcorpus original y localizado serían:

-Subcorpus original: «Para cualquier aclaración o duda, **puede ponerse en contacto** mediante correo electrónico en la dirección:» [Belnature]

---

<sup>452</sup> Según se especifica en la Ley Orgánica 15/1999.

-Subcorpus localizado: «Si tiene alguna pregunta, observación o inquietud sobre nuestras prácticas de privacidad, **póngase en contacto con nosotros** enviándonos un mensaje de correo electrónico a...» [Estee Lauder].

En el caso de la subunidad de contacto referente a los derechos del usuario sobre los datos personales almacenados por la empresa, un ejemplo dentro del subcorpus original sería:

Podrá ejercer su derecho de acceso, rectificación, cancelación y oposición al tratamiento de sus datos personales, en los términos y condiciones previstos en la propia LOPD. Para su mayor comodidad también **podrá ejercer los derechos antes mencionados mediante la siguiente dirección de correo electrónico:** [Opticaprats].

### **10.3.7. Análisis de unidades y subunidades significativas en la sección de condiciones generales (F3)**

El conjunto de unidades significativas que conforman la sección **condiciones generales** (F-3) presenta las mayores diferencias entre ambos subcorpus del bloque **legal** (F). Las diferencias se cristalizan en la mayor frecuencia en los textos localizados de tres unidades significativas y tres subunidades integrantes de esta sección. En orden de frecuencia se sitúan: la unidad significativa referente a la limitación de responsabilidad (F-3-I), la unidad sobre modificaciones o cambios en los términos o estipulaciones (F-3-III), y la referente a las modificaciones o cambios de contenidos (F-3-II). En lo que respecta a las posibles subunidades, la primera se centra en la limitación de responsabilidad por problemas debidos a virus (F-3-I-a), por errores en el sitio o interrupciones del servicio (F-3-I-b), y por daños y perjuicios (F-3-I-c). En los textos originales, la única unidad significativa recurrente en los textos originales es la referente a la limitación de responsabilidad general (F-3-I) (67,92%). Pasamos a revisar cada una de estas unidades y subunidades más frecuentes.

### **10.3.7.1. Unidad significativa sobre la limitación de responsabilidad (F-3-I)**

Esta unidad significativa recoge la limitación o exoneración de responsabilidad en caso de que se pretendan emprender acciones legales en contra del emisor del texto. En ambas secciones del corpus, esta unidad presenta la mayor frecuencia en la sección de condiciones generales (F3), tanto en originales (67,92%) como en localizados (86,92%). Las distintas subunidades identificadas en la misma representan aspectos que potencialmente podrían derivar en responsabilidades a las que la empresa renuncia expresamente. Éstas subunidades son los problemas ocasionados por posibles virus informáticos (F-3-I-a), problemas por errores en el sitio web o por la interrupción del servicio (F-3-I-b), por daños y perjuicios ocasionados por el uso del sitio web (F-3-I-c), y por posibles contenidos incorporados al sitio en foros o demás áreas interactivas (F-3-I-d).

Esta unidad significativa se encuentra íntimamente relacionada con la exención o exclusión de garantías (F-3-IX), asociada principalmente a los sitios localizados, con un 30,43% de frecuencia frente al 5,66% en los sitios originales. La asociación de estas dos unidades significativas queda patente en la siguiente forma lingüística convencional, «Exclusión [limitación] de garantías y responsabilidad», presente en los títulos de estos subtextos legales.

El análisis de las formas lingüísticas convencionales asociadas a esta unidad significativa se realizó por medio de las distintas formas del lema [responsab\*], ya que su forma sustantiva y adjetiva constituyen las palabras de mayor frecuencia en estos textos: 0,21% del sustantivo «responsabilidad», y 0,20% de la forma adjetiva y sustantiva «responsable»<sup>453</sup>. La siguiente tabla comparativa presenta las distintas

---

<sup>453</sup> Estas palabras ocupan el puesto 22 y 24 respectivamente en la lista de palabras organizada por frecuencia del subcorpus legal original obtenida con la herramienta informática Wordlist.

unidades fraseológicas en las que aparece este lema organizadas por su frecuencia en el subcorpus original:

Unidad fraseológica	Frec. original	Frec. localizado
no se hace responsable de ...	22	25
no será responsable de ...	20	23
no se responsabiliza de ...	18	28
no asume responsabilidad alguna por ...	10	18
no es responsable de cualquier...	9	12
no tiene responsabilidad ...	1	5
se libera de toda responsabilidad...	1	0
no adquiere responsabilidad...	1	0
no acepta responsabilidad alguna...	1	0
rechaza la responsabilidad sobre...	1	0
queda exonerada de responsabilidad...	1	0

**Tabla 37. Análisis contrastivo de la formas lingüísticas asociadas al lema [responsab\*] en subcorpus legal.**

Contrastivamente, las unidades fraseológicas identificadas en el subcorpus localizado muestran ciertas diferencias referentes por un lado (1) al mayor uso de formas en primera persona del plural, (2) al mayor uso del tiempo verbal futuro, y finalmente en (3) la aparición de una ingente cantidad de combinaciones inexistentes en el subcorpus original que incluimos en la Tabla 38.

Si profundizamos en las diferencias significativas, en primer lugar se encuentra el uso de la primera persona del plural en las formas verbales asociadas a esta unidad. Este uso, inexistente en los textos originales, representa el 5,94% del total en el subcorpus localizado, principalmente por medio de las formas «no somos responsables» o «no nos responsabilizamos». El tiempo verbal futuro asociado a este lexema presenta igualmente ciertas diferencias con un 18,03% de uso en los textos originales y un 20,54% en los textos localizados. A pesar de que la diferencia porcentual es escasa, el uso del tiempo verbal futuro en los textos originales se concentra casi exclusivamente en la forma «no será responsable», mientras que en los textos originales, el tiempo futuro se usa igualmente en otros verbos como «no admitirá responsabilidad», «no asumirá responsabilidad», «no tendrá

responsabilidad», «no podrá ser considerado responsable», «no se responsabilizará», o «no será su responsabilidad».

Formas exclusivas de los textos localizados	
no está sujeto a responsabilidad no admitirá responsabilidad legal no asumirá responsabilidad no asumimos responsabilidad no puede asumir responsabilidad no podrá ser considerado responsable no nos hacemos responsables de acuerda no hacer responsable límite de la responsabilidad límite a la responsabilidad exclusión de la responsabilidad limitación de la responsabilidad	no asume ninguna responsabilidad no tendrá responsabilidad <u>mantiene la no responsabilidad</u> no nos responsabilizamos no podrá responsabilizarse a no se responsabilizará acepta que no sean responsables no somos responsables será su responsabilidad no tendrá responsabilidad no haya incurrido en una responsabilidad

**Tabla 38. Formas lingüísticas asociadas al lema [responsab\*] exclusivas del subcorpus legal localizado.**

Otra diferencia observada en los textos localizados la constituye el uso del determinante frente a la palabra responsabilidad en ciertos casos: «límite de la responsabilidad», «límite a la responsabilidad» o «exclusión de la responsabilidad». La forma lingüística convencional identificada en los textos originales, «límite [exclusión] de responsabilidad» no incluye en ningún caso el uso de este determinante, con lo que el uso opcional del mismo podría considerarse propio de los textos localizados.

Finalmente, identificamos en los subtextos localizados la unidad fraseológica «no asume ninguna responsabilidad» en el 6% de las formas asociadas a este lema, con una menor presencia en los originales (1,6%). La inclusión de este adjetivo indefinido frente al sustantivo «responsabilidad» se correspondería en este caso a un calco sintáctico del inglés referente a la expresión *any responsibility*. En los textos originales esta unidad fraseológica quedaría reflejada en «no asume responsabilidad alguna», cuya frecuencia en los textos localizados es 5,4%.

### 10.3.7.1.1. Subunidad significativa sobre la limitación de responsabilidades por virus (F-3-I-a)

La propagación de virus informáticos y sus distintas variedades<sup>454</sup> a través del sitio web implica posibles daños económicos a los usuarios. Si bien las empresas toman las medidas oportunas, la naturaleza cambiante y evolutiva de tales virus les impide controlarlos de manera efectiva. Este hecho conlleva la aparición de una subunidad significativa que exime de responsabilidad a la empresa por posibles daños causados por virus (F-3-I-a). Su carácter recurrente lo atestigua la existencia de una subunidad significativa presente en el 58,70% de textos localizados y el 41,51% de originales. Esta subunidad se refiere tanto a los daños potenciales por virus que partan del sitio web, como a la apelación a no difundir ni introducir en el sitio cualquier programa informático lesivo. En parte, esta exención de responsabilidad se fundamenta en que el propio usuario es quien debe proteger su sistema informático.

1, ya sea por configuración inadecuada, presencia de **virus** informáticos o cualquier otra causa ajena a Ge  
 2 SOLUTIONS no controla ni garantiza la ausencia de **virus** ni de otros elementos en los contenidos que p  
 3 de toda clase que puedan deberse a la presencia de **virus** o a la presencia de otros elementos lesivos en  
 4 da naturaleza que puedan deberse a la presencia de **virus** o a la presencia de otros elementos  
 5 ión por parte de Deloitte de controlar la ausencia de **virus** o cualquier otro elemento informático dañino, c  
 6 lar y no controla con carácter previo, la ausencia de **virus** o **elementos en los contenidos que puedan** pro  
 7 lar y no controla con carácter previo, la ausencia de **virus** o **elementos en los contenidos que puedan** pro  
 8 eb www.unionfenosa.es, ni garantiza la ausencia de **virus** u **otros elementos en los mismos que puedan**  
 9 ace referencia en la web, ni garantiza la ausencia de **virus** u **otros elementos** en los mismos que pued  
 10 das ajenas al control de ATENTO; (v) la presencia de **virus** u **otros elementos que puedan** producir alteraci  
 11 ace referencia en la web, ni garantiza la ausencia de **virus** u **otros elementos** en los mismos que puedan  
 12 no se responsabiliza por la eventual existencia de **virus** u **otros elementos** nocivos, introducidos por

**Figura 112. Extracto de concordancias de la palabra «virus» en subcorpus legal original.**

En los textos originales, las formas lingüísticas convencionales asociadas a esta subunidad que incorporan el término «virus» gravitan en torno a las siguientes unidades fraseológicas: «la presencia de virus», «la ausencia de virus», y «la existencia de virus». El 26,54% de las apariciones del término «virus» se acompañan de la siguiente forma lingüística convencional: «virus u otros

<sup>454</sup> Como los troyanos, gusanos o bombas lógicas.

elementos que puedan producir alteraciones en el [sistema informático, software, sistema físico]». Este carácter convencional queda patente en el extracto de concordancias originales de la palabra «virus».

En los textos localizados, el término «virus» presenta una mayor variación en sus patrones colocacionales. De las unidades fraseológicas que identificamos en los textos originales, sólo hemos encontrado una correspondencia en la falta de garantías por «la ausencia de virus». Ni la forma «la presencia de virus», o «la existencia de virus» se encuentran representadas. Igualmente, la forma lingüística convencional «virus u otros elementos que puedan producir alteraciones» no se encuentra en los textos localizados. No obstante, hemos identificado una unidad fraseológica recurrente de aparición exclusiva en los textos localizados, «está libre de virus», que constituiría un préstamo léxico de la unidad fraseológica inglesa *free of viruses*. Por último, la unidad fraseológica de mayor representación en los textos localizados sería «virus u otros componentes dañinos», aunque existe cierta variación como, «u otro elemento perjudicial», «u otro elemento destructivo».

#### **10.3.7.1.2. Subunidad significativa sobre errores o interrupciones en el sitio (F-3-I-b)**

En esta subunidad significativa se expresa la limitación o exención de responsabilidad tanto en el caso de interrupciones o errores al acceder al sitio web, como por los distintos errores que pudieran presentarse. De entre los posibles errores adicionales, los más recurrentes son (1) los errores tipográficos, (2) los ortográficos, (3) los de contenido, (4) las omisiones, junto con referencias a (5) los errores de seguridad u otros cualquiera causados por virus, y (6) los errores de traducción. A pesar de la distinta naturaleza en los mensajes de error por la interrupción del servicio y el resto de posibles «errores», estos dos aspectos aparecen unidos en su formulación, de ahí que los hayamos recogido en una misma

subunidad significativa. Esta subunidad es recurrente en los textos localizados (56,52%), mientras que en los originales se presenta como ocasional (35,85%).

Desde nuestra perspectiva, cabe reseñar que en el 6,52% de estos textos se recoge la limitación de responsabilidad frente a los posibles errores de traducción, ya que éstas se presentan como un mero servicio al usuario sin garantías de calidad. El siguiente segmento de UPS será esclarecedor en este sentido:

Las traducciones a otros idiomas, en caso de haberlas, **se facilitarán únicamente a efectos prácticos**, y la versión inglesa podrá visionarse de las siguientes maneras [UPS].

En los textos originales las formas lingüísticas convencionales asociadas a esta subunidad son «no se hace responsables de posibles errores», «no garantiza la inexistencia de interrupciones o errores en el acceso a...». En los textos localizados identificamos la unidad fraseológica «no se hace responsable de errores» en sólo dos casos. En estos textos traducidos la palabra «error» aparece relacionada con «omisión», de hecho, en el 18,56% de casos la palabra «error» aparece ligada en la unidad fraseológica «error u omisión». Otra relación de interés se encuentra en la expresión «está libre de error» (11,32%). De entre los términos recurrentes que gravitan en torno al dominio del concepto «error», es exclusivo de los textos localizados la aparición de compuestos «faltas de precisión» y de «inexactitudes /imprecisiones/ incorrecciones técnicas».

En los textos localizados esta subunidad significativa aparece ligada a una forma lingüística en la que se presenta el sitio web en sus condiciones actuales y sin garantías. Para su formulación se utiliza un préstamo léxico que consiste en la inclusión de la forma «tal cual» o «tal como es», que convencionalmente se presenta entre comillas y supondría un préstamo tipográfico. La misma puede estar acompañada de «según disponibilidad». Esta forma lingüística aparece en el 34,78% de textos localizados, aunque en el caso de incluir las distintas variaciones en su



traducción, su frecuencia ascendería al 45,65% de sitios web. Las siguientes líneas de concordancia muestran la alta frecuencia de esta forma lingüística exclusiva de los textos localizados, así como sus posibles variantes de traducción:

N	Concordance
9	lementos (sin carácter limitativo) se proporcionan 'TAL CUAL' y están 'DISPONIBLES' sin manifes
10	cionamiento, y acepta que el sitio se ofrece 'tal cual'. Sealed Air Corporation no ofrece ningú
11	RAMENTE INFORMATIVOS Y SE PRESENTAN 'TAL CUAL'. POR CONSIGUIENTE, AL OFREC
12	TENIDOS DE ESTE SITIO WEB SE OFRECEN 'TAL CUAL'. MEDIANTE ESTE DOCUMENTO,
13	Garantías* EL SERVICIO SE PROPORCIONA 'TAL CUAL'. ADM NO OFRECE NINGÚN TIPO
14	Í CONTENIDOS SE PROPORCIONAN EN BASE 'TAL CUAL', SIN GARANTÍAS DE NINGUNA CL
15	DUCTOS DE TERCEROS, SE PROPORCIONAN 'TAL CUAL' Y "SEGÚN DISPONIBILIDAD" PAR
16	su contenido. Los juegos en fase beta se ofrecen "tal cual" y "según disponibilidad" y no le propor
17	te sitio es responsabilidad suya. El sitio se ofrece "tal cual" y "según disponibilidad". Nos reservam
18	abilidad LOS MATERIALES SE SUMINISTRAN 'TAL CUAL' SIN NINGÚN TIPO DE GARANTÍA,
19	FRECEN EN ESTE SITIO SE PROPORCIONAN 'TAL CUAL' SIN GARANTÍA DE NINGÚN TIPO,
20	ADOS CON LOS MISMOS SE PROPORCIONAN 'TAL CUAL' SIN GARANTÍA DE NINGÚN TIPO.
21	sable de los mismos, los cuales se proporcionan "tal cual" ('as is') sin garantías de ningún tipo, ex
22	ro no está limitada a, la información de contacto (tal como su nombre, dirección postal, dirección
23	es tipográficos. Todo en este sitio se proporciona "tal como es", sin garantía de ningún tipo, ya se
24	l contenido de este sitio web se ofrece al usuario 'TAL COMO ESTÁ', SIN GARANTÍA DE NIN

**Figura 113. Extracto de concordancias de la unidad fraseológica exclusiva de los textos localizados «tal cual» en subunidad significativa sobre errores o interrupciones en el sitio (F-3-I-b).**

Su presencia recurrente exclusiva en textos localizados le confiere su carácter de préstamo léxico y sintáctico, que como ejemplifica la línea de concordancia 21, queda patente mediante la inclusión de su versión original en inglés, «"tal cual" ("as is")» [Stryker]. En este extracto de línea de concordancia, observamos asimismo el uso frecuente de mayúsculas como recurso tipográfico, testigo de la gran importancia que se confiere a estos segmentos en los textos de origen y que se transfiere a los localizados<sup>455</sup>. Este uso de mayúsculas en subtextos legales es exclusivo de los textos localizados, y sería contrario a las recomendaciones de guías de estilo<sup>456</sup> (véase 5.9.5). De esta manera, el uso de mayúsculas en los subtextos legales en el corpus localizado podría considerarse un anglicismo ortotipográfico<sup>457</sup>.

<sup>455</sup> El mayor uso de las mayúsculas en la expresión de esta subunidad significativa apuntaría a su gran importancia en el contexto sociocultural estadounidense.

<sup>456</sup> Como mencionamos anteriormente, el uso de mayúsculas está asociado con el grito y es inapropiado en los sitios web corporativos, adicionalmente, ralentiza la lectura en un 10% (Nielsen y Loranger 2002: 238).

<sup>457</sup> El uso de texto en mayúsculas en los subtextos legales originales se reduce a los títulos.

Dada la alta frecuencia de la forma convencional «as is» en los textos localizados, podemos inferir por su alto grado de convencionalización en los textos de origen en el contexto sociocultural de partida. No obstante, son varias las distintas versiones traducidas de esta forma convencional que hemos identificado y que presentamos por su interés:

- El Sitio web de Agribands es operado por Cargill sobre una base **“COMO SE ENCUENTRA”, “COMO ESTÉ DISPONIBLE”** [Cargill]
- Este sitio, los servicios, la información y los materiales del mismo así como cualquier software que se ponga a disposición en el sitio, se proveen **“en el estado en que se encuentran”** [Genworth Financial]
- La información de esta publicación, incluyendo texto, imágenes y enlaces, SON PROPORCIONADAS EN SU **"ESTADO ACTUAL"** POR CATERPILLAR INC. PARA COMODIDAD DE SUS DISTRIBUIDORES Y CLIENTES [Caterpillar]
- Sin limitar lo precedente, todo en el Sitio Web le es proporcionado en **el estado en que se encuentra** [Mattel.es]
- LOS MATERIALES (INCLUIDO EL SOFTWARE) Y LOS SERVICIOS PROPORCIONADOS EN ESTE SITIO SE SUMINISTRAN **«EN LAS CONDICIONES QUE SE ENCUENTRAN»**, SIN GARANTÍAS DE NINGÚN TIPO [HP.es]

De los ejemplos anteriores podemos deducir que las variaciones en la traducción de esta forma lingüística convencional en los textos de partida se cristalizan en las palabras «estado» o «condiciones», seguido de la oración adjetival subordinada «como se encuentra», o de la forma adjetival «actual». Es igualmente significativo el uso de mayúsculas en la oración que la contiene o, como se observa en la primera línea, sólo en la forma lingüística convencional. Este hecho presentaría una interesante relación entre la importancia que el emisor confiere a tal segmento y las decisión traductora de mantener una traducción literal, incluyendo el uso tipográfico de mayúsculas o comillas, o incluso con el mantenimiento de la forma en inglés. Finalmente, hemos de indicar que esta forma se encuentra además relacionada con la unidad significativa de exclusión de garantías (F-3-IX), presente en el 30,43% de estos textos y el 5,66% de los textos originales.

### **10.3.7.1.3. Subunidad significativa sobre la limitación de responsabilidad por daños y perjuicios (F-3-I-C)**

Esta sería la última subunidad significativa de carácter recurrente en los textos localizados, y supone la exención o limitación de responsabilidad por los posibles daños y perjuicios que se derivaran del sitio. Su frecuencia de aparición asciende al 50,17% en estos textos y 24,53% en originales. Si analizamos el contenido proposicional de esta unidad, los daños y perjuicios se podrían derivar del (1) acceso, uso, mantenimiento, calidad, disponibilidad o fiabilidad del sitio, (2) de la pérdida o intromisión en los datos personales almacenados por la empresa o (3) cualquier daño resultante por virus u otros problemas no imputables a la misma.

El resultado del análisis contrastivo arroja diferencias considerables. En primer lugar, la unidad fraseológica más extendida en las subunidades originales, «los daños y perjuicios», presenta un claro contraste en su frecuencia, un 52,7% frente al 18,48% en subtextos localizados. En éstos existe la posibilidad del cambio de la conjunción copulativa «y» por la disyuntiva «o», «daños o perjuicios». En segundo lugar, la única relación convencional en los textos originales asociada a la expresión anterior serían las distintas formas del lema [responsab\*], tales como «no se hace responsable por los daños y perjuicios», «no se responsabiliza de los daños y perjuicios», o «excluye cualquier responsabilidad por los daños y perjuicios de cualquier clase [de toda naturaleza]». Por su parte, los textos localizados ofrecen un mayor rango colocacional para esta unidad fraseológica.

De entre los posibles daños recogidos, la terminología empleada en los textos originales se limita a las unidades fraseológicas «daños directos o indirectos», «daño emergente y/o por lucro cesante». Esta unidad fraseológica se identificó en el 5,6% de los textos originales y en el 2,17% de los localizados. En estos últimos identificamos una mayor variación en la terminología existente dentro del campo

conceptual de los daños: «directos», «indirectos», «especiales», «punitivos», «punitorios», «fortuitos», «ejemplares», «emergentes», «extraordinarios», «resultantes», «incidentales», «causales», «consecuenciales», «ejemplares», «puntuales», «cuantificables», «emergentes». Ciertos adjetivos como «punitorios» constituirían errores de traducción. De las dieciocho posibilidades anteriores, solo dos se encuentran representados en los textos originales, «directos», e «indirectos». No obstante, el análisis contrastivo muestra que en los textos originales el término «daños» suele ir acompañado en mayor medida de un compuesto nominal que abarque la totalidad de daños posibles, como «de cualquier tipo», «de cualquier naturaleza», «de toda naturaleza», o «de toda clase». Estos resultados supondrían la existencia de una mayor posibilidad de daños recogida en el sistema jurídico estadounidense, que consecuentemente se refleja en estas unidades fraseológicas.

#### **10.3.7.2. Unidad significativa sobre modificaciones al contenido del sitio (F-3-II)**

La empresa emisora se reserva el derecho de modificar o cambiar los contenidos del sitio web en cualquier momento, y esta facultad se cristaliza en la aparición de la presente unidad significativa (F-3-II). Su frecuencia de aparición es 54,35% en los textos localizados y 41,55% en los originales. Esta unidad se encuentra relacionada con la unidad referente a los cambios o modificaciones en los textos legales en sí (F-3-II), que analizaremos en el siguiente epígrafe.

Para el análisis de las formas lingüísticas asociadas a esta unidad se seleccionó la forma lingüística recurrente «se reserva el derecho». Esta selección se debió a que ocupa el quinto puesto en la lista de frecuencia global por cuatro elementos en el subcorpus original y la segunda posición en el localizado (véase Tabla 27, p. 455). Esta unidad fraseológica aparece mayoritariamente relacionada con ciertas unidades significativas de la sección de **condiciones generales** (F3), como la unidad

de modificación o cambios de contenidos (F-3-II), la unidad significativa de modificación de los términos legales (F-3-III), la unidad sobre la restricción al acceso a discreción de la empresa (F-3-VI), o la subunidad significativa referente a facultad de la empresa de suspender el acceso al sitio web (F-3-I-a). A continuación presentamos la tabla resumen contrastiva de las posibles colocaciones de esta unidad fraseológica:

Unidad significativa de modificación o cambio en contenidos (F-3-II)		
	Original	Localizado
Se reserva el derecho a	-revisar las páginas [1] -actualizar contenidos[1] -realizar cambios en portal [1] -actualizar, modificar o eliminar información [1]	-modificar, borrar o reorganizar los contenidos [2] -efectuar cambios al sitio [2] -modificar el contenido del sitio [1] -modificar la presentación [1] -actualizar o modificar la información contenida en su Web [1] -introducir modificaciones en el web[1] -realizar cambios en el portal [1] -cambiar el contenido presente en este Sitio[1] -cambiar, modificar, agregar o eliminar contenidos [1] -realizar correcciones o cambios [1] -inhabilitar elementos [1] -corregir cualquier error, falta de precisión [1] -eliminar el contenido [1]
Se reserva el derecho de	-modificar el sitio [1] -efectuar modificaciones [1] -cambio o cancelación de sitio [1] -modificar [1]	-modificar los contenidos [3] -agregar, eliminar o modificar información [1] -efectuar modificaciones [1] -inhabilitar elementos [1] -cambiar, <b>adicionar</b> o eliminar [1] -modificar, añadir o suprimir [1]
Se reserva la facultad de	-efectuar modificaciones [3]	

**Tabla 39. Tabla contrastiva de colocaciones de «se reserva el derecho» en la unidad (F-3-II).**

El análisis contrastivo de esta unidad fraseológica muestra que los textos localizados presentan sólo dos formas, «se reserva el derecho a» y «se reserva el derecho de», mientras que en los textos originales se recoge una mayor variación con tres formas adicionales: «se reserva el derecho para [...]», «se reserva la facultad de [...]»<sup>458</sup>.

<sup>458</sup>También identificamos «se reserva la posibilidad de [...]», aunque esta forma está asociada en los textos originales a «ejercer acciones judiciales» [2] y a «ejercitar su derecho de acceso, rectificación, cancelación y oposición» [2].

La forma lingüística de mayor frecuencia de uso en los textos localizados sería «se reserva la facultad de efectuar modificaciones», existiendo la posibilidad de variación con «se reserva el derecho de efectuar modificaciones». Esta última cuenta con una sola aparición en los textos localizados. La forma lingüística de mayor frecuencia en los textos localizados sería «se reserva el derecho de modificar los contenidos», con dos casos adicionales de «se reserva el derecho de modificar, borrar o reorganizar los contenidos». Los textos localizados muestran una mayor tendencia a incluir varios términos referentes tanto a los posibles cambios en el contenido del sitio web («modificar», «cambiar», «añadir», «actualizar» o «agregar»), como en la eliminación de contenidos («eliminar» o «suprimir»).

### **10.3.7.3. Unidad significativa sobre modificaciones en términos legales (F-3-III)**

Ésta última unidad significativa recurrente en los textos localizados recoge la facultad de efectuar modificaciones al texto legal de manera unilateral. Dado el carácter jurídico y vinculante de estos textos, esta potestad unilateral de modificación entraña la consecuente apelación a leer y revisar con cierta frecuencia el texto. En este sentido, esta unidad estaría ligada a la subunidad sobre la apelación a leer el texto (F-1-I-b), y a la unidad sobre la aceptación de términos legales (F-1-II).

Aunque no de forma exclusiva, las formas lingüísticas asociadas a esta unidad incorporan asimismo la locución «se reserva el derecho a-de», o paralelamente a la unidad anterior (F-3-II), aparece además en los textos originales la locución «se reserva la facultad de». En ambos subcorpus es mayoritario el uso del verbo «modificar» o de la locución «efectuar modificaciones» para completar la unidad fraseológica. La forma lingüística convencional mayoritaria en los textos originales la constituye «se reserva el derecho de/a modificar [la(s) presente(s)]

condiciones/política». En segundo lugar se encontraría la forma lingüística «se reserva la facultad de efectuar modificaciones a [las presentes] condiciones». Los textos localizados muestran igualmente una preferencia por la forma lingüística convencional más usada en los textos originales, «se reserva el derecho a modificar [estos-a] términos-condiciones». No obstante, esta forma suele incluir el pronombre demostrativo «estos» o «esta» en lugar del adjetivo «presentes», más frecuente en los textos originales. En consecuencia, este análisis revela que la presencia de la colocación «estas condiciones» o «estos términos y condiciones» sería propia de textos localizados, ya que los textos originales muestran mayoritariamente «las presentes condiciones» o «la presente política». Las diferencias se cristalizan igualmente en la mayor variación en los verbos usados en la formulación de esta subunidad, con la aparición de «actualizar», «añadir partes» o «ampliar». Contrastivamente, los textos originales incorporan un solo verbo en la expresión de esta subunidad.

Unidad significativa de modificación o cambios en las condiciones o términos legales		
	Original	Localizado
Se reserva el derecho a	<ul style="list-style-type: none"> <li>-modificar las presentes condiciones[2]</li> <li>-modificar la presente política [2]</li> <li>-modificar las condiciones [1]</li> <li>-cambiar las condiciones [1]</li> </ul>	<ul style="list-style-type: none"> <li>-actualizar o modificar estos términos y cond. [2]</li> <li>-cambiar o modificar los términos [2]</li> <li>-modificar el acuerdo [2]</li> <li>-modificar o ampliar las condiciones [1]</li> <li>-modificar las condiciones [1]</li> <li>-modificar los términos [1]</li> <li>-modificar, añadir o eliminar partes de la declaración [1]</li> <li>-modificar, agregar o eliminar disposiciones [1]</li> <li>-realizar cambios en la normativa [1]</li> <li>-revisar y modificar la presente declaración [1]</li> </ul>
Se reserva el derecho de	<ul style="list-style-type: none"> <li>-modificar en cualquier momento las condiciones [1]</li> <li>-modificar las presentes condiciones [1]</li> </ul>	<ul style="list-style-type: none"> <li>-modificar estas condiciones [2]</li> <li>-actualizar esta directiva [1]</li> <li>-modificar este acuerdo [1]</li> <li>-cambiar dichos términos [1]</li> </ul>
Se reserva la facultad de	<ul style="list-style-type: none"> <li>-efectuar modificaciones a las presentes Condiciones [3]</li> <li>-de modificar las condiciones [1]</li> <li>-de modificación de la presente política [1]</li> </ul>	

**Tabla 40. Tabla contrastiva de colocaciones de «se reserva el derecho» en la unidad (F-3-III).**

Las formas lingüísticas asociadas a esta subunidad no se concretan exclusivamente en las formas indicadas anteriormente, sino que se encuentran formulaciones de esta subunidad diferentes como «UPS podrá revisar y actualizar en cualquier momento las presentes Condiciones de Uso» [UPS]. Como se indicó anteriormente, esta unidad puede incorporar una subunidad referente a la revisión periódica de los textos legales por posibles cambios: «ADM se reserva el derecho de modificar este Acuerdo en cualquier momento [...] debería visitar esta página periódicamente para revisar el entonces actual<sup>459</sup> Acuerdo de Términos de Uso al que está vinculado» [ADM].

### **10.3.8. Resumen contrastivo de formas lingüísticas convencionales asociadas a la sección legal (F)**

Con un afán descriptivo, presentamos a continuación una tabla contrastiva de los segmentos textuales más frecuentes en los textos originales, que contrastamos con segmentos localizados. En estos últimos hemos resaltado en negrita los posibles errores o inadecuaciones.

---

<sup>459</sup> Error de traducción.



Sección y unidad significativa ORIGINAL	LOCALIZADO
F1. AVISOS LEGALES	
F-1- I. Presentación	
	Bienvenida (F-1-I-a)
	¡Bienvenido a Boeing.es, el Sitio Web de la compañía Boeing sobre las actividades de Boeing en España (el "Sitio")! .
	Leer texto (F-1-I-b)
Le rogamos que lea con atención los términos de este acuerdo antes de visualizar los contenidos	<b>¡MUY IMPORTANTE! LEA ESTE ACUERDO LEGAL ANTES DE VISITAR LA WEB.</b>
F-1- II. Aceptación de términos legales	
	Aceptar (F-1-II-a)
El usuario expresa su comprensión y aceptación plena y sin reservas del contenido de esta "Nota Legal".	El hecho de acceder a este sitio web y <b>recorrer sus páginas</b> implica la total aceptación de estas condiciones generales
	Abandonar el sitio (F-1-II-b)
Si estas condiciones no le son aceptables, se le pedirá que abandone el sitio	Si no acepta estos Términos y condiciones (que pueden enmendarse de vez en cuando) o está disconforme con este sitio Web, <b>su único y exclusivo remedio es interrumpir su uso.</b>  <b>Si está a disgusto con el Sitio web</b> de Agribands o cualquier contenido del sitio o con estos términos o condiciones, <b>el único recurso exclusivo</b> que usted tiene es dejar de usar el Sitio web de Agribands
F-1- III. Registro legal	
	Inscripción de la empresa (F-1-III-a)
ADECCO TT, S.A., Empresa de Trabajo Temporal, debidamente inscrita en el Registro Mercantil de Madrid, al tomo 8171, folio 153, hoja M-131861, inscripción 1	Compañía inscrita en el Registro Mercantil de Madrid <b>en el Tomo 18.606, Folio 191, Hoja M- 323639»</b>
	Domicilio social (F-1-III-b)
[...] es titular [...] S.A. con domicilio social en Bodión, s/n, 15930 - Boiro (A Coruña)	[...] una sociedad con domicilio social en Manuel Tovar, 35, 28034 Madrid (España),[...]
	Código de Identificación Fiscal (F-1-III-c)
[...] con C.I.F. A-83319400 y domicilio en Madrid.	C.I.F.: A-80781727 Sociedad Unipersonal.
F-1- IV. Ley Aplicable y Jurisdicción	
La prestación del servicio del Sitio Web y las presentes Condiciones Generales de Uso del Sitio Web se rigen por la ley española.	Las <b>leyes</b> del Estado de California, sin contar las <b>normas de conflicto entre leyes, gobiernan</b> este Acuerdo  Para toda cuestión litigiosa o que incumba al sitio "web" de la <b>Firma PwC</b> , será de aplicación la legislación española
F-1- V. Dónde almacenan	
	Mattel controla y hace funcionar este Sitio Web desde su sede en El Segundo, California, Estados Unidos y no puede asegurarle que estos materiales sean apropiados o disponibles para ser utilizados en <b>otras localidades.</b>

## Capítulo 10. Análisis contrastivo de formas lingüísticas convencionales

Sección y unidad significativa ORIGINAL	LOCALIZADO
F-1- VI. Propiedad intelectual-uso del contenido	
La presente página web está protegida por los derechos de propiedad intelectual y éstos son de exclusiva propiedad de la empresa.	Los materiales de este Sitio están protegidos por los derechos de propiedad intelectual y <b>cualquier</b> uso no autorizado de <b>cualquier</b> material de este Sitio es susceptible de incumplir la normativa de propiedad intelectual, marcas y otras <b>regulaciones</b> .
	Autorización por escrito (F-I-VI-a)
Queda expresamente prohibida la reproducción total o parcial de este sitio web sin el permiso expreso y por escrito de Conor Sports, S.A.	No podrá realizarse ningún uso de los mismos que no haya sido previamente autorizado por escrito por Johnson & Johnson Vision Care.
	Qué material está protegido (F-I-VI-b)
están protegidos por los derechos de autor, los logos, combinaciones de colores, la selección y la forma de presentación, el código fuente del sitio web, los menús, los botones de navegación, el código HTML, los applets de Java, los textos, imágenes, gráficos, así como cualquier otro contenido del sitio web relativo a los productos, servicios, contenidos y/o herramientas facilitados por Panda Software	Los contenidos del Portal, como servicios, fotografías, gráficos, imágenes, iconos, tecnología, software, links y demás contenidos audiovisuales o sonoros, así como su diseño gráfico y códigos fuente ( <b>en adelante, los "Contenidos"</b> ), son propiedad intelectual de MSD.
F-1- VII. Propiedad el sitio web	
Los datos identificativos del responsable y titular del portal [...] son:	Este sitio Web <b>es publicado</b> por Johnson & Johnson, S.A. División Lifescan, siendo el mismo responsable exclusivo de sus contenidos.
F-1- VIII. Destinatario del sitio web	
Este Sitio Web se dirige exclusivamente a residentes españoles.	Este sitio está diseñado para residentes en España y sus contenidos se han configurado para cumplir las regulaciones y leyes vigentes.
F-1- IX. Uso de las marcas registradas	
Deloitte se reserva todos los derechos sobre las marcas, nombres comerciales o demás signos distintivos, las patentes, y la propiedad intelectual, con respecto al contenido y diseño de este Portal de Internet.	Todos los nombres de productos farmacéuticos que <b>pusieron</b> aparecer en este sitio Web son marcas registradas propiedad de ó licenciadas a Merck & Co., Inc., sus subsidiarias ó afiliadas.
F-1-X. Fecha de entrada en vigor o revisión	
La información que aparece en este website es la vigente en la fecha de su última actualización.	Esta Declaración de privacidad se actualizó por última vez el <b>24 de Maj de 2002</b> .
F2. PROTECCIÓN DE DATOS PERSONALES	
F-2-I. Cumplimiento de leyes españolas	
	Ley española LOPD 1999 (F-2-I-a)
El presente Aviso legal ha sido elaborado con respeto y cumplimiento de las disposiciones recogidas en la Ley Orgánica 15/1999, de 13 de Diciembre, de Protección de Datos de Carácter Personal	BMS cumple con la Ley Orgánica 15/1999, de protección de datos de carácter personal y con cualquier otra normativa vigente en la materia

Sección y unidad significativa ORIGINAL	LOCALIZADO
	Ley 34/2002, de 11 de julio, de Servicios de la Sociedad de la Información y de Comercio Electrónico. (F-2-I-b)
Para dar cumplimiento con lo establecido en la Ley 34/2002, de 11 de julio, de servicios de la sociedad de la información y de comercio electrónico, a continuación se indican los datos de información general de este sitio web:	En cumplimiento de lo dispuesto en el artículo 10 de la Ley 34/2002, de 11 de julio, de Servicios de la Sociedad de la Información y Comercio Electrónico, se pone a disposición de los Usuarios la información general relativa a las Empresas del Grupo Aon en España:
F-2-II.Recogida	
Al registrarse para recibir información, usted proporcionará voluntariamente la información personal que se le solicite en cada caso.	<b>La información a través de este sitio</b> que pueda identificarlo personalmente únicamente se recoge en caso de que usted ofrezca dicha información de manera voluntaria.
F-2-III.Derechos de acceso, rectificación, cancelación y oposición	
Posibilidad de ejercer los derechos de acceso, rectificación, cancelación y oposición.	En cualquier momento, usted podrá ejercitar sus derechos de acceso, rectificación, cancelación, información de valoraciones y oposición en los términos establecidos por la legislación vigente.
F-2-IV.Enlaces hacia o desde el sitio web	
La inclusión en el portal dia.es de links o enlaces con otros portales, tiene una finalidad meramente informativa, y no supone que DIA% recomiende y/o garantice dichos portales, sobre los que no ejerce control alguno, ni es responsable del contenido de los mismos.	En este sitio web el usuario podrá encontrar diversos enlaces que le conducirán a páginas web independientes de ADP GSI ESPAÑA, S.A.. Su única finalidad es la de mostrar algunos ejemplos en Internet, <b>del</b> trabajo realizado.
F-2-V.Uso de datos personales	
Toda información sobre clientes recopilada en el sitio web se utilizará para estudios propios, con fines de promoción o ampliación de jungiberica.es	El motivo principal por el cual Autodesk recopila información personal es ofrecer a los usuarios una experiencia personalizada al utilizar sus sitios Web y sus productos y servicios. Asimismo, esta información puede utilizarse para notificarle aspectos relacionados con el uso de nuestros productos y servicios.
F-2-VI.Contacto	
Para cualquier aclaración o duda, puede ponerse en contacto mediante correo electrónico en la dirección:	Si tiene alguna pregunta, observación o <b>inquietud</b> sobre nuestras prácticas de privacidad, póngase en contacto con nosotros enviándonos un mensaje de correo electrónico a:
F-2-VII.Cookies	
	Definición (F-2-VII-a)
Una cookie es un pequeño elemento informativo enviado por el sitio web contactado y guardado en disco duro por el explorador en su ordenador	Una cookie es <b>un dato</b> que una página web puede enviar al explorador del usuario y que posteriormente se puede guardar en el sistema de éste.
F-2-VIII.Revelación	
Jung Electro Ibérica S.A., no revelará la información personal que usted proporcione en este sitio web a terceros, salvo cuando la revelación sea exigida por la legislación vigente, por orden judicial, o por una autoridad competente gubernamental o de otra índole	Se podrán igualmente difundir datos de carácter personal que Ud. haya suministrado a través de este sitio a terceros que no estén afiliados a Johnson & Johnson, si bien únicamente para ...

## Capítulo 10. Análisis contrastivo de formas lingüísticas convencionales

Sección y unidad significativa ORIGINAL	LOCALIZADO
F-2-IX.Seguridad	
Ingenieros Asesores, S.A. ha implementado el nivel de seguridad que permiten las actuales tecnologías y mantiene una constante vigilancia sobre las condiciones de seguridad con las que se gestionan y utilizan las bases de datos.	<b>Ponemos en práctica</b> numerosas medidas de seguridad para proteger la seguridad de su información personal, tanto on line como off line.
	Riesgos (F-2-IX-a)
No obstante, estas medidas no son infalibles y, por ello, Renfe-Operadora no puede asegurar totalmente la ausencia de dichos elementos nocivos.	Sin embargo, puesto que es imposible garantizar absolutamente la seguridad en las transmisiones de datos que se realizan por Internet, no podemos asegurarle la seguridad de la información <b>que nos envía o que nosotros le enviamos</b> sobre productos y servicios en línea.
F-2-X.Notificación de revelación	
Cualquier cesión a terceros de los datos personales de los usuarios de este portal, será comunicada debidamente a los afectados especificando la identidad de los cesionarios y la finalidad con que se van a tratar los datos que se ceden	Excepto cuando se permita explícitamente <b>ó lo</b> requiera la ley, <b>ó</b> esté previsto y aceptado en esta política, su información personal no será compartida sin su permiso.
F-2-XI.Protección de menores	
Gebasa no se responsabiliza del uso que los menores hagan de la web. Gebasa no es responsable de que los contenidos de la web puedan herir la sensibilidad de los menores	Los sitios web de CIGNA no destinan contenidos a menores de edad ni pretenden recabar información personal sobre menores de edad.  Sin limitar lo anterior, el Servicio no está disponible para menores. <b>Si no cumple con los requerimientos, no use el Servicio.</b>
F-2-XII.Direcciones IP	
El registro de dicha dirección IP sirve sólo para fines exclusivamente internos, como son las estadísticas de acceso a este sitio web.	Los Sitios Web de UPS recogen direcciones IP a efectos de administrar el sistema, brindar seguridad y realizar análisis estadísticos. Una dirección IP es un número asignado automáticamente a una computadora cuando ésta se conecta a Internet.
<b>F3. CONDICIONES GENERALES</b>	
F-3-I.Limitación de responsabilidad	
	Problemas por virus (F-3-I-a)
[...] ni garantiza la ausencia de virus u otros elementos en los mismos que puedan producir alteraciones en el sistema informático (hardware y software), los documentos o los ficheros del usuario, incluyendo cualquier responsabilidad por los daños de cualquier clase causados al usuario por este motivo.	Aunque GLASER se esforzará en mantener la Página Web de GLASER <b>libre de virus</b> , GLASER no garantiza que la misma esté absolutamente <b>libre de virus</b> .
	Errores del sitio – interrupción (F-3-I-b)
Jazztel no garantiza la inexistencia de interrupciones o errores en el acceso a la Web Corporativa, en su contenido, ni que éste se encuentre actualizado, aunque desarrollará sus mejores esfuerzos para, en su caso, evitarlos, subsanarlos o actualizarlos	<b>GARANTÍA RELATIVA A RETRASOS, INTERRUPTIONES, ERRORES U OMISIONES EN EL SERVICIO O EN CUALQUIER PARTE DEL MISMO</b>

Sección y unidad significativa ORIGINAL	LOCALIZADO
	Por daños y perjuicios (F-3-I-c)
SECURITAS no otorga ninguna garantía ni se hace responsable, en ningún caso, de los daños y perjuicios de cualquier naturaleza que pudieran traer causa de	En cualquier caso, sea cual fuere su causa, la Firma PwC no asumirá responsabilidad alguna ya sea por daños directos o indirectos, daño emergente y/o por lucro cesante
	Exactitud de contenidos (F-3-I-d)
Deloitte no se encuentra obligada a comprobar la veracidad, exactitud y actualidad de la información suministrada a través de su página web.	Las informaciones contenidas en este sitio pueden contener inexactitudes o errores tipográficos, que pueden cambiarse, corregirse o suprimirse en cualquier momento.  El material incluido en este sitio web ha sido compilado a partir de fuentes diversas y el sitio puede contener inexactitudes técnicas o errores tipográficos.
	Por contenidos de usuarios (F-3-I-d)
ATENTO no edita los Contenidos de terceros publicados en el Sitio Web ni se hace responsable de la licitud, fiabilidad, veracidad, exactitud, exhaustividad y actualidad de dichos Contenidos	HP no asume ninguna responsabilidad como consecuencia de dichas Comunicaciones en caso de que las mismas infrinjan la normativa de Comercio Electrónico, Propiedad Intelectual, Protección de Datos, Derecho al Honor, a la Intimidad o a la Propia Imagen, Orden público o cualquier otra que resulte de aplicación, comprometiéndose el usuario a mantener indemne a HP a tal respecto.
F-3-II. Modificaciones o cambios de contenidos	
Telefónica, S.A. se reserva la facultad de efectuar, en cualquier momento y sin necesidad de previo aviso, modificaciones y actualizaciones de la información contenida en su web [...]	[...] se reserva el derecho a modificar, borrar o reorganizar cualquiera de los contenidos de esta página en el momento en que lo considere apropiado.
F-3-II. Modificación o cambio de términos	
Caja Madrid se reserva la facultad de modificar unilateralmente las condiciones y términos de uso de este website.	MSD se reserva el derecho a modificar estos términos eventualmente cuando lo estime necesario
F-3-IV. Uso adecuado	
El usuario se compromete a utilizar el Web y sus servicios y contenidos sin contravenir la legislación vigente, la buena fe, los usos generalmente aceptados y el orden público	El Usuario <b>se obliga a no</b> utilizar el Portal ni los servicios ofertados en o a través del mismo para la realización de actividades contrarias a las leyes, a la moral, al orden público, lesivas de los derechos e intereses de terceros o que de cualquier otra forma puedan dañar, inutilizar, sobrecargar, deteriorar o impedir la normal utilización del Portal.
F-3-V.uso personal y privado	
Se permite el uso de este sitio Web y de su contenido para uso privado no comercial.	El contenido del Sitio web de Agribands, y el sitio en su conjunto, está dirigido exclusivamente <b>para uso personal y no comercial</b> de los usuarios de nuestro sitio
F-3-VI. Restringir el acceso	
Panda Software se reserva el derecho de revocar la autorización para utilizar el contenido de su sitio web, contenidos y/o herramientas en cualquier momento	HP se reserva el derecho de dar por terminada la autorización para utilizar cualquier servicio del Sitio

Capítulo 10. Análisis contrastivo de formas lingüísticas convencionales

Sección y unidad significativa ORIGINAL	LOCALIZADO
F-3-VII. Publicar materiales ilegales, polémico, pornográfico, amenazante...	
Difundir contenidos o propaganda de carácter racista, xenófobo, pornográfico, apología del terrorismo o contrarios a los derechos humanos.	Sin limitación a lo que sigue, acepta no transmitir, distribuir, remitir, comunicar o almacenar la información ni ningún otro tipo de material sobre, a o a través del Servicio que: [...] sea obsceno, difamatorio, amenazante, vejatorio, abusivo, <b>aborrecible</b> , difamatorio o <b>embarazoso</b> a cualquier otra persona o entidad según lo determine ADM bajo su propia discreción; [...] sea sexualmente explícito.
F-3-VIII. uso del usuario con fines ilícitos o lesivos - ilícito, negligente, fraudulento	
Asimismo, queda prohibido, el uso del Web con fines ilícitos o lesivos contra Sistema AZUD, S.A. o cualquier tercero, o que, de cualquier forma, puedan causar perjuicio o impedir el normal funcionamiento del Web.	Usted se abstendrá de utilizar el sitio web y/o sus contenidos con fines ilícitos, prohibidos por estas Condiciones de Uso, lesivos de los derechos e intereses de terceros, o que de cualquier forma puedan dañar, inutilizar, sobrecargar, deteriorar o impedir la normal utilización del sitio web [...]
F-3-IX. Exclusión de garantías	
CEMENTO NATURAL, S.L. puede modificar sin previo aviso cualquier información aquí recogida ésta no está sujeta a ninguna garantía de ningún tipo.	XEROX no asume garantía de ningún tipo, ya sea expresa o implícita, sobre la exactitud o precisión de la información, ni que la misma sea completa.
F-3-X. NIELSEN RATINGS	
Este sitio utiliza las herramientas de medición de audiencia de Nielsen//NetRatings, que permiten analizar el comportamiento y perfil de los visitantes anónimos [...]	
F-3-XI. Material enviado es público	
Cualesquiera informaciones, ideas o materiales, de cualquier naturaleza, que sean remitidos a CEMENTO NATURAL, S.L. por el Usuario serán considerados públicos y no confidenciales, y podrán ser utilizados libremente por CEMENTO NATURAL, S.L.	Ford se reserva el derecho a utilizar cualquier material enviado <b>al site</b> para fines propios, incluidos textos e imágenes, ya sea en el <b>site</b> o de cualquier otra forma, incluyendo fines publicitarios.
F-3-XII. Expectativas futuras	
	Cualquier declaración contenida en el sitio en relación a las previsiones de futuro de VeriSign se considerará "declaración de previsión" de acuerdo con las leyes de valores federales de Estados Unidos. No se garantiza que se vayan a obtener los resultados futuros y es posible que los resultados reales difieran considerablemente de las previsiones, estimaciones y <b>datos de resumen</b> contenidos en el sitio

Tabla 41. Cuadro resumen de formas lingüísticas asociadas a la superestructura del bloque legal (F).

#### **10.4. Resumen y discusión sobre análisis adicional de los bloques «contacto» y «legal»**

Una vez finalizado nuestro análisis contrastivo de estos dos bloques, nos es imprescindible detenernos a discutir los resultados obtenidos y, ya que el cuerpo de resultados es suficientemente amplio y representativo, proceder a contrastarlos con las hipótesis de partida.

Retomamos aquí el repetido *desideratum* del sector de la localización de producir textos similares a los originales de la región meta, «the ultimate aim of releasing a product that looks like it has been developed in country» LISA (2004: 11). Este aspecto cristalizó en nuestra hipótesis de trabajo sobre las diferencias cuantificables empíricamente debidas a los distintos contextos de producción (Baker 1995; 1996; 1998). Tras el análisis contrastivo que hemos realizado en este capítulo, podemos sin lugar a dudas concluir que, a pesar de las múltiples similitudes, los textos localizados presentan diferencias en todos sus niveles que resumimos en los siguientes aspectos:

(1) **La influencia del entorno sociocultural de partida en los textos localizados.** En primer lugar, las diferencias existentes entre ambos perfiles textuales quedan ampliamente recogidas en los textos pertenecientes al bloque **legal** (F), diferencias que atribuimos a los distintos contextos socioculturales de producción. Nuestro análisis superestructural relativo a las unidades y subunidades con mayor diferencial de frecuencia entre ambos perfiles textuales muestra datos reveladores:

La diferencia más significativa en el conjunto surge de una media de palabras en los textos localizados superior en 200% a los originales. Este hecho se proyecta en la presencia de la subunidad significativa la apelación a leer el texto (F-1-I-b), que conlleva el mayor diferencial de frecuencia entre ambos textos, 52,01%. De igual modo, el carácter de texto multinacional conlleva la mayor frecuencia de inclusión

de la unidad significativa sobre la legislación aplicable, ya que la española no es la implícita (F-1-V).

<b>Textos localizados &gt; originales</b>	<b>% dif.</b>	<b>Textos originales &gt; localizados</b>	<b>% dif.</b>
-Apelación a leer el texto (F-1-I-b)	52,01	-Registro legal inscripción	-38,80
-Modificación o cambio de términos (F-3-III)	51,56	-Ley española LOPD 1999	-36,01
-Enlaces hacia o desde el sitio web (F-2-IV)	43,56	-CIF	-22,40
-Legislación Aplicable y Jurisdicción (F-1-V)	42,00	-Derechos de acceso, rectificación, cancelación y oposición	-18,91
-Aceptación de términos legales (F-1-II)	34,99	-Ley 34/2002, de Serv. de la Sociedad de la Inf. y de Com. Elect.	-16,12
-Protección de menores (F-II-11)	31,58		

**Tabla 42. Unidades y subunidades con mayor diferencia de frecuencia entre ambos subcorpus legales.**

En todo caso, el hecho más significativo sería que todas las unidades con mayor frecuencia en los textos originales se refieren al marco legal español, tanto en las leyes vigentes aplicables, como en el estatus jurídico de las empresas en su inscripción en el registro o su CIF. En este sentido, los textos originales incluyen referencias a la ciudad en la que se encuentran los tribunales elegidos, mientras que los localizados mencionan exclusivamente jurisdicciones de países o regiones.

Estos resultados atestiguan que la presencia exclusiva de unidades o subunidades exclusivas de los textos localizados revela el contexto sociocultural de origen. En el bloque legal, éstos presentan en exclusiva una sección de bienvenida al texto (F-1-I-a), que desde un punto de vista lógico sería inapropiado puesto que el usuario ya ha visitado otras partes del sitio web.

La influencia del entorno de partida queda patente en nuestro análisis del bloque contacto (B). En los formularios de contacto (B-III) hemos mostrado cómo sólo los textos localizados presentan el campo «apellido» en singular (33,33%), reflejo de



entorno estadounidense, al igual que surgen los campos «tratamiento» (19,05%) e «inicial» (9,52%), inexistentes en los textos originales.

(2) Diferencias en el **plano léxico**. En primer lugar, el análisis de palabras clave o *keywords* del subcorpus legal localizado muestra diferencias significativas en el contenido léxico, y que reflejan las divergencias en la superestructura presentadas anteriormente. Las palabras «privacidad», «términos» y «declaración» presentan una frecuencia significativamente superior en los textos localizados, fruto de la mayor frecuencia de las secciones que las representan en la superestructura. Como ejemplo, la sección de privacidad (F-2) está presente en el 13,37% de los textos originales y en el 70,52% de los localizados.

Por su parte, los textos originales muestran una mayor influencia de la terminología convencional empleada en la legislación vigente en España. En la análisis de la Tabla 31 se muestra la influencia de la «Ley Orgánica 15/1999, [de 13 de diciembre], de Protección de Datos de Carácter Personal», en la que los primeros puestos los ocupan combinaciones que incluyen esta terminología. Igualmente, el análisis del compuesto sintagmático «propiedad intelectual», incluido en la denominación de la ley española en vigor, presenta una frecuencia de 52,38% en los textos originales y sólo 31,81% en los localizados. En este sentido, el contexto legal estadounidense quedaría patente en la mención a ciertas leyes de este país o en la inclusión de realidades jurídicas específicas de este contexto (F-3-I-c), tal como la inclusión de dieciséis posibles clases tipos de daños inexistentes en el corpus original.

Por último, quizás la diferencia de mayor interés sea el mayor uso en los textos originales de ciertos préstamos no adaptados como «email» o «link». En los formularios de contacto (B-III), el campo que incluye el correo electrónico presenta este anglicismo en un 85,36% de textos originales frente al 47,62% en

localizados. De igual modo, el anglicismo «link» está presente en el 21,38% de las unidades sobre enlaces originales (F-2-VI), e inexistente en los localizados. Este hecho podría apuntar a una mayor corrección en los textos localizados en este sentido, o a su carácter de textos localizados producto de un proceso de traducción. No obstante, el abuso de préstamos y calcos léxicos, sintácticos, tipográficos etc. en general es más extendido en los textos localizados.

(3) Diferencias en el **plano morfosintáctico**. En primer lugar, una de las diferencias más significativas en este plano es la mayor explicitación del pronombre «usted» en los textos localizados. Este pronombre tiene una frecuencia global en los textos de 0,44%, frente a 0,065% en los localizados. En contraste, este pronombre es 6,76 veces más frecuente en los textos localizados.

En segundo lugar, las diferencias referentes a los usos verbales serían igualmente significativas. Por ejemplo, el uso de formas en modo imperativo es consistentemente superior las mismas unidades o subunidades en los textos localizados. En la unidad sobre la apelación a la lectura (F-1-I-b), los textos localizados muestran un 62,55% de uso imperativo del verbo «leer», y un 25,92% de su forma infinitiva, mientras que esta tendencia se invierte en los textos localizados con un 9,09% de uso del imperativo y un 81,81% de uso de la forma infinitiva, sobre todo en perífrasis verbales. Resultados similares se obtuvieron en la subunidad de envío y validación de los formularios (B-III-i), o en la unidad de ofrecimiento de la información de contacto (B-I). En ésta, los textos localizados muestran una preferencia absoluta por las formas verbales en modo imperativo en la unidad fraseológica «póngase en contacto», mientras que en los originales prima la modulación con distintas perífrasis verbales.

En esta misma área se sitúa el mayor uso de formas verbales futuras en los textos localizados. En el análisis de la unidad fraseológica «someterse a la legislación» en

la unidad (F-I-V), los textos localizados mostraron un 37,5% de uso de formas futuras, mientras que este tiempo no se emplea en los originales. En esta misma unidad, el análisis de la forma lingüística convencional «regirse por la legislación» muestra un índice de uso de formas en futuro del 68,42% en los textos localizados y de 36,36% en los originales. Esta mayor frecuencia de las formas verbales futuras en segmentos que cumplen la misma función entre ambos textos se repite en nuestro análisis de los usos del lema [responsab\*] dentro de la unidad significativa F-III-2.

(4) Mayor **variación léxica y morfosintáctica** en los textos localizados. Ésta constituiría una de las principales diferencias que hemos observado en nuestro análisis. De hecho, la mayor variación léxica y morfosintáctica queda patente tanto en aquellos segmentos textuales de mayor frecuencia en uno u otro componente del corpus, como en aquellos cuya frecuencia sea similar. Como ejemplos de variación en el plano léxico, el análisis del préstamo «email» es esclarecedor. En los formularios de contacto (B-III) se presentan tres posibles variaciones para este concepto, mientras que son cinco las identificadas en los textos localizados. En este mismo sentido, el compuesto sintagmático nominal «número de teléfono» aparece en el 18,76% de los textos localizados y es inexistente en el otro componente del corpus. En esta unidad encontramos otro caso de gran variación, el segmento de borrado de los datos (B-III-H-2). En los textos originales el término «borrar» aparece en el 84,21% de los casos, mientras que en los localizados su frecuencia es de 27,27%, con la aparición de siete opciones alternativas inexistentes en los primeros.

Otro ejemplo de variación se recoge en el análisis de las posibles colocaciones del verbo «regirse» en la unidad F-I-V. En los textos originales este verbo se presenta seguido de la preposición «por» el 89,81% de casos Sin embargo, en los textos localizados esta preposición se encuentra sólo en el 42,1% de casos, apareciendo

además «de acuerdo con», 21,01%, «de conformidad con», 15,78% y finalmente «según», con un 10,52%.

Finalmente, otros ejemplos claros de variación los hemos encontrado en la unidad sobre el titular del sitio web (F-I-VI), en la gran cantidad de formas alternativas para limitar la responsabilidad (F-3-1). En la Tabla 38 se recogen veintitrés posibles formulaciones inexistentes en el corpus original.

(5) Presencia de **convenciones tipográficas** del entorno sociocultural de partida. En los textos localizados hemos identificado un número significativo de convenciones tipográficas propias de los textos legales estadounidenses con presencia en las traducciones. Entre ellas, destacan (1) el uso del paréntesis y las comillas ("...") en la identificación de las partes contratantes o los términos que se definen en el documento, «Intel Corporation ("Intel") proporciona los materiales de este sitio web (el "sitio") como un servicio para sus clientes» [Intel], «Por el hecho de utilizar o acceder al sitio web, usted y la entidad de la cual usted es representante autorizado ("Usted") se comprometen a obligarse por estas condiciones de uso» [UPS], así como el uso de comillas en conjunción con anglicismos como «as is», (2) o el uso de mayúsculas en oraciones o párrafos completos para resaltar su importancia, un uso sin presencia en los textos originales y no recomendado en los textos web (Price y Price 2002):

«Sin limitar lo precedente, todo en el Sitio Web le es proporcionado en el estado en que se encuentra, SIN GARANTÍA DE CLASE ALGUNA, EXPRESA O IMPLÍCITA, INCLUYENDO, SIN LIMITARSE A LAS MISMAS, LAS GARANTÍAS IMPLÍCITAS DE COMERCIALIZACIÓN, DE SU UTILIZACIÓN PARA FINES CONCRETOS O RELATIVAS A INEXISTENCIA DE INFRACCIONES.»

Como resumen, podemos concluir que en el proceso de localización se producen unos textos con unas características diferenciadas que hemos podido cuantificar empíricamente mediante la metodología de la Lingüística de corpus. En este

sentido, nuestra hipótesis de partida quedaría confirmada: los textos localizados presentan diferencias recurrentes que hemos constatado. No obstante, en la caracterización y descripción contrastiva de los textos localizados hemos ido observando una diferencia contrastiva primordial que marca lingüísticamente los mismos. Nos referimos a los errores de traducción que hemos ido mencionando a lo largo de nuestro análisis. Como paso final en nuestro análisis, creemos pertinente analizar qué tipo de errores hemos identificado, así como examinar si la presencia de ciertos errores se debe a deficiencias específicas en sus procesos de localización en general, y no exclusivamente a la tarea de traducción en sí.

## **11. LOS ERRORES DE LOCALIZACIÓN Y LAS CONVENCIONES DE GÉNERO EN LOS TEXTOS LOCALIZADOS**

El análisis contrastivo que hemos realizado ha supuesto una exploración sobre el papel que desempeñan el género, las diferencias superestructurales y sus formas lingüísticas convencionales en el proceso de traducción web. En principio, ambos conjuntos textuales deberían responder a un mismo género textual, si bien, existen diferencias sustanciales que hemos puesto de relieve en nuestro análisis contrastivo. Entre estas diferencias, una referencia constante a lo largo de los capítulos previos se ha concretado en los errores o inadecuaciones presentes en el componente localizado del corpus. Para completar nuestro estudio y como último paso, procederemos a explorar si éstos podrían constituir marcadores que diferencien textos originales y localizados pertenecientes a un mismo género. En caso positivo, los distintos contextos de producción conllevarían la presencia de inadecuaciones o errores exclusivos de los textos localizados en virtud de su dependencia de un contexto de producción marcado. Este hecho constituiría otra divergencia intratextual entre ambos conjuntos textuales, y podría apuntar a la posible existencia de géneros textuales diferenciados como resultado de su dependencia de un proceso de traducción - localización.

En general, el estudio de los errores de traducción se ha realizado en Traductología con objetivos mayoritariamente didácticos o evaluativos (Nord 1997a, 1996, 1991; Hurtado Albir 1999; López Rodríguez y Tercedor Sánchez 2004, 2008; Larose 1998; Martínez Melis y Hurtado Albir 2001; Lörscher 1991; Seguinot 1990; Kussmaul 1995). Puesto que existe una amplia bibliografía en esta área, dejaremos

conscientemente a un lado las discusiones existentes sobre la naturaleza de los errores de traducción, sus implicaciones, su importancia en el proceso cognitivo traductor, el impacto o graduación de cada tipo de error<sup>460</sup>, o las distintas perspectivas desde las que se ha abordado. Puesto que la noción de error de traducción está íntimamente ligada a la de problema de traducción (Nord 1997a; Lörcher 1991), definiremos operativamente los errores de traducción desde una perspectiva cognitiva y funcional como un problema de traducción que no se ha resuelto convenientemente<sup>461</sup> (Nord 1991: 96-100). Por su parte, un problema de traducción constituiría «an objective problem which every translator [...] has to solve during a particular translation task» (Nord 1991: 151).

En cualquier caso, en el contexto de la localización hemos de diferenciar los errores de traducción (Nord 1991; Lörcher 1991) desde esta perspectiva cognitiva - funcional, de aquellos fruto de deficiencias durante el proceso de localización o internacionalización, y que denominaremos «errores de localización». Estos errores, cuya existencia se ha apuntado anteriormente (Pierini 2007), no son atribuibles directamente a la tarea específica de traducción, sino más bien a deficiencias en el ciclo de la localización o en la colaboración entre los diferentes agentes que intervienen en el proceso (véase 2.4). El control de este segundo tipo de errores se ha asociado a la adquisición de mayores competencias tecnológicas por parte de los localizadores, evitando así en este ciclo una pronunciada división de tareas que puede suponer un mayor número de errores (Bass 2006).

---

<sup>460</sup> En este sentido, consúltese por ejemplo Martín Melis y Hurtado Albir (2001: 281-282).

<sup>461</sup> Los problemas de traducción se encuentran asimismo asociados a las estrategias de traducción. Desde una perspectiva cognitiva Lörcher apunta «Translation problems are of eminent importance because they function as both the starting point and the reason for the use of translation strategies. As a consequence, translation strategies only occur when translation problems exist» (Lörcher 1991: 201). En este mismo sentido el investigador define las estrategias de traducción como: «a translation strategy is a potentially conscious procedure for the solution of a problem which an individual is faced with when translation a text segment from one language to another» (ibid: 76).

Dada la gran cantidad de errores identificados en los textos localizados, el objetivo del presente capítulo será exclusivamente determinar qué tipo de errores diferencian contrastivamente los textos localizados de los originales, haciendo especial hincapié en aquellos relacionados con las convenciones de género. Para establecer una base sobre la que articular este análisis, hemos recurrido a la tipificación de errores con objetivos didácticos de Hurtado Albir, que a nuestro parecer supone una propuesta suficientemente clara para los objetivos de la presente revisión<sup>462</sup>. Defendemos, además, su pertinencia por que en la misma se incluyen explícitamente las inadecuaciones relacionadas con el género textual (Hurtado Albir 2001: 305-306). La investigadora agrupa a los errores en tres tipos diferenciados: (1) las inadecuaciones relacionadas con la comprensión del TO<sup>463</sup>, (2) las inadecuaciones que afectan a la expresión en la lengua de llegada<sup>464</sup>, y por último, (3) las inadecuaciones pragmáticas, aquellas que surgen por una desconexión con la finalidad de la traducción<sup>465</sup>. Dentro de este último grupo se imbricarían aquellos errores de traducción referentes al género textual y a sus convenciones lingüísticas asociadas. Éstas son objeto específico de análisis en investigaciones contrastivas basadas en géneros (Gamero 2001: 184-196).

Volvemos, en primer lugar, al punto de partida y criterio de mayor rango jerárquico de nuestro análisis intratextual, la superestructura (véase 5.3.1). Si nos centramos en este aspecto, nuestro análisis ha hecho patentes las diferencias superestructurales entre los sitios web originales y localizados (véase, 9.1). Dados

---

<sup>462</sup> La investigadora señala que esta tipificación se dirige a la traducción escrita con el objetivo de constituir un baremo de uso didáctico en la docencia de la traducción (Hurtado Albir 2001: 305 1999: 120) el objetivo de esta tipificación.

<sup>463</sup> Los problemas por la comprensión del texto original pueden ser omisiones, falsos sentidos, no mismo sentido, adición, supresión, referencia extralingüística mal solucionada e inadecuación de variación lingüística (tono, dialecto, idiolecto, etc.).

<sup>464</sup> Errores en la ortografía y puntuación, gramática, léxico, aspectos textuales (coherencia, progresión temática, elementos de referencia, conectores) y redacción (formulación defectuosa o poco clara, falta de riqueza expresiva, pleonasmos, etc.).

<sup>465</sup> Nord (1997) señala a estos errores como los de mayor importancia. Hurtado Albir (2001), basándose en las propuestas funcionalistas de esta autora señala que las inadecuaciones en este área están en relación con el tipo de encargo, el destinatario al que va dirigida), el método elegido, el género textual y sus convenciones, etc.



los resultados obtenidos podemos concluir que al traducir y localizar los sitios web corporativos no se respetan las convenciones del género referentes a la superestructura en la lengua de llegada, el español. En los textos localizados existe una frecuencia en la presencia de bloques y secciones que reflejan claramente la estructura de los TO estadounidenses. Esta tendencia es además más acusada en los textos legales (véase 10.3). Nuestros resultados están en consonancia con los presentados por Gamero en otro género textual<sup>466</sup> y en la combinación alemán-español (2001: 187).

En referencia a esta tendencia, secundamos la opinión de la investigadora que defiende la importancia de respetar las convenciones superestructurales de cada género ya que, a nuestro parecer, los textos localizados deberían en principio ajustarse lo más posible a las expectativas del receptor. No obstante, somos conscientes de las peculiaridades y limitaciones impuestas en proceso de localización, puesto que habitualmente el encargo de traducción puede impedir o desaconseja implícitamente cambios superestructurales. Este hecho se debe a que los cambios requieren la asistencia de desarrolladores y programadores, con lo que se aumenta el tiempo y los recursos necesarios para producir la versión localizada del texto.

### **11.1. El proceso de localización y los errores de traducción**

En general, el análisis de los errores de traducción en los sitios web localizados ha de contextualizarse dentro de un proceso global caracterizado por la autoría múltiple del texto, la gradual desaparición del texto en su conjunto como unidad textual de traducción (Jiménez 2007; Biau y Pym 2006; Bowker 2006; Shreve 2006; Bedard 2000), la influencia de las herramientas informáticas empleadas (Biau y Pym 2006), o la existencia de una etapa de edición o control de calidad con unas

---

<sup>466</sup> El género textual es el manual de instrucciones.

características específicas que no se ha investigado a fondo (Bouffard y Caignon 2006). La relativa mayor presencia de errores de traducción a todos los niveles en textos localizados se ha relacionado en investigaciones previas con:

- La ausencia de una fase de control de calidad de suficiente importancia (Corte 2002; Bouffard y Caignon 2006).
- Los aspectos organizativos relativos a plazos de entrega muy limitados en el tiempo (Bass 2006; Bolaños 2003).
- La relajación de las normas gramaticales en los textos web (Posteguillo 2003).
- La desaparición de la figura del texto unitario que conlleva una traducción dentro del microcontexto (Shreve 2006).
- Problemas tecnológicos asociados a las fases de localización o internacionalización, y no al proceso de traducción en sí (Bass 2006).
- Problemas de calidad en los textos originales a localizar (Bass 2006).
- La traducción de textos web por traductores de distinto grado de profesionalización o falta competencias específicas (Pierini 007; Reineke y Sánchez de León 2005: 345).

En este capítulo nos limitaremos a describir y tipificar los errores identificados que son de interés desde un punto de vista contrastivo. Continuando con el carácter descriptivo de nuestra investigación, no profundizaremos en las posibles deficiencias específicas que causan estos errores ya que una investigación a fondo quedaría fuera de las posibilidades de nuestro estudio. En un primer lugar, y puesto que el punto de partida de nuestra investigación lo constituye el género textual y sus convenciones, analizaremos los errores de traducción atribuibles a las inadecuaciones pragmáticas en los textos localizados.

## **11.2. Las formas lingüísticas convencionales y los menús de navegación**

Los menús de navegación ocupan una posición prominente en el análisis del conjunto de formas lingüísticas convencionales asociadas a la superestructura del género. El respeto a las convenciones establecidas en los mismos es de suma importancia dado su carácter metatextual y metacomunicativo (Nielsen y Loranger 2006; Nielsen y Tahir 2002; Brink *et al.* 2003). Hemos demostrado previamente el alto grado de convencionalización en la terminología en el género sitios web corporativo español (véase 9.1.3), si bien nuestros resultados apuntan a un grado de convencionalización terminológica inferior en comparación con los sitios corporativos en el contexto geográfico estadounidense (Nilsen y Tahir 2002). Como apuntamos en el capítulo 9, las unidades léxicas de mayor frecuencia para la denominación de bloques y secciones presentan una media en su frecuencia de uso del 58,26% en los textos originales, toda vez que la media de frecuencia de estas mismas formas convencionales se reduce al 34,38% en los textos localizados (véase 9.1.3). La media de frecuencia de la unidad léxica con mayor frecuencia propia de los textos localizados presenta un porcentaje del 43,49%, con lo que hemos constatado constatar la mayor variación lingüística y menor uso de las formas lingüísticas convencionales en los textos localizados.

Esta relativa falta de atención a las convenciones textuales en estos subtextos conlleva en parte la presencia de un porcentaje significativo de inadecuaciones. El análisis de la terminología empleada en los menús de navegación arrojó un porcentaje de formas lingüísticas inadecuadas del 6,82%, si bien, este porcentaje incluye exclusivamente (1) las inadecuaciones debidas a la comprensión del TO y (2) las inadecuaciones que afectan a la expresión en la lengua de llegada. No hemos incluido aquellas unidades léxicas que aparecen en inglés, ya que nos resulta imposible desde nuestro punto de vista discernir qué unidades léxicas constituyen

una decisión del traductor-localizador y qué unidades léxicas en inglés responden una decisión consciente que forme parte del encargo de traducción. Dos ejemplos serían «My Sun» [Sun] para el área de clientes o «Pro» [Apple]<sup>467</sup>, un área dirigida a profesionales. En cualquier caso, sólo consideramos anglicismos no adaptados<sup>468</sup>, como «links» o «email» o «home», a aquellos términos que estén representados ambas secciones del corpus comparable. Así, defendemos que el mantenimiento de una unidad léxica en inglés en un texto localizado puede responder a una estrategia consciente del iniciador<sup>469</sup> y no podría considerarse a priori como un error de traducción.

### **11.3. Errores de traducción en los menús de navegación**

Dada la importancia del análisis de unidades léxicas en menús de navegación a lo largo de nuestra investigación, sobre todo por su alto grado de convencionalización, comenzaremos analizando los errores identificados en estos subtextos.

#### **11.3.1. Errores de coherencia terminológica**

La coherencia terminológica constituye uno de los parámetros principales de calidad según la norma alemana de calidad en traducción DIN 2345 (Bass 2006)<sup>470</sup>. La existencia de estos errores de coherencia terminológica en localización queda recogida en investigaciones previas (Corte 2002; Bolaños 2003). Éste sería uno de los errores recurrentes de mayor importancia que hemos identificado, dado el

---

<sup>467</sup> «Pro» constituiría una abreviación convencional del sustantivo «Professional».

<sup>468</sup> Denominamos «anglicismos no adaptados» siguiendo a Aguado de Cea (1994) y a Belda (2003), aunque este último se decanta a favor de «préstamos». Aguado de Cea (1994: 8) aboga por utilizar «anglicismo» en lugar de «préstamo» ya que este término no incorpora la lengua de procedencia del término en cuestión.

<sup>469</sup> Normalmente, se alude a razones de marketing para el mantenimiento de términos en inglés.

<sup>470</sup> El resto de parámetros de calidad son: presentar una traducción completa, un estilo y uso gramatical apropiados y el seguimiento de una guía de estilo acordada entre ambas partes (Sturz 1998, cf. Bass 2006)

valor metatextual de las unidades léxicas en los menús de navegación. Este tipo de inadecuaciones se suele reflejar en la falta de coherencia terminológica entre los menús de navegación y el título u otra referencia metatextual de una página determinada o subtexto determinado. Estas inadecuaciones están íntimamente relacionados con las formas lingüísticas convencionales del género en cuestión, puesto que independientemente de la existencia de un traductor único o múltiple, el conocimiento y puesta en valor de las convenciones de género establecidas en la lengua meta reduciría, a nuestro parecer, este tipo de errores<sup>471</sup>.

Presentamos con carácter ilustrativo una lista representativa aunque no exhaustiva, de los errores de coherencia terminológica que se han identificado en estas unidades léxicas de carácter metadiscursivo.

- Falta de coherencia por el uso indistinto de las preposiciones «acerca de» y de «sobre» en la unidad léxica relacionada con el bloque empresa (C). Este error se observó en los sitios web de EMC y de Black&Decker. La unidad léxica «acerca de [la empresa]» constituiría, según nuestros resultados, un anglicismo de frecuencia en los textos localizados. La forma lingüística convencional en los textos originales sería el término «empresa» o «quiénes somos» (véase 9.1.2.3.1):



---

<sup>471</sup> En este sentido, esperamos que la presente tesis o una posible guía de estilo que podríamos desarrollar, sea de utilidad práctica en la labor diaria de los localizadores web.



Figura 114. Falta de coherencia en los menús de navegación. Sitios webs de EMC y de Black&Decker.

- Co-presencia de formas traducidas y originales en inglés en los textos localizados. Un ejemplo sería el uso de «mapa del web» y «site map» de forma indistinta en el sitio de Arrowiberica:

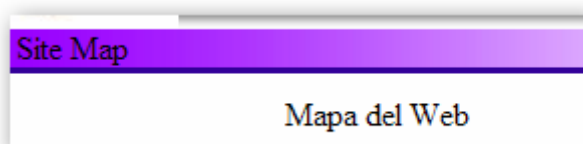


Figura 115. Falta de coherencia en los menús de navegación. Sitio web de EMC.

- La sección de protección de datos (F2) presenta los mayores índices de variación. (1) En una misma página web de Xerox, se emplean cuatro unidades léxicas diferentes referentes a una misma sección o subtexto: «política de privacidad», «política de confidencialidad y seguridad», «directiva de privacidad», «declaración de privacidad». (2) El sitio web de CSC presenta un enlace en inglés hacia esta sección, «Privacy Policy», el título de la sección es «Política de privacidad» y en el texto se refiere a la misma como «Política de intimidad», tres unidades léxicas diferentes. (3) En el sitio web de EA el enlace hacia el bloque legal reza «política de privacidad» mientras que el título del bloque textual es «normativa de privacidad». Finalmente, (4) ilustramos este punto con una captura del sitio web de Mattel, que muestra «política de privacidad» y «declaración de privacidad», junto con la variación en la

traducción que se aprecia en el anglicismo puro y adaptado «online» y «en línea».

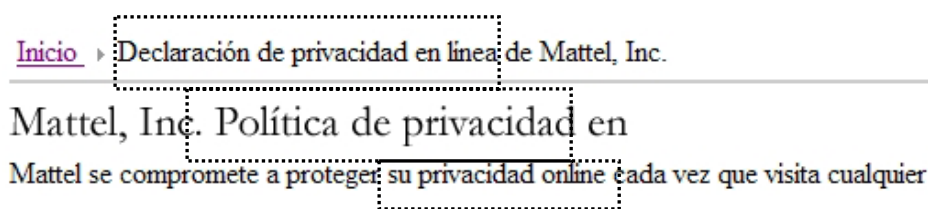


Figura 116. Falta de coherencia en los menús de navegación. Sitio web de Mattel.

- Otro ejemplo ilustrativo lo constituye la unidad léxica en el menú de navegación de ADM, «cumplimiento y ética», que en la página en sí reza «conformidad y ética».

### 11.3.2. Errores de expresión en la lengua de llegada en menús de navegación

Este análisis se realizará según el orden de bloques y secciones identificadas y se señalarán, de modo ilustrativo y no exhaustivo, las unidades léxicas erróneas presentes en los menús de navegación. En cada caso señalaremos una posible corrección al error identificado.

- Bloque B, «contacto». (1) «Contacte a Verisign» [Verisign], un error sintáctico ya que esta unidad léxica debería incluir la preposición «con». (2) Las siguientes unidades léxicas «Póngase en contacto con nosotros» [Alcoa] y «Entre en contacto con nosotros» [Lucent] constituirían errores pragmáticos ya que los segmentos textuales en menús de navegación han de ser lo más breves y sucintos posibles. La forma lingüística convencional en los textos originales sería «contacto» (véase 9.1.2.2.1).

- Bloque C, «empresa». La unidad léxica «acerca de nosotros/ [nombre de empresa]» constituiría un error de traducción, puesto que, según nuestros datos constituiría un anglicismo de frecuencia, calcando además la estructura sintáctica de la lengua de partida (véase 9.1.2.3.1). Las unidades léxicas de mayor grado de convencionalización en los textos originales, «Empresa», o «Quiénes somos», supondrían mejores opciones en este caso. Otros errores serían:
  - «Datos cortos», traducción literal de «quick facts», un anglicismo léxico.
  - «Corporativo» sería un uso inadecuado de una forma adjetival, puesto que esta forma se expresa convencionalmente a través de sustantivos. La solución sería de nuevo la forma lingüística más convencional: «La empresa».
  - Sección C1, «Localización». Uso del término «Localidad» [Bostoncientifics], que sería un sentido inadecuado al referirse a una población.
  - Sección C2, «Experiencia de la empresa». En los textos localizados son recurrentes los anglicismos léxicos que incorporan la palabra «casos», inexistente en el corpus original en estos segmentos «Casos de éxito», «Casos de clientes», «Casos concretos», «Estudios de casos»<sup>472</sup>. Aunque no se identificó en el corpus original, «historias de éxito» supondría una traducción más adecuada. Igualmente se podría emplear el término «clientes» para señalar la experiencia de la empresa, puesto que es la unidad léxica de mayor uso en esta sección en el corpus original.
- Bloque D, «noticias». La traducción «librería» [Ernst&Young] constituiría un barbarismo al ser un falso amigo del inglés «library», en lugar de «biblioteca». Según el corpus original «publicaciones» sería la unidad léxica más apropiada.

---

<sup>472</sup> Esta unidad léxica sería una traducción literal de «case studies».



- Bloque E, «Productos y Servicios». En la sección de ofertas (D3) se observa «Gangas de Dell», que consideraríamos inadecuado tanto por la carga semántica asociada a la palabra «gangas»<sup>473</sup>, como por la cacofonía que se produce en el uso de la preposición «de» frente al nombre de la empresa.
  
- Bloque F, «legal». (1) Por el carácter de forma lingüística convencional de la unidad léxica en singular «aviso legal», consideramos una inadecuación de traducción la presencia de su forma plural «avisos legales» [Alcoa, Xerox]. (2) En la siguiente unidad léxica, «Comunicados legales» [Hospira], el primer elemento está relacionado con el bloque noticias (C), ya que los comunicados constituirían textos informativos que emite la empresa (véase 9.1.2.4.1).
  
- Bloque G, «zonas de clientes». Dentro de las zonas de clientes se han identificado errores en la unidad léxica que autentifica a los usuarios, como «acredítese», errores sintácticos como «ingresar a... », o juegos de palabras que confundirían al usuario como «Usted S.A.» [Cisco] y que se desaconsejan por no ser unívocos y faltos de ambigüedad (véase 5.9.2).
  - Sección G1, «trabajo». El error más recurrente en esta sección lo constituye el anglicismo léxico «carreras» [Alcoa, Symantec], o unidades léxicas que incluyen este anglicismo como «carrera profesional» [PriceWaterHouseCooper]. Se identificó igualmente un anglicismo tipográfico en el uso de la arroba, @, con el sentido anglosajón que se lee tal como la preposición «at» o «en» en español, «trabaje@dell», [Dell], «Trabajo@EMC» [EMC] (véase Figura 74, p. 402).
  - Sección G10, «Inversores». Inadecuación al usar la forma adjetiva «inversionista» en lugar de la forma sustantiva «inversores», mayoritaria en los textos originales: «relaciones con los inversionistas» [AMD], «relaciones

---

<sup>473</sup> La definición de «ganga» en el diccionario de la RAE incluye esta posible carga semántica negativa «f. Cosa apreciable que se adquiere a poca costa o con poco trabajo. U. m. en sent. irón. **para designar algo despreciable, molesto.**».

de los inversionistas» [Intel]. Esta última incluiría un error sintáctico en el uso de la preposición «de» que resulta en un contrasentido.

- Bloque H, «interactividad».
  - Sección H1, «mapa web». La unidad léxica «mapa de la página web» [Caterpillar] constituye un error puesto que el mapa representa la superestructura del conjunto del sitio web, y no de una sola página. Igualmente, «mapa del site» [Bristol-Myers] supondría la introducción de un anglicismo puro en este compuesto sintagmático.
  - Sección H2, «buscar». El carácter verbal y sustantivo de la forma inglesa «search», conlleva la mayor inclusión de la forma sustantiva en los textos localizados que en los originales (véase 9.1.2.8.1). No obstante, el error de mayor importancia supone la inclusión del verbo «Ir» en el botón que activa la búsqueda, que se corresponde con el inglés «Go!». Tal como el siguiente ejemplo de ADM:



Figura 117. Error en la sección H2, «buscar». Sitio web de ADM.

En cualquier caso, los siguientes casos del sitio web de Otis y Alcoa incorporan la preposición «a» tras el verbo «ir», en un claro ejemplo de un error de localización debido a una posible deficiencia en la fase de internacionalización. Este botón interactivo debería colocarse en cualquier caso antes del cajón de búsqueda. Esta inadecuación, a activación de una búsqueda con «Ir a», o «Ir» aparece en el 4,2% de los sitios web localizados.



Figura 118. Error en la sección H2, «buscar». Sitios web de Otis y Alcoa.

## 11.4. Otros errores de expresión en la lengua de llegada en los sitios web localizados

En general, los textos localizados presentan gran cantidad de errores tanto en el uso de las formas convencionales como inadecuaciones en la expresión en la lengua de llegada. La naturaleza de este abundante conjunto de errores abarca tanto el plano léxico semántico, el morfosintáctico y el estilístico. No obstante, la presencia de ciertos errores como los ortotipográficos no es específica de los sitios web localizados, puesto que hemos identificado un número significativo de errores con esta naturaleza en los textos originales. Como ejemplo, la presencia de ciertos préstamos léxicos o la ausencia de tilde constituirían dos errores de expresión recurrentes en ambos conjuntos textuales. En cualquier caso, la naturaleza de los errores de traducción es diferenciada y queda recogida en los siguientes aspectos que organizamos según los niveles de análisis textual léxico-semántico, morfosintáctico y estilístico:

### 11.4.1. Nivel léxico-semántico

Este tipo de errores puede conllevar dificultades de comprensión para el usuario del texto, ya que estos errores pueden comprometer la necesaria claridad y falta de ambigüedad recomendada en estos textos (véase 5.9.1).

En primer lugar mencionamos la presencia de **préstamos léxicos adaptados, no adaptados y barbarismos**<sup>474</sup>. La presencia de préstamos léxicos quedó ilustrada en el epígrafe sobre los errores en los menús de navegación. En el caso de los barbarismos, siguiendo a Martínez de Sousa los definimos como «palabras

---

<sup>474</sup> Entendemos por barbarismo «palabras generadas según normas morfológicas y fonológicas que no son propias de una lengua o que no han sido admitidas por el uso considerado correcto» (Martínez de Sousa 2007: 141).

generadas según normas morfológicas y fonológicas que no son propias de una lengua o que no han sido admitidas por el uso considerado correcto» (Martínez de Sousa 2007: 141). Son muchos los ejemplos de barbarismos léxicos identificados, como «**Tope** de la Página» [Kimberly-Clark] que reflejaría el original «Top of page», que se podría traducir por «Principio»/«ir al principio», o «Si quieres **reportar** enlaces rotos» (Sun). Igualmente, se encuentran palabras como «apoyar» [lucent.com] en el menú de navegación, traducción de «support», que en español se traduciría por «asistencia» o «ayuda»<sup>475</sup>. Otro de los barbarismos léxicos más recurrentes en la combinación inglés – español consiste en la traducción de «experience», en el sentido de una vivencia previa, como en el caso de «la experiencia del café» [Starbucks] o « ¿Ha experimentado un problema con...?»[AmericanExpress]. Existen igualmente barbarismos «adaptados» como se observa en este caso, «Avaya no es responsable del contenido de ninguna página **linkada**» [Avaya].

Dentro de los barbarismos, se incluiría el uso de **neologismos** anglosajones de difícil traducción al español, pero que se mantienen en este idioma sin razón aparente. En su traducción se han observado tres tendencias. Por un lado (1) los textos pueden incluir el término en español seguido del término o unidad léxica en inglés entre comillas, como «algunos datos pueden quedar en el sitio en forma residual debido a la existencia de archivos de soporte ("**backup**")» [UPS], «o los cuales se proporcionan "tal cual" ("**as is**") sin garantías de ningún tipo» [Stryker]. (2) Una segunda posibilidad es la inclusión del barbarismo entre comillas para señalar su falta de adaptación o incorporación a nuestra lengua, «se reserva el derecho a modificar en cualquier momento la presentación (incluyendo el "**look & feel**")»,» [Stryker]. Por último, (3) la forma más común de incorporar estos barbarismos o préstamos no adaptados sería su inclusión en inglés sin marca tipográfica alguna de su procedencia: «Externalizar el **hosting**» [ADP], «Aplicación

---

<sup>475</sup> Apoyar en este caso constituiría un sinónimo del calco «soportar», que Aguado de Cea recomienda «ayuda» o «asistencia» en su lugar (1994: 237).

de **Best Practices** Mundiales» [ADP], «Cómo reducimos los costes de implantación gracias a nuestro **know how**» [ADP], «EDS España reduce en un 90% los costes de **Help Desk** interno con **eTrust® Admin. De CA.**» [CA] «**Trials** y Descargas» [CA], «Mattel es libre de utilizar cualesquiera ideas, conceptos, **know-how** o [...]» [Mattel].

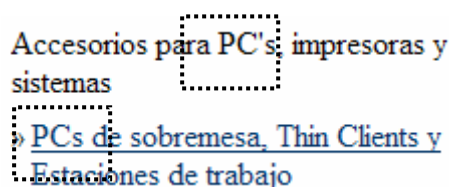
En segundo lugar, hemos identificado errores en el uso de **siglas, abreviaturas, y acrónimos**. Los manuales de estilo desaconsejan el uso de siglas y acrónimos en el contexto multinacional de los sitios web por su falta de univocidad y las dificultades de comprensión (véase 5.9.2). De entre las inadecuaciones identificadas, las siglas sin explicación son comunes en los corpus localizados, sobre todo en bloques y secciones no dirigidos a especialistas<sup>476</sup>. Este hecho dificultaría la comprensión del texto, tal como «Éstas dependen de la disminución de las insatisfacciones de los clientes, reflejadas en el **CQI**, y del aumento de su fidelidad, reflejada en el **CLI**.» [Sun]. Al igual que ocurre en los textos técnicos (Belda 2003), las siglas empleadas suelen ser las anglosajonas, con la posible inclusión de su traducción al español aún cuando no se corresponda en la formación con la unidad léxica traducida «Véase **PDA** (Ayudante Digital Personal)» [Motorola].

Se identificaron asimismo errores en la formación de siglas, sobre todo en su puntuación y en su pluralización. En el caso de la puntuación, en la actualidad es posible incluir abreviaciones sin puntos intermedios, tal como la misma frecuencia de uso de de CIF y C.I.F analizado anteriormente (véase 10.2.2.2), el siguiente ejemplo constituiría un error al incluir puntuación extra «Cambios para entrar a **E.E.U.U.** » [American Airlines]. Ésta sería una abreviatura que se escribe doblada por ser una palabra en plural, con lo que no serían apropiados los dos puntos adicionales (Martínez de Sousa 2007: 210).

---

<sup>476</sup> Como quedó patente en el análisis del bloque comunicativo «zonas especiales» (G), este género puede incorporar secciones dirigidas a especialistas, como desarrolladores, inversores, etc., que se expresarían por medio de un lenguaje técnico especializado. El uso frecuente de siglas y abreviaciones sería propio de este tipo de textos (Jiménez Serrano 2002; Vázquez del Árbol 2006).

Los errores por la pluralización de siglas conllevan la frecuente inclusión de anglicismos ortotipográficos, como la presencia del apóstrofe anglosajón. Martínez de Sousa (2007) recomienda no añadir una «s» en formar el plural. En todo caso, en el siguiente ejemplo se identifican las dos posibles tendencias, que supone además una inconsistencia en la pluralización de las siglas, con calcos de la estructura inglesa en la pluralización mediante el uso del apóstrofe, «PC's», junto con el recurso más frecuente en español que sería la simple adición de la s, PC (Aguado de Cea 1994: 18). El siguiente ejemplo se extrajo del sitio web de HP:



**Figura 119. Inconsistencia en la pluralización de siglas. Ejemplo de HP.**

Los textos localizados muestran asimismo barbarismos en siglas como «ID olvidado» [UPS]<sup>477</sup> o «número ID» [Acuvue.es]. Esta sigla se correspondería en español con el compuesto sintagmático nominal «nombre de usuario».

Por último, los errores de coherencia terminológica son muy abundantes, y redundan en la menor precisión el texto. Como ejemplo, en sitio web de Fedex se emplea indistintamente en menús de navegación y en cuerpo de página «seguimiento» y «rastreo», ambas traducciones de «tracking».

#### **11.4.2. Nivel morfosintáctico**

En este nivel los errores más frecuente gravitan entorno a la presencia de préstamos sintácticos y el calco de estructuras sintácticas anglosajonas, como «Ir

---

<sup>477</sup> Según Aguado de Cea (1994) este caso constituiría un anglicismo de siglas.

atrás» [Goodyear], traducción de «Go back». En ciertos casos, los calcos sintácticos de estructuras anglosajonas resultan en textos de ínfima calidad y de difícil comprensión, como el siguiente caso de la sección legal (F):

Este sitio web **es operado por** Cargill, Incorporated ("Cargill"). AL USAR ESTE SITIO WEB, USTED **CONVIENE Y ACEPTA LOS TÉRMINOS** Y CONDICIONES AQUÍ ESTIPULADOS. El uso que usted haga de este Sitio web se rige por estos Términos y Condiciones. Tómese unos minutos para revisar estos Términos y Condiciones. Al usar el sitio web de Agibrands, **usted está de acuerdo en seguir** estas reglas y **sujetarse a éstas**. Si no **está de acuerdo por estos** Términos y Condiciones, no **use** el sitio web de Agibrands. [NutrimentosPurina.es].

En este extracto textual se identifican errores que abarcan todos los niveles, aunque los errores morfosintácticos serían los de mayor peso. Los calcos sintácticos son abundantes «es operado por», «está de acuerdo en seguir», «está de acuerdo por estos Términos». A nivel oracional observamos su escasa longitud ya que constituyen una copia de la estructura oracional original. El uso de mayúsculas para enfatizar una parte del texto se ha identificado exclusiva y recurrentemente en las secciones legales en los textos localizados, a pesar de no recomendarse desde los manuales de estilo (véase 5.9.5), e igualmente se identifica una falta de coherencia en el uso de «Sitio web» con y sin mayúscula. El uso recurrente del determinante en la expresión «estos Términos y Condiciones» o «estas reglas» constituiría otro anglicismo sintáctico dada su alta frecuencia, ya que como vimos en el análisis de los subtextos legales, se corresponde con el uso del adjetivo «presentes» (véase 10.3.7.3).

En general, los errores gramaticales producen un mayor impacto negativo en la valoración del texto por parte de los usuarios. Éstos son abundantes en los sitios web localizados, de ahí que creamos pertinente la realización de futuras investigaciones sobre la fase de Control de Calidad en la localización web (Bouffard y Caignon 2006). En el corpus localizado hemos identificado:

(1) Errores de concordancia sujeto-verbo «Adobe and Acrobat **es una marcas** de Adobe Systems Inc.» [Hospira].

(2) Uso erróneo de formas verbales reflexivas, «**Contáctese** con UPS.»<sup>478</sup> [UPS]

(3) Uso erróneo de preposiciones, «Si tiene alguna pregunta sobre la presente declaración de privacidad, [...], puede contactar **a:**» [UPS], o «se encuentran protegidos por la legislación aplicable **a la** materia» [American Express].

(4) O finalmente errores de género gramatical: «Necesitamos poder contar con **una** equipo de expertos.» [ADP].



Figura 120. Error sintáctico relacionado con la variedad dialectal del español no propia de España. Sitio web de EMC.

Por cuento respecta a los sujetos verbales, el error más recurrente ya señalado en otras investigaciones lo constituye la falta de coherencia en el tratamiento del usuario (Tercedor 2005; Bolaños 2003), que supone la aparición del tratamiento formal e informal hacia éste en una misma página web. Los dos siguientes ejemplos son capturas de pantalla de Sun.com y de Motorola que ilustran este aspecto:



<sup>478</sup> El diccionario de la Real Academia no recoge forma reflexiva alguna de este verbo. Tras un análisis de esta forma verbal en Google.com, el uso reflexivo de esta forma verbal se ha identificado mayoritariamente con la variedad dialectal del español propia de Argentina y Chile.



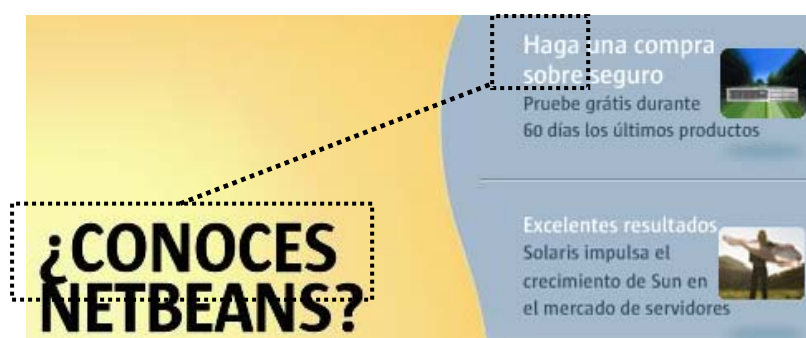


Figura 121. Error en la falta de coherencia en el uso del tratamiento formal e informal en la misma página. Sitios web de Motorola y Sun.

En el primer ejemplo se incluye en la misma página la forma verbal del imperativo «Mantén» junto con «Mándenlos». En el segundo caso esta falta de coherencia se produce en dos segmentos textuales que pertenecen a un mismo tipo textual, el exhortativo publicitario, y en la que se mezclan igualmente las formas verbales imperativas en tú y usted, «conoces» y «haga». En los textos originales las diferencias en el trato dentro de un mismo sitio web se relacionan con el cambio en las relaciones de poder, tal como cuantificamos estadísticamente en la preferencia por el tratamiento de «tú» en la sección «trabajo» (G1) (véase 4.2.1.3.1 y 9.1.2.7).

Dentro de los errores referentes al sujeto verbal, la mayor explicitación del pronombre «usted», y en menor medida «tú», podría identificarse como un anglicismo de frecuencia en los textos localizados. De hecho, el análisis expandido de la sección legal (F) arrojó un porcentaje de uso de este pronombre 6,5 veces superior en textos localizados que en originales. El mismo segmento textual del sitio de Cargill presentado al principio de este epígrafe constituiría otro ejemplo, ya que el pronombre «usted» aparece explicitado en dos ocasiones.

### 11.4.3. Nivel estilístico

El proceso de traducción conlleva la posible incorporación de **estructuras oscurecidas**, con la profusión de oraciones subordinadas y una estructura rebuscada: «Sin embargo, como empresa multinacional, los datos que recopilamos sobre usted pueden ser distribuidos internacionalmente a través de CSC, y la naturaleza abierta de Internet es tal que los datos pueden circular por redes que no disponen de medidas de seguridad, por lo que personas que no eran las destinatarias de dichos datos podrían tener acceso a éstos y utilizarlos» [CSC]. Estas estructuras serían contrarias a la necesaria claridad y brevedad recomendada en los manuales de estilo.

Desde nuestra perspectiva, se consideraría inadecuada la incorporación de oraciones cortas con un estilo más directo y breve propio de los textos anglosajones. La adopción de este estilo puede conllevar como resultado errores tales como la falta de modulación en la expresión del ruego. El siguiente ejemplo ilustra una posible inadecuación por el uso de un estilo directo en las secciones legales: «Si no está de acuerdo con los términos aquí descritos, **debería abandonar** el site ahora mismo.» [Avaya]. En lo que respecta a la brevedad oracional, ésta puede conllevar la inclusión de falsos sentidos ««Sun tiene el compromiso corporativo de alcanzar una excelencia reconocida en sus productos y servicios, que resulte en: **Empleados orgullosos, Clientes encantados, Inversores satisfechos**» [Sun]. Finalmente, este estilo puede resultar en contrasentidos: «Artículos prohibidos: No enviamos los siguientes artículos: Bebidas alcohólicas, **Pieles de animales (no domesticados)**» [UPS].

A pesar de que los textos web se caracterizan por una reducción en el nivel de formalidad, (Posteguillo 2003; Crystal 2001), existen errores pragmáticos relacionados con la falta de adecuación a un mayor nivel de formalidad propio de ciertas secciones en los sitios web originales. Éste menor grado de formalidad se refleja tanto en la selección léxica como en la sintaxis empleada: «Usted está sujeto

a los siguientes Términos y condiciones, **que podemos actualizar de vez en cuando** sin notificarle a usted de ello» [Hospira], «**Cualquier cosa** en este site tiene copyright» [Avaya]., o «Cualquiera que sea tu área de experiencia, **queremos gente** que:» [PriceWaterHouse].

#### 11.4.4. Errores ortotipográficos

Quizás los errores ortotipográficos sean los más recurrentes en este tipo de textos, y se centran especialmente en los errores ortográficos o cacografías (Martínez de Sousa 2007). Éstos podrían responder en ciertos casos a erratas o al uso de anglicismos ortotipográficos. Ambas inadecuaciones de expresión en la lengua de llegada tienen un impacto negativo en la apreciación global del sitio web por parte del usuario, y por extensión, la empresa en general (Jeney 2007: 46; Bolaños 2003). En general, apuntan a la presencia de deficiencias en la etapa de Control de Calidad, ya que estos errores ortográficos serían fácilmente evitables con los correctores ortográficos. A continuación presentamos un listado ilustrativo:

- Errores ortográficos o cacografías:
  - Acentuación: «Le agradecemos su interes en Colgate-Palmolive»
  - Ortografía errónea: «En caso de desear **entar en contacto** con nosotros y hasta poder establecer un servicio de enlace mediante correo electrónico», «Estamos comprometidos a la **protección** del medio ambiente» [Colgate Palmolive]

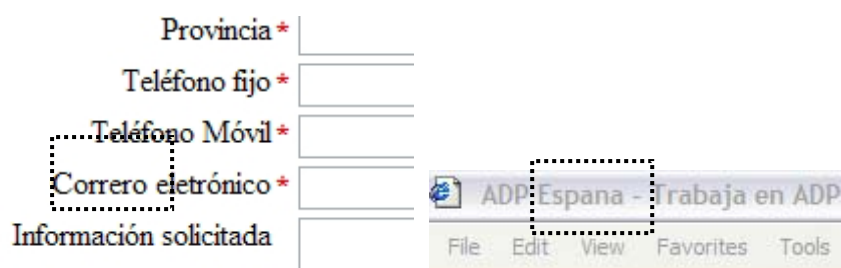


Figura 122. Ejemplos de errores ortográficos o cacografías, incluyendo la sustitución de la «ñ» por la letra «n».

- Anglicismos ortotipográficos (Martínez de Sousa 2007: 152-157):
  - Uso inadecuado de mayúsculas en sustantivos y adjetivos que forman parte de títulos de obras o partes de obras. Los textos muestran además inconsistencias en la inclusión de este anglicismo:

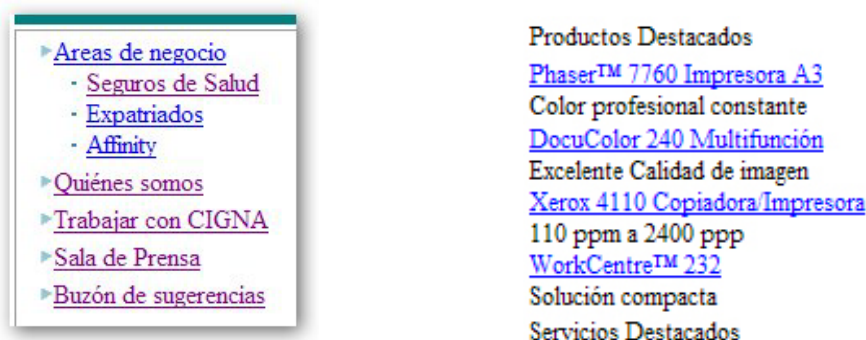


Figura 123. Inconsistencia en el uso inadecuado de mayúsculas en títulos en Cigna y Xerox.

- Uso de mayúsculas con los meses del año o idiomas, como «Por favor tenga en cuenta que todas las páginas de PeopleBank y formularios sólo están disponibles en **Inglés.**» [Selaed Air]. Los siguientes ejemplos en capturas de pantalla sobre el sobre el uso de mayúsculas para los meses del año:



Figura 124. Uso erróneo de mayúsculas en fechas en Electronic Arts.

- La escritura de las décadas y los años:

**1950's** Purina continua creciendo expandiéndose a los mercados de Europa y las Américas.

Figura 125. Anglicismo ortográfico en la descripción de las décadas y años en Purina.es

- La puntuación de las cantidades con el uso de coma para los miles y el punto en el decimal:

Auditoría	10.872	9.154	18.8%	14.6%
Consultoría	3.520	2.890	21.8%	16.6%
Legal y Fiscal	4.606	4.210	9.4%	5.5%
<b>Ingresos Netos por Servicios Profesionales Continuados</b>	<b>18.998</b>	<b>16.254</b>	<b>16.9%</b>	<b>12.6%</b>
Gastos Facturados a los Clientes	1.259	1.094	15.1%	8.6%

Figura 126. Anglicismo ortotipográfico y error en cifras con el uso de punto para marcar tanto los millares como los decimales.

- Formato de números - (Versión imprimible PDF - 2.18 Mb)
- Formato de cifras monetarias - «\$8.500 millones.» [ADP], «tiene una facturación de aproximadamente \$2.500 millones por sus ventas» [Hospira]
- Uso mayoritario de las comillas anglosajonas o el de las en lugar de las latinas, « ». Hemos identificado además el uso directo de las comillas simples en lugar de reservarlo para el último lugar (DPD, I edición 2005): «se proporcionan 'TAL CUAL' y están 'DISPONIBLES '» [Ford].
- En la grafía de las abreviaturas, escribir mayúscula en la abreviatura si le corresponde minúscula constituiría un anglicismo tipográfico: «En un vuelo típico, un 747-400 despeg a 290 **Km.** /h, vuela a una velocidad de 910 **Km.** /h y aterriza a 260 **Km.** /h» [Boeing].
- Inexistencia de los símbolos de interrogación o exclamación de apertura «Obtenlo ya!», «No se lo pierda!» [Sun].

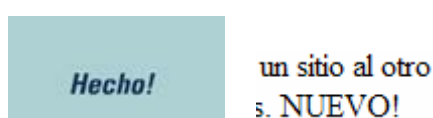


Figura 127. Error ortotipográfico en la falta de signo inicial de exclamación en sitios de Motorola y Mattel.

- Uso del signo & en lugar de la conjunción copulativa «y» - «Preguntas & Respuestas frecuentes» [Lexmark .es], «Educación & Investigación [...] Servicios & Soluciones» [Sun]. En los textos originales su uso se reduce a los nombre de empresas anglosajonas, como puede ser «Adquisición de la firma inglesa **Vogan&Company Ltd.**» [Arrocerías Herba], o «Newsletter "Accounting Roundup" publicada por **Deloitte & Touche** de Estados Unidos el 3 de febrero de 2003» [Deloitte & Touche].
- Finalmente, hemos de señalar las inconsistencias tipográficas, tal como el uso diferenciado de comillas:

Declaración de Privacidad "On Line", de HP, cualquier material, información o comunicación de otro tipo que el usuario transmita o publique en este sitio ("Comunicaciones") se considerará información no confidencial y no propietaria. Todo material, información o comunicación de otro tipo («Comunicaciones») [HP].

### **11.5. Reflexiones finales sobre los errores de traducción en textos localizados**

En el presente capítulo hemos presentado tan solo una selección ilustrativa de los errores de traducción relacionados con el incumplimiento de las convenciones referentes a las formas lingüísticas siguiendo el estudio de Gamero (2001), así como otros errores de expresión en la lengua de llegada. El alto porcentaje de errores identificados en la mayoría de sitios web podría apuntar a la menor importancia que se le otorga a los textos web en su conjunto, una circunstancia de interés ya que el corpus localizado está compuesto por los sitios web de empresas estadounidenses y, en principio, con mayores recursos disponibles para la localización. Esta menor importancia queda patente en la existencia de cláusulas en las secciones legales que eximirían de responsabilidad al emisor por cualquier error

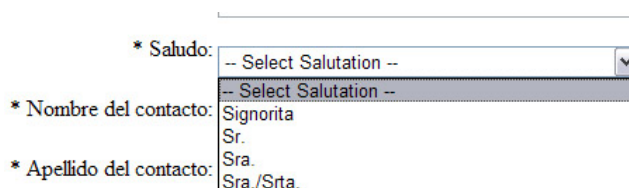
en el sitio web (F-3-I-b), presentes en el 27,36% de los sitios web localizados que contienen un bloque legal y sólo en el 11,04% de los originales. Es igualmente significativa la presencia de exenciones de responsabilidad por las traducciones de textos localizados, presentes en el 3,15% de los sitios localizados. En estas cláusulas se presentan los textos traducidos como un simple servicio al usuario: «se proporcionan para conveniencia del usuario Únicamente» [Agribrands] o «se facilitarán únicamente a efectos prácticos» [UPS]. Las traducciones se llegan a presentar sin garantía alguna de que sean fieles y precisas, deslegitimando directamente la labor de los traductores-localizadores, «Cargill no realiza declaración o manifestación alguna de que las traducciones sean fieles y precisas» [Cargill].

En nuestra opinión, la relación de mayor interés desde un punto de vista traductológico sería la correlación existente entre la traducción de bloques, secciones, unidades comunicativas de mayor frecuencia en los TO, reflejados en las diferencias superestructurales entre ambos conjuntos textuales, y la mayor presencia de errores de traducción en estos segmentos textuales en los que no existen formas lingüísticas convencionalizadas en la lengua de llegada. El caso de la sección «Política de privacidad» (F2) constituye uno de los mejores casos en este sentido. Esta unidad léxica se encuentra altamente convencionalizada en los textos anglosajones, como atestigua su presencia de esta sección en el 70,52% de textos localizados y sólo en el 13,37% de textos originales. En este caso, los errores identificados están relacionados con la falta de coherencia terminológica, llegando incluso a aparecer cinco formas diferentes en un mismo subtexto legal (véase 11.3.1).

En este sentido, el segmento textual que presenta la mayor diferencia porcentual en su frecuencia de uso entre ambos subcorpus dentro del bloque legal (F) lo constituye la subunidad significativa que apela a leer el texto (F-1-I-b) (véase

10.3.1). Los errores identificados en los textos localizados son principalmente pragmáticos (Hurtado Albir 2001), y se centran en la formulación de la exhortación con formas breves y concisas, así como el uso mayoritario de formas verbales en modo imperativo que resultar en una descortesía en el contexto español. Unos ejemplos ilustrativos serían « ¡MUY IMPORTANTE! LEA ESTE ACUERDO LEGAL ANTES DE VISITAR LA WEB» [Kimberly-Clark], «Antes de utilizar este sitio lea atentamente estos términos y condiciones.» [Genworth Financial]. En los textos originales esta subunidad significativa se suele expresar indirectamente por medio de formas como «el usuario ha de leer» o «el usuario deberá leer atentamente», o en caso de dirigirse directamente al usuario se presenta con una modulación en el ruego con verbos como «rogar», « le rogamos que lea con atención...» [Jungibérica].

Otro ejemplo significativo en este sentido sería la subunidad significativa «Tratamiento» (B-III-a-2) de presencia exclusiva en textos localizados. En la formulación de esta subunidad, el 50% de segmentos localizados incorporarían errores de traducción debidos a la falta de adaptación al contexto sociocultural española. Un caso significativo e ilustrativo lo constituiría esta misma subunidad en el sitio web de la empresa Symantec. Los errores identificados son múltiples, y se deben tanto al proceso de traducción como al proceso de localización e internacionalización. Los errores atribuibles a deficiencias en el proceso de localización serían tanto la permanencia de segmentos en inglés tales como «Select Salutation», como la incorporación de una palabra italiana «Signorita» atribuible a la integración de los segmentos localizados en la página web.



\* Saludo: -- Select Salutation --

\* Nombre del contacto: Signorita

\* Apellido del contacto: Sr.  
Sra.  
Sra./Srta.

Original en inglés de <http://www.symantec.com/es/es/partners/index.jsp>



\*Salutation: -- Select Salutation --  
\*Contact First Name: -- Select Salutation --  
\*Contact Last Name: Mr.  
Mrs.  
Middle Initial: Ms.

**Figura 128.** Subunidad significativa «tratamiento» (B-III-a-1) exclusiva de los textos localizados, con errores de traducción.

En lo que respecta a las inadecuaciones atribuibles al proceso de traducción-localización, el título de esta subunidad, «saludo», constituiría un barbarismo debido a un falso amigo, «salutation». Igualmente, la presencia de la forma singular de «apellido» constituiría una inadecuación pragmática a las normas socioculturales de la cultura de llegada: en el análisis de estas mismas subunidades significativas en textos originales todas las formas se encuentran en plural, «apellidos» (véase Tabla 27, p. 455). Finalmente, nos es de interés la traducción del tratamiento anglosajón «Ms.», que el inglés no denota estado civil alguno por «Sra./Srta.». Esta traducción en el contexto de un menú multiopción confundiría a los usuarios, ya que la opción de «señorita» que denota el estado civil de soltera, no aparecería por sí sola debido al error de localización que incluyó este segmento en italiano.

Este conjunto de circunstancias confirman nuestra posición sobre la relativa falta de atención a la calidad por parte de ciertas empresas a este proceso en general, así como deficiencias recurrentes durante la fase de control de calidad en los procesos de localización web (Bouffard y Craignon 2006; Corte 2002). Los errores que hemos identificado se corresponden con los parámetros por los que se establecería la calidad en la norma alemana sobre calidad de la traducción DIN 2345 (Bass 2006). Nuestra posición contrasta con las conclusiones de Bolaños *et al.* (2007: 141) que defienden la calidad en los sitios web originales: «La redacción y presentación de los textos [sitios web corporativos] es muy cuidada en las dos lenguas, con una ausencia casi absoluta de erratas y errores lingüísticos en general. Esto se debe a una estrategia de estas empresas para reforzar su credibilidad ante el cliente». Asumimos que nuestra perspectiva desde una metodología de la Lingüística de

corpus nos ha permitido identificar cuantitativamente esta mayor existencia de errores tanto en los textos localizados como en los originales. En este sentido y como reflexión final, creemos pertinente defender los beneficios que reportaría el uso de corpus de control y evaluativos amplios en las tareas de control de calidad en localización. Por consiguiente, en el futuro pondremos nuestro propio corpus convenientemente etiquetado a disposición de la comunidad científica y empresarial.



**PARTE IV:**  
**CONCLUSIONES**



## 12. CONCLUSIONES

Al inicio de la presente tesis nos planteamos el reto de investigar el proceso de localización web con miras a poner a disposición de traductores, docentes e investigadores un detallado análisis de las peculiaridades de este proceso mediante el estudio contrastivo de su producto, el texto localizado. Para conseguir este objetivo nos fue indispensable asentar nuestro estudio sobre unas bases teóricas firmes que validaran el análisis contrastivo, así como desarrollar un marco metodológico que tomara como base investigaciones previas, si bien adaptándolas para dar cabida a un nuevo modelo textual y comunicativo.

Si retomamos las motivaciones que originaron la presente tesis, podemos concluir que los resultados obtenidos podrán ser de utilidad como aportación al diccionario de géneros textuales propuesto por Hurtado Albir (2001) y Gamero (2001), así como a trabajos previos de carácter evaluativo y didáctico (Nord 1997b; 2003), al haber puesto de relieve cuantitativa y descriptivamente deficiencias y patrones de interferencia en los textos localizados de forma contrastiva con textos originales españoles. En este mismo sentido, nuestro trabajo también fue motivado por la línea de investigación más amplia sobre la traducción de géneros textuales propuesta por Hurtado Albir (2001: 505), y en este momento nos sentimos autorizados a concluir que los traductores de textos web no se ajustan necesariamente al conjunto de «convenciones estructurales y lingüísticas» establecidas en la lengua de llegada, el español en nuestro caso.

En primer lugar, y puesto que nuestra investigación se basa en textos producto de un proceso de localización, nos fue indispensable analizar cómo se desarrolla este proceso desde una perspectiva multidisciplinar. El ciclo de la localización posee unas características que pueden incidir directamente sobre cómo se desarrolla

proceso traductor y, consecuentemente, influir en su producto. El capítulo II lo dedicamos a revisar la escasa bibliografía en este ámbito con miras a dilucidar cómo se constituyen y desarrollan estos procesos de localización. En las publicaciones en este ámbito hemos constatado una defragmentación con múltiples perspectivas de estudio (Pym 2003), de ahí que nuestro objetivo se circunscribiera a aproximarnos a este fenómeno de forma multidisciplinar, tratando de tender puentes entre el sector de la Localización y la Traductología con miras a un beneficio mutuo (Jiménez 2008a, 2007a). Como resultado, hemos definido qué constituye el proceso de localización mediante un análisis de los aspectos que lo diferencian de un concepto amplio de traducción, incidiendo en las restricciones impuestas por el ciclo tecnológico en el que se enmarca y su dependencia del mismo (Lommel 2007). Tras nuestro análisis, hemos llegado a la conclusión de que las principales características que definirían esta nueva modalidad de traducción (Hurtado Albir 2001) son:

- (1) El medio en el que se sitúan los textos objeto de localización y su almacenamiento digital.
- (2) La disgregación del texto web como resultado del contexto tecnológico que lo organiza según criterios informáticos y no necesariamente comunicativos.
- (3) La existencia de dos estructuras superpuestas, la de presentación y la de organización o *back-end*, que suelen forzar al traductor a operar sobre una representación textual distinta al texto final que recibe el usuario. Este aspecto influye en el propio proceso de traducción.
- (4) La necesaria cooperación entre traductores y desarrolladores/especialistas en programación web.
- (5) Finalmente, hemos puesto de relieve las diferencias entre el proceso de localización web y localización de software, ya que este último se desarrolla más homogéneamente por su gran complejidad técnica (Esselink 2006), mientras que el

primero no se desarrolla de forma tan estructurada debido, en parte, a su menor complejidad tecnológica.

A partir de nuestro análisis del proceso de localización constatamos que el concepto de «texto», como unidad operativa de procesamiento durante la tarea de traducción, ha ido desapareciendo gradualmente (Bowker 2006; Mossop 2006; Biau y Pym 2006; Pym 2004a). Por consiguiente, el capítulo III lo dedicamos a revisar este concepto como entidad unitaria clave tanto en el proceso traductor como en nuestra investigación empírica. En este sentido, su importancia se debe a que nuestro objetivo consistió en contrastar unidades textuales lo más concretas y similares posibles, siendo necesaria su máxima comparabilidad (Bernardini y Baroni 2003). Desde una perspectiva traductológica, nuestra concepción de texto web la hemos identificado con el concepto de «hipertexto» propuesto por Storrer (2002), remarcando la naturaleza unitaria de cada sitio web como entidad en la que ha de recaer la coherencia y cohesión textuales (Jiménez 2007d). Estos parámetros de textualidad se sitúan como esenciales en el establecimiento de un texto en general (Beaugrande y Dressler 1981; Hatim y Mason 1990; Baker 1992), aunque en los textos web ocupa un lugar primordial la coherencia (Foltz 1997; Fritz 1999; Jiménez 2007d). Como resumen de nuestra visión teórica sobre el texto web hemos de señalar que:

- (1) Un sitio web constituye un texto unitario (Jiménez 2007b) con excepción de las «hiperwebs»: los géneros «portal» o «buscador» (Storrer 2002).
- (2) El carácter unitario de estos textos es una naturaleza inherente sobre la que recae la textualidad (Beaugrande y Dressler 1981: 3).
- (3) Un sitio web corporativo es un todo coherente (Foltz 1996; Fritz 1999) y cohesivo.



(4) El texto web engloba elementos visuales, sonoros, estructurales, tipográficos (Göpferich 1995), así como elementos dados por los estándares de base como las etiquetas y atributos html (véase 8.1.3.1).

(5) La estructura jerárquica de los textos web puede estar convencionalizada (Nielsen y Tahir 2002), es mutlilinear (Janoschka 2003), y depende de cada cultura (Neubert y Shreve 1992; Nord 1997b).

(6) Finalmente, los sitios web, desde una perspectiva funcionalista, constituirían mayoritariamente «textos complejos» (Reiss y Vermeer 1984), puesto que pueden incorporar ejemplares de otros géneros textuales.

En el capítulo IV realizamos una revisión crítica de las tipologías textuales sobre las que se han desarrollado investigaciones traductológicas de carácter contrastivo. En especial, nos centramos en la pertinencia de usar los conceptos de «tipo» o «género» textual como marco para nuestro análisis. En la presente tesis nos decantamos por el concepto de «género», que representa un concepto menos explorado en Traductología en comparación con el de «tipo textual» (López Rodríguez 2001). El género posee una larga tradición en los estudios literarios y surge por la capacidad de hablantes de una lengua de reconocer características y funciones de unos textos prototípicos que se repiten en una ocasión social determinada. Éstos suelen convencionalizarse con miras a mejorar su efectividad comunicativa (Hatim y Mason 1990).

Los géneros textuales combinan tres aspectos: los formales, los socioculturales y los cognitivos. En lo que respecta a los aspectos formales, los textos representan formas convencionalizadas que se expresan a través de movimientos o etapas (Swales 1990; Bhatia 1993; Nwogu 1997) reconocidas por los hablantes de una comunidad de discurso, y que pueden variar entre diferentes lenguas y culturas (Paltridge 1997; Gamero 2001). Debido a las peculiaridades de la textualidad web, estos movimientos los hemos asociado con las distintas páginas independientes que

---

componen un sitio web (Askehave y Nielsen 2005). Los aspectos socioculturales se centran en mejorar la efectividad comunicativa en una ocasión social determinada que se repite y, en este caso, se produce en un nuevo marco de comunicación que denominamos Modelo de Comunicación de Masas Interactivo (Janoschka 2003). Los géneros dependen de la cultura en la que se generan y sus convenciones pueden diferir entre culturas (Nord 1997b: 54), de ahí su importancia en la labor del traductor (íbid: 52; García Izquierdo 2002). Por último, los aspectos cognitivos se basan en los propósitos y expectativas de los participantes en un intercambio comunicativo. Nuestro análisis nos ha llevado a atribuir al género *sitio web corporativo* una función principal expositiva con una función secundaria exhortativa. Adicionalmente, esta vertiente cognitiva se centra en las expectativas de los usuarios que el emisor y traductor de un texto han de tener en cuenta y, en nuestro caso, nos es imprescindible añadir las relacionadas con la funcionalidad textual (Nielsen y Tahir 2002). Esta funcionalidad se ha definido por medio de los conceptos «interacción» e «interactividad» (Janoschka 2003), es decir, los procesos comunicativos que se entablan entre máquina y usuario o entre el emisor del texto y receptor. En nuestra delimitación de los rasgos característicos de todo género hemos adoptado una versión adaptada de la propuesta de caracterización de Silvia Gamero (1998, 2001), ya que ésta se fundamenta en géneros técnicos en los que podríamos situar la gran mayoría de subtextos integrantes de los sitios web corporativos.

La metodología de análisis que hemos adoptado quedó recogida en el capítulo IV y se fundamenta en la corriente traductológica que introduce la Lingüística de corpus para el estudio del producto y proceso de la traducción (Ej. Baker 1995; Laviosa 2002; Kenny 2001). En nuestro caso, el núcleo lo ha constituido un amplio corpus comparable de sitios web corporativos que recopilamos de forma sincrónica (Baker 1995). Este tipo de corpus, combinado con la base contrastiva que proporciona nuestro estudio descriptivo basado en el género textual y sus distintas

convenciones asociadas, nos ha permitido cuantificar las diferencias existentes a distintos niveles entre los sitios web originales y localizados. Para la presente tesis recopilamos un corpus comparable, con una sección representativa de la población textual que se define por el género web «sitio web corporativo» español, con una segunda sección representativa de este mismo género textual pero en su localización hacia el español desde una única lengua y contexto geográfico: el inglés estadounidense de las empresas cuya sede se encuentra en aquel país. A su vez, a partir de este corpus hemos elaborado varias versiones y subcorpus que responden a distintas necesidades metodológicas (véase Tabla 5, p. 282). Nuestro objetivo es poner este corpus a disposición de la comunidad científica, traductores profesionales que traten con este género textual, así como de profesores de traducción (Bowker 1998), ya que somos conscientes de que las posibles explotaciones de nuestro corpus van más allá de nuestra propia tesis. Adicionalmente, somos conscientes que los posibles análisis a partir de nuestro corpus están abiertos a distintas perspectivas y futuros trabajos. En especial, la sección original de nuestro corpus puede ser de interés con objetivos (1) evaluativos y de calidad (Bowker 2001) o (2) como corpus de control<sup>479</sup> en investigaciones monolingües o multilingües, tanto comparables como paralelas, y que abarquen distintas combinaciones lingüísticas.

Nuestro análisis de resultados ha girado en torno a dos ejes básicos, el descriptivo y el cuantitativo, siguiendo el *continuum* analítico que parte de parámetros extratextuales, pasando a aspectos intratextuales de mayor a menor orden jerárquico (Gamero 2001). Nuestro análisis extratextual en el capítulo VII ha puesto de relieve las diferencias existentes entre los ejemplares textuales recopilados en ambas secciones del corpus. Así, a pesar de estar dirigidos a una misma comunidad de discurso delimitada por el concepto tecnológico *locale*,

---

<sup>479</sup> No obstante, según Bowker un corpus de control ha de ser de escasas dimensiones: «a high-quality subcorpus consisting of texts that have been hand-picked primarily for their conceptual content. It is very small by corpus linguistics standards, containing only four or five texts with a total word count in the area of 5, 000 words» (2001: 351).

---

existen ciertas diferencias referentes al entorno sociocultural en el que se emiten los textos, sobre todo en su marco jurídico y socioprofesional. Las investigaciones sobre géneros textuales postulan la necesaria correspondencia en el conjunto de rasgos extratextuales al ser definitorios en sí del género (ibid), sobre todo por su influencia en la variación y convencionalización de sus elementos intratextuales (Göpferich 1995; Fernández Sánchez 2004).

La cuestión de mayor interés que surge en la conclusión de esta tesis sería, ¿deberíamos asumir la existencia de géneros textuales paralelos dependiendo de su carácter de texto original o traducido? Hasta este momento las investigaciones basadas en géneros textuales de corte traductológico se han basado principalmente en corpus paralelos con textos originales en ambos contextos socioculturales (Bolaños 2002a; Gamero 2001; Fernández Sánchez 2004; Vázquez del Árbol 2006; *inter alia*). Además, las perspectivas teóricas sobre el género postulan que las diferencias en rasgos extratextuales producen géneros diferenciados o subgéneros (Bhatia 1993), en parte por la gran influencia que éstos ejercen en la convencionalización de los elementos intratextuales en cada género (Gamero 1998). Quizás una aportación novedosa de nuestra tesis resida en la revisión del estatus de pertenencia a un mismo género de textos traducidos o localizados, sobre todo tras la identificación de ciertos rasgos prototípicos de los textos localizados que no comparten los textos originales. Este fenómeno de interferencia o asimilación en textos traducidos a las convenciones de género habituales en el contexto sociocultural de partida ha sido reconocido durante años como una característica de estos textos (Reiss y Vermeer 1984: 195; Toury 1995), aunque desde el punto de vista del género, los sitios web localizados en su conjunto comparten una serie de peculiaridades que por su alta frecuencia podríamos caracterizar como recurrentes. En especial, son de interés las diferencias relativas a la dependencia de la superestructura de origen, la presencia de ciertas convenciones lingüísticas específicas de su contexto sociocultural de partida o la

mayor variación denominativa y estilística que hemos justificado por el proceso de traducción en sí.

¿Sería por tanto pertinente establecer la existencia de dos géneros paralelos, el sitio web corporativo español y el sitio web corporativo localizado al español? ¿Es factible a partir de nuestros resultados ajustar los textos localizados a los textos originales de modo que sean mínimas las diferencias entre ambos tipos de textos y se diluyan las divergencias que los separan? A nuestro parecer, podemos ampliamente confirmar que nuestra tesis constituye un gran paso para que el sector de la localización consiga cumplir con su *desideratum* relativo a producir traducciones que se reciban como textos originales, de ahí que en caso de usar nuestros resultados con fines evaluativos hayamos situado este proceso dentro de los tipos de traducciones instrumentales de Nord (1997), comunicativas de Reiss y Vermeer (1984), o traducciones encubiertas de House (1977).

En los capítulos VIII y IX abordamos en primer lugar un análisis descriptivo del género con el fin de establecer qué características textuales, discursivas, pragmáticas o estructurales poseen estos textos web en España. A la vez, a lo largo de estos capítulos llevamos a cabo un análisis contrastivo por el que pusimos de relieve las posibles divergencias existentes entre ambas secciones del corpus. Puesto que sólo dentro de este nivel de análisis nos fue posible cuantificar estas posibles divergencias, retomaremos en este punto las hipótesis de partida para contrastarlas con los resultados obtenidos. La concepción de nuestra tesis estuvo determinada por el mencionado *desideratum* de producir textos similares a los originales de la región de destino, «with the ultimate aim of releasing a product that looks like it has been developed in country» (LISA 2004: 11). Este aspecto cristalizó en nuestra hipótesis de trabajo que volvemos a recoger:

- Las localizaciones son (1) el resultado de un proceso de producción textual sujeto a unas condiciones específicas y un contexto de producción diferenciado que (2) resultará en textos con ciertos aspectos cuantificables empíricamente por medio de la metodología de la Lingüística de corpus. (Baker 1995; 1996; 1998).

El primero de los puntos ha quedado ampliamente confirmado en las diferencias extratextuales entre ambos componentes del corpus (capítulo VII), especialmente a través de las diferencias referentes al contexto sociocultural y, por ende, el contexto socioprofesional de la producción textual. Estas diferencias en el contexto de producción entre redactores/desarrolladores originales y localizadores influyen en ciertas divergencias como: la mayor falta de homogeneidad terminológica, fraseológica, sintáctica y estilística. En lo que respecta a la segunda de las implicaciones de esta hipótesis de partida, tras nuestro análisis intratextual podemos concluir que aun perteneciendo al mismo género, los sitios web localizados presentan unas características diferenciadas que hemos ido revelando a lo largo de nuestro análisis de resultados y discusión. Recogemos a continuación las conclusiones a través de los siguientes puntos:

- a) Los textos localizados son el producto de un proceso de localización y poseen unas características diferenciadas y recurrentes que hemos analizado por medio de la metodología de la Lingüística de corpus. En este sentido, la aplicación de una base metodológica que parta del marco teórico que proporciona el concepto de «género» nos ha permitido realizar un estudio contrastivo más ajustado a las peculiaridades textuales de este tipo de textos. El género nos ha permitido disgregar y analizar contrastivamente aquellos subtextos y formas lingüísticas convencionales (Gamero 2001) que hemos seleccionado por su interés, tales como los títulos de las páginas web o el bloque comunicativo «contacto» (B). Por otro lado, esta decisión metodológica

nos ha permitido eliminar el ruido que se suele producir por la redundancia léxica típica de los sitios web, en parte debida a la repetición del texto «interfaz» en hipertextos (ej. menús de navegación) o la repetición de subtítulos de la estructura profunda sobre los que actúa el traductor, como los títulos de las páginas web (véase 9.2.1) o las descripciones en los atributos *Keywords* o *Description* (véase 9.2.3).

- b) Los perfiles textuales convencionales de ambos conjuntos textuales difieren considerablemente en su superestructura prototípica, el aspecto de mayor rango jerárquico en el análisis intratextual de géneros textuales (Göpferich 1995; Gamero 2001). Las diferencias superestructurales más significativas son las relativas a los subtítulos legales de distinto rango. Estas diferencias superestructurales han quedado patentes tanto en el análisis superestructural global que llevamos a cabo en el capítulo VIII, con una reveladora diferencia del 57,15% de aparición de la sección comunicativa «privacidad» (F2), como en el análisis expandido del bloque legal en el capítulo IX. En este último análisis ha quedado patente que las mayores diferencias se encuentran en aquellas unidades y subunidades significativas referentes a la mención específica del marco legal español (véase Tabla 42, p. 529). En este sentido, las subunidades representativas identificadas en este bloque con mayor índice de divergencia entre ambos conjuntos de subtítulos serían aquellas relativas a la realidad jurídica española, tales como la Ley Española de Protección de Datos (F-2-I-a) o la inscripción en el Registro Mercantil (F-1-III-a). Estas diferencias superestructurales se manifiestan además en los distintos niveles, tal como hemos constatado en el análisis adicional del bloque comunicativo «contacto» (B) en donde identificamos dos subunidades significativas exclusivas de los textos localizados, «iniciales» (B-III-A-1) y «tratamiento» (B-III-A-2). Estos resultados en su conjunto confirman ampliamente la influencia del contexto

sociocultural en la configuración del género y, consecuentemente, confirmarían a este nivel nuestra dos primeras hipótesis:

- Las restricciones que impone cada género textual y su dependencia de una cultura determinada.
- Nuestra segunda hipótesis relativa a las restricciones que impone el proceso de localización en la labor del traductor en la que la unidad textual operativa sobre la que trabaja el traductor no suele ser el texto completo y, por consiguiente, se reduce la posibilidad de que se realicen los cambios superestructurales necesarios.

En este sentido, el contexto tecnológico en el que se produce la traducción conlleva mayoritariamente textos «clonados» de los originales (Larose 1998: 172), en los que la estructura textual es completamente similar a la original<sup>480</sup>. La mayor longitud de los textos legales localizados,  $\pm 200\%$ , y la subsiguiente mayor frecuencia de la subunidad que apela a leer el texto legal (F-1-I-b) reflejan esta mera clonación de los subtextos legales localizados. Podríamos concluir que el proceso global de localización impediría en cierta manera el ejercicio traductor de experto en comunicación transcultural desde una perspectiva funcionalista (Reiss y Vermeer 1984; Holz-Mänttari 1984; Nord 1997b), en cuyo contexto el texto de llegada constituiría una oferta de información que se produce sobre la base que proporciona la oferta de información constituida por el texto de origen. Consecuentemente, en nuestra tesis hemos elaborado un marco que ayudaría a adaptar este tipo de textos al contexto sociocultural hacia el que se dirigen.

- c) Una de las conclusiones de mayor relevancia en nuestro análisis de los subtextos legales es la escasa presencia de traductores con especialización en

---

<sup>480</sup> Este hecho, el mantenimiento global de la superestructura sin cambios algunos, se repite en otros tipos de traducción como la del género «folleto turístico» (Nobs 2006).



traducción jurídica que podemos constatar tras nuestro análisis. En particular, hemos constatado la relativa falta de uso de la terminología convencionalizada en el contexto sociocultural meta, la fraseología propia de estos textos, o la presencia de gran cantidad de inadecuaciones o errores de traducción de todo tipo según la tipología adoptada (Hurtado Albir 2001). En este sentido, en estos textos existe un contrasentido entre el afán de presentarlos como textos originales (LISA 2003), y la aparición de cláusulas en las que se deslegitima la traducción en los subtextos legales al aducir que la única versión legalmente vinculante es la redactada en inglés o incluso la exclusión de responsabilidad por errores de traducción en el texto. Desde un punto de vista didáctico, esta conclusión posee serias implicaciones en la formación de los futuros localizadores, ya que como observamos en el reciente Congreso sobre la enseñanza en este ámbito en la Universidad Lessius en Amberes<sup>481</sup>, la gran mayoría de titulaciones en localización se decantan por una formación en las competencias técnicas, obviando hasta cierto punto la incorporación de una mayor variedad de tipos de traducción como la técnica, jurídica, etc. en estos programas de estudio.

- d) Las diferencias superestructurales se proyectan en la aparición de una mayor variación denominativa, fraseológica y estilística en los textos localizados que en los originales. Estas diferencias se muestran asimismo en los contenidos proposicionales y léxicos que hemos identificado por medio de los listados de palabras en ambas secciones del corpus. Los análisis contrastivos de las unidades léxicas empleadas en los menús de navegación para los bloques textuales «contacto», «empresa» y «privacidad» son muy clarificadores en este sentido. Si partimos del supuesto de que las unidades léxicas en estos subtextos se encuentran altamente convencionalizadas dentro del contexto anglosajón

---

<sup>481</sup> El Congreso «Teaching Localization for Global Business Readiness» organizado por la Universidad Lessius y la organización Localization Industry Standard Association del 10 al 12 de diciembre del 2007 en la que contribuimos con una conferencia (Jiménez 2007b).

(Nielsen y Tahir 2002), observamos que las posibles unidades localizadas muestran una pronunciada variación denominativa que no comparten los textos originales. En el caso del bloque «contacto», las unidades léxicas localizadas muestran un 51,11% de uso del pronombre «nosotros» o su forma enclítica «-nos», reflejo de la forma anglosajona «contact us», mientras que en las unidades léxicas originales su uso se reduce al 15,86%. Igualmente, el carácter verbal de esta forma anglosajona se refleja en un 71,11% de uso de formas verbales en el corpus localizado frente al 41,38% de los originales. En el bloque comunicativo «empresa», el dato más significativo sería el mayor porcentaje de uso de las preposiciones «acerca» y «sobre» en las unidades léxicas que reflejarían la convencional «about us»<sup>482</sup> anglosajona. En su conjunto, el 47,51% de unidades léxicas localizadas usarían estas preposiciones frente al 10,58% de originales. Desde un punto de vista cognitivo del proceso de traducción, sería de interés la amplia preferencia por la selección de la preposición «acerca de...» frente a «sobre» en los textos localizados, una tendencia que se invierte en los originales. Este hecho podría apuntar a la incidencia de la vertiente cognitiva en el proceso (Lörcher 1991; Danks y Shreve 1997; Tirkkonen-Conditt 2004) mediante un efecto de facilitación léxica<sup>483</sup> en la localización web. La atribución de diferencias entre textos originales y traducidos a factores cognitivos durante el proceso de traducción ha sido defendida por Tirkkonen-Conditt (2004) y Malmkjaer (2008), y nuestros resultados apoyarían el impacto de esta vertiente en el producto final. Por último, el caso de la sección de «privacidad» demuestra una relación directa entre la falta de convencionalización de una sección comunicativa en una cultura y la mayor dispersión y variación denominativa en el conjunto de elementos que componen las unidades léxicas localizadas.

---

<sup>482</sup> Con un 71% de uso en los textos originales anglosajones (Nielsen y Tahir 2002)

<sup>483</sup> Para un tratamiento más extenso del fenómeno de facilitación léxica o *lexical priming*, consúltese Brock (1986).

- e) Las diferencias entre ambos conjuntos textuales abarcan los niveles: léxico, morfosintáctico y estilístico. Queda patente en nuestro análisis que los textos localizados presentan hasta cierto punto la interferencia de los TO en inglés (Toury 1995; Laviosa 1996), sobre todo mediante la presencia de préstamos, patrones de co-ocurrencia léxica que reflejan su origen del inglés, frecuencias atípicas de elementos léxicos y sintácticos, o la presencia de un fenómeno de explicitación (Baker y Olohan 2000), como sería el caso de la elipsis del sujeto o *null subjects* que repercute en la mayor frecuencia de aparición explícita en los textos localizados de pronombre personales en función de sujeto.
- f) Hemos constatado que en los textos web se relajan hasta cierto punto las normas de uso de la lengua, así como la corrección lingüística y de estilo, un efecto que en Internet ha sido denominado *informality* (Posteguillo 2003). En concreto, en los textos originales son significativas las erratas ortotipográficas, especialmente en lo que respecta al uso de la tilde. Por su parte, los textos localizados presentan un conjunto de errores e inadecuaciones recurrentes que, dada la importancia económica de las empresas, apuntan a la pertinencia de futuros estudios sobre la fase de control de calidad web (Pierini 2007; Bouffard y Cragnon 2006). Si nos centramos en las inadecuaciones que atañen específicamente a los textos web, hemos de señalar aquellas que (1) se deben a deficiencias en la fase de localización-internacionalización, o (2) la falta de coherencia terminológica y sintáctica en los textos web. Estos errores se deben al entorno tecnológico basado en memorias de traducción y la desaparición de la figura del «texto» desde una perspectiva comunicativa y cognitiva en este proceso de traducción. Por otro lado hemos de señalar que (3) estos textos presentan asimismo errores debido a una fase posterior a la tarea de traducción en sí, el ensamblaje del texto final realizado por expertos técnicos y que hemos denominado «errores de localización», tales como la presencia errónea de otras lenguas (véase Figura 128), o errores sintácticos de los que el traductor en sí no

sería responsable, tales como la traducción de *go* por «ir a!» en el botón que apunta a la sección comunicativa «buscar» (H2), aunque sí que sería su responsabilidad en caso de realizar la fase de control de calidad.

- g) En cualquier caso, hemos de concluir que parte de los sitios web localizados de las grandes empresas presentan un nivel de calidad que nos sentimos autorizados a señalar sería inaceptable en textos impresos. Este hecho entronca con el concepto en textos web de «oralidad conceptual» (Janoschka 2003) o *informality* (Posteguillo 2003), en el que hasta cierto punto se relajan las normas de los textos en medio escrito.

En resumen, hemos identificado gran cantidad de divergencias en el análisis contrastivo intralingüístico que situarían nuestros resultados en el ámbito de investigaciones previas sobre corpus comparables en distintas combinaciones lingüísticas (Olohan y Baker 2004; Tirkkonen-Conditt 2004; Mauranen 2000; Kenny 2001; Laviosa 1996; *inter alia*). No obstante, hemos de apuntar que la gran mayoría de estas investigaciones se basan en textos periodísticos o literarios, de ahí que podamos añadir al cuerpo de conocimiento sobre este tipo de investigaciones en una modalidad diferente. En este sentido, nuestros resultados contrastan con investigaciones previas que presentan como «universal de traducción»<sup>484</sup> a la controvertida «convencionalización» (Kenny 1999; Toury 1995: 268), una posible tendencia a que los textos traducidos sean más convencionales que los originales. Los resultados obtenidos demuestran que los textos localizados presentan una mayor variación y unos menores niveles de convencionalización en general. Finalmente, hemos identificado ciertas diferencias que hemos de atribuir a peculiaridades del proceso de localización, tales como la falta de homogeneidad terminológica y estilística que resulta de la combinación de segmentos textuales de

<sup>484</sup> Consúltese Mauranen (2008: 40-41) para una revisión de las investigaciones que defienden o desestiman esta controvertida tendencia.

diverso tamaño traducidos por más de un traductor, y que se conjugan en el texto final sin una revisión global de calidad.

Nuestra investigación ha pretendido contribuir a elaborar una base teórica y metodológica para futuras investigaciones sobre textos web, sobre todo por que esta incipiente disciplina está siendo investigada con una carencia de bases sólidas y coherentes (Foralón 2006), que hasta cierto punto impide su progreso e intercambio de resultados entre las distintas perspectivas de estudio (Dunne 2006; Pym 2003). En nuestra tesis hemos pretendido reivindicar la importancia de conceptos clave en el desarrollo de la Traductología como el «texto», «las tipologías textuales», «el género» u otros fruto de la evolución dentro de este ámbito, como «el corpus web» o «la superestructura hipertextual». Nuestros resultados han sido y serán utilizados en la docencia en Localización, en el desarrollo de guías de traducción web, en sistemas de evaluación de calidad, y mejoras del proceso traductor por parte del sector de la localización.

Como conclusión, con nuestro estudio empírico hemos pretendido aportar información sobre las incidencias del proceso de localización en los textos traducidos que esperamos ayude a producir textos web plenamente funcionales y coherentes. Esperamos que nuestro estudio sea de utilidad a la comunidad docente y científica, puesto que podemos afirmar que el desarrollo de este proceso se multiplicará inexorablemente en los años venideros (LISA 2004; Schäler 2005). Asimismo, y una vez demostrado que no es factible la quimera de producir textos «similares» a los originalmente redactados en un contexto sociocultural específico sobre todo por restricciones impuestas por el entorno tecnológico, sociocultural, cognitivo y económico del proceso (Baker 1999: 285)<sup>485</sup>, esperamos que el sector de

---

<sup>485</sup> Mona Baker señala las restricciones sociales, culturales, ideológicas y cognitivas como las principales en el proceso traductor (1999: 285). No obstante, debido al entorno tecnológico en el que se desarrolla la gran mayoría de la traducción profesional en el siglo XXI (Mossop 2006), hemos de reclamar la importancia de las restricciones «técnicas» o «tecnológicas» que hemos demostrado ejercen una influencia determinante en este proceso.

la localización se concierne plenamente sobre la necesaria atención al «texto», al «género textual», y que sitúe las vertientes textual, comunicativa y cognitiva en este amplio ciclo al mismo nivel que las consideraciones tecnológicas.

### **12.1. Futuras vías de investigación**

Como colofón de la presente tesis nos es imprescindible proyectar nuestra investigación hacia el futuro exponiendo las posibles vías de investigación.

(1) En un primer lugar, hemos de señalar que el marco teórico y metodológico desarrollado nos permitiría llevar a cabo investigaciones contrastivas o descriptivas sobre otros géneros web de importancia en la práctica de la traducción. De entre ellos, y puesto que el autor de la presente tesis desarrolla su labor profesional en centros universitarios estadounidenses, el siguiente género de interés para un estudio descriptivo-contrastivo serían los géneros web institucionales, los de mayor importancia específica en este contexto sociocultural dado su gran volumen de traducción para la población hispanohablante. En la actualidad, la gran mayoría de sitios institucionales suelen localizarse a una variedad en principio neutra del español por un conjunto de traductores cuyo perfil socioprofesional no se encuentra tan definido como en otras áreas geográficas. La combinación de estos dos hechos hace que este fenómeno reclame su estudio pormenorizado desde una perspectiva traductológica.

(2) Otra posible vía está relacionada con la variación dialectal en la WWW (Bouffard y Craignon 2006) y el desarrollo en esta área geográfica de una variedad denominada español internacional o neutro (Castro 1996; García Izquierdo 2006). Su desarrollo está íntimamente ligado a una ingente cantidad de procesos de traducción desde el inglés estadounidense hacia un español «neutro», y realizados por traductores con diverso grado de profesionalización y una variedad dialectal

propia dependiente de su región de procedencia. En este mismo sentido, durante el proceso de defensa de nuestro trabajo de investigación tutelada se nos animó a estudiar las diferencias en la influencia del proceso de localización entre una región lingüística más protectora de su variedad dialectal, como la española, con otras más en contacto directo con el ámbito estadounidense, como México. Tal investigación sería de gran interés ya que partiríamos de la hipótesis de que la proximidad física entre dos regiones o *locales* puede afectar al grado de aceptación de traducciones extranjerizantes o la mayor adopción de préstamos (Bouffard y Craignon 2006).

(3) En un tercer lugar, la solidez de nuestros resultados y la posible adscripción de varias tendencias identificadas a ciertos «universales de traducción» (Baker 1993, 1996) o «leyes de traducción» (Toury 1995, 2004), tales como la explicitación, la simplificación o la interferencia (Mauranen 2008), se afianzaría si comparásemos los resultados obtenidos en la dirección inversa, en traducciones que partan desde el español hacia el inglés. Este estudio sería factible con nuestro corpus original sin procesar del que separamos aproximadamente 9000 páginas web localizadas al inglés a partir de textos originales españoles. Con este material hemos elaborado un corpus paralelo no estudiado aún y que, tras su análisis preliminar, podemos adelantar que existe un proceso de asimilación similar a las convenciones de género español en la dirección inversa, en los textos localizados al inglés. Sirva como ejemplo ilustrativo una unidad léxica identificada que sólo aparece en menús de navegación de sitios web localizados desde español peninsular hacia el inglés, «legal warning», un calco directo de la unidad léxica «aviso legal» convencionalizada en España. Igualmente, otro ejemplo sería «who are we?» en el bloque comunicativo «empresa» (C), calco directo de «¿Quiénes somos?» (16,8%). Estos ejemplos ilustrativos mostrarían la relativa falta de conocimiento de las convenciones de género establecidas en los sitios web anglosajones por parte de los traductores de estos textos. De igual modo, sería de interés en esta futura vía de

---

investigación el análisis del contexto laboral en el que se desarrollan en España las traducciones hacia el inglés.

(4) Finalmente, nos es imprescindible abordar el posible uso didáctico de los resultados obtenidos en nuestra tesis. En este sentido, partimos de la base de que la traducción web debe formar parte de la formación integral de todo futuro traductor (Jiménez 2008b; Prieto, Tercedor y López 2007; Bolaños *et al.* 2005; Días Fouçes 2004; Gouadec 2003), y por extensión, los sitios corporativos representan el género web traducido por antonomasia en cursos de localización. De ahí que los resultados obtenidos en la presente tesis supongan un compendio de las mayores dificultades e interferencias que entraña este tipo de traducción. En especial, nuestros resultados demuestran la necesaria sensibilización de los futuros traductores hacia la integridad del texto global, el sitio web en su conjunto, como entidad textual mínima sobre la que recaiga la cohesión y coherencia global (Jiménez 2007d). En particular, las investigaciones empíricas sobre el proceso de traducción han puesto de manifiesto la tendencia de los estudiantes de traducción a emplear menos estrategias globales y traducir en el microcontexto que los expertos (Jääskeläinen y Tirkonen-Condit 1991), toda vez que el indispensable uso de herramientas informáticas en localización conlleva un procesamiento que fuerza a procesar el texto al nivel oracional (Shreve 2006; Dragsted 2002; Webb 2000). Estos dos aspectos apoyarían nuestra defensa de usar textos completos, el sitio web en su conjunto (Jiménez 2008b), así como la sensibilización de los futuros traductores hacia una perspectiva global que enmarque el conjunto del contexto textual (Hatim y Mason 1990), en el que ocupan un lugar primordial los aspectos textuales y convenciones dadas por el género textual (Nord 1997b; Gamero 2001). Finalmente, nuestra tesis demuestra que los géneros digitales constituyen géneros o textos «complejos» (Martin 1995; Hanks 1996; Reiss y Vermeer 1984), puesto que pueden incluir en su interior otros géneros textuales. Este hecho dificulta su traducción e implica que la competencia textual/genérica requerida para su



traducción ha de ser amplia (Hurtado Albir 2001: 477). Como resultado, la formación en localización web debe incorporar, además de la competencia tecnológica, los distintos tipos de traducción relacionados con los distintos bloques comunicativos que hemos identificado: traducción jurídica, publicitaria, económica, técnica, periodística. De hecho, los resultados del capítulo 10 demuestran la falta de competencia en traducción jurídica de un alto número de los traductores-localizadores encargados de traducir los bloques legales (F).

En otra de sus posibles vertientes, la explotación de nuestro corpus textual con objetivos didácticos se centraría en el uso en la docencia del perfil prototípico del género que hemos obtenido, así como las distintas convenciones textuales asociadas tanto estructurales, fraseológicas o terminológicas. Nuestros resultados facilitarían el proceso de selección textual en las clases de localización (Jiménez 2008; Reinke 2005), así como la selección de los problemas fraseológicos, terminológicos o pragmáticos de mayor dificultad con los que familiarizar a los estudiantes. Finalmente, sería posible la explotación de nuestro corpus, que pondremos a disposición de la comunidad científica, como corpus de referencia o evaluativo en la docencia en traducción (Zanettin *et al.* 2003; Zanettin 1998, 2000; Varantola 2002; Bowker 2001). En especial, el corpus podría disgregarse y, consecuentemente, sería posible usar la sección original del corpus a modo de corpus monolingüe especializado de referencia en las clases de localización (Bowker 1998).

(5) Aún dentro de las posibles aplicaciones a la docencia de la localización, quizás la aportación de mayor calado sería la futura elaboración de una guía de estilo de la localización web en la que se recogerían, desde una perspectiva descriptiva y no necesariamente prescriptiva, las peculiaridades de los textos web en España. Esta futura guía web complementaría el proyecto de guía de estilo web en curso de la

empresa Microsoft (Karsch 2007)<sup>486</sup>, ya que la perspectiva tras este guía es de arriba-abajo o *top-down* (Schäler 2005), partiendo desde los textos originales en inglés, mientras creemos haber dejado patente que nuestra perspectiva de estudio, abajo-arriba, es más productiva.

Para terminar, nos gustaría retomar la cita inicial de nuestra tesis. Nos volvemos a hacer eco de las palabras del experto en usabilidad Jakob Nielsen con las que reclama la importancia vital del «texto» en la WWW: «**plain text** is the foundation of most web information» (Nielsen y Tahir, 2002: 48). Nuestra tesis confirma que, a pesar de los avances tecnológicos desde el advenimiento de Internet y las dificultades técnicas presentes en localización, nos hemos de desestimar sino más bien atender en igualdad de condiciones a los aspectos textuales, comunicativos y cognitivos que forman parte intrínseca de todo proceso de traducción.

---

<sup>486</sup> Terminóloga asociada a la empresa Microsoft. Comunicación personal el 11 de diciembre del 2007 tras la conferencia del doctorando en la conferencia «Teaching Localization for Global Business Readiness» organizada por la Localization Industry Standard Association en Amberes, Bélgica.



**PARTE V:**  
**BIBLIOGRAFÍA**  
**Y**  
**APÉNDICES**



---

## 13. BIBLIOGRAFÍA

- Adam, J. M. 1985. «Cadre théorique d'une typologie séquentielle». *Études de Linguistique Appliquée*, 83: 7-18.
- Adkisson, H. 2002. *Identifying de-facto standards for E-commerce websites*. Tesis de Master. Universidad de Washington, EEUU.
- Agost, R. 1996. *La traducció audiovisual: el doblatge*. Tesis doctoral. Universitat Jaume I.
- Agost, R. 1999. *Traducción y doblaje: palabras, voces e imágenes*. Barcelona: Ariel.
- Alamán, T. 2003. «The banner ad as a netvertising subgenre: an analysis of form and content». En S. Posteguillo, E. Ortells, J. R. Prado, A. Bolaños y A. Alcina (eds.). *Internet in linguistics, translation and literary studies*. Castellón de la Plana: Publicaciones de la Universidad Jaume I. 375-390.
- Alejo González, R. 2005. «Textual metadiscourse in commercial websites». *Ibérica*, 9: 33-52.
- Altanero, T. 2006. «The localization job market in academe». En A. Pym, A. Perekstenko y B. Starink (eds.). *Translation Technology and its Teaching*. Tarragona: Intercultural Studies Group. 31-36.
- Anderson, D., B. Benjamin y B. Paredes-Holt. 1998. *Connections: a guide to on-line writing*. Boston: Allyn and Bacon.
- Apple Computer Inc. 1992. *Macintosh Human Interface Guidelines*. Boston: Addison-Wesley Professionals.
- Archibald, J. (ed.). 2004. *La localisation: problématique de la formation*. Montreal: Linguatéc Editeur.
- Arntz, R. 1990. «Überlegungen zur Methodik einer "Kontrastive Textologie"». En R. Arntz y G. Thorne, *Übersetzungswissenschaft Ergebnisse und Perspektiven*. Tübinga: Narr. 393-404.
- Arntz, R. 1988. «Steps towards a translational oriented tipology of technical texts». *Meta*, 33: 468-471.
- Askehave, I. y A. E. Nielsen. 2005. «Digital genres: a challenge to traditional genre theory». *Information Technology and People*, 18 (2): 120-141.

- Atkins, J. S. Clear, y N. Olster. 1992. «Corpus Design Criteria». *Literary and Linguistic Computing*, 7 (1): 1-16.
- Austermühl, F. 2006. «Training Translators to Localize». En A. Pym, A. Perekstenko y B. Starink (eds.). *Translation Technology and its Teaching*. Tarragona: Intercultural Studies Group. 69-81.
- Austin, J. L. 1975. *How to do things with words*. Oxford: Clarendon.
- Baker, M. 2004. «A corpus-based view of similarity and difference in translation». *International Journal of Corpus Linguistics*, 9 (2): 167-193.
- Baker, M., 1999. «The Role of Corpora in Investigating the Linguistic Behaviour of Professional Translators». *International Journal of Corpus Linguistics*, 4 (2): 281-298.
- Baker, M. 1996. «Corpus-based Translation Studies: The Challenges that Lie Ahead». En H. Somers (ed.). *Terminology, LSP and Translation: Studies in Language Engineering in Honour of Juan C. Sager*. Amsterdam - Philadelphia: John Benjamins.
- Baker, M. 1995. «Corpora in Translation Studies: An overview and some suggestions for future research». *Target*, 7 (2): 223-243.
- Baker, M. 1993. «Corpus Linguistics and Translation Studies: Implications and Applications», in M. Baker, G. Francis y E. Tognini-Bonelli (eds.). *Text and Technology: In Honour of John Sinclair*. Amsterdam-Filadelfia: John Benjamins.
- Baker, M. 1992. *In Other Words: A coursebook on translation*. Londres/Nueva York: Routledge.
- Baroni, M. y S. Bernardini. 2003. «A preliminary analysis of collocational differences in monolingual comparable corpora». *Actas de la conferencia Corpus Linguistics 2003. UCREL technical paper 16*. Lancaster: UCREL Lancaster University. 82-91.
- Barthes, R. 1977. *Image – Music – Text*. Nueva York: Hilla and Wang.
- Bassnett, S. 1991. *Translation Studies*. London/New York: Routledge.
- Bassnett, S. y A. Lefevere. 1990. *Translation, history and culture*. Londres-Nueva York: Pinter Publishers.
- Bass, S. 2006. «Quality in the Real World». En K. Dunne (ed.). *Perspectives on Localization*. Amsterdam-Filadelfia: John Benjamins. 69-84.

- 
- Bazerman, C. 1994. «Systems of genres and the enactment of social intentions». En A. Freedman y P. Medway (eds.). *Genre and the new rhetoric*. London: Taylor and Francis. 79-101.
- Beaugrande, R. de. 1978. *Factors in a theory of poetic translating*. Amsterdam: Rodopi.
- Beaugrande, R. de y W. Dressler. 1981. *Introduction to Text Linguistics*. Londres: Longman.
- Bédard, C. 2000. «Mémoire de traduction cherche traducteur de phrases». *Traduire*, 186: 41-49.
- Belda Medina, J. R. 2003. *El lenguaje de la informática e Internet y su traducción*. Alicante: Universidad de Alicante.
- Bell, R. T. 1991. *Translation and Translating: Theory and practice*. Londres: Longman.
- Berkenkotter, C. y T. Huckin. 1995. *Genre knowledge in disciplinary communications. Cognition, culture, power*. New Jersey: Lawrence Erlbaum.
- Bermúdez Bausela, M. 2005. «Teaching Localization in Spanish Universities». *Localization Focus*, septiembre 2005.
- Bernal Merino, M. 2006. «The Translation of Video Games». *Jostrans, Journal of Specialized Translation* 6.
- Bernárdez, E. 1982. *Introducción a la Lingüística del texto*. Madrid: Espasa Calpe.
- Berners-Lee, T. 1999. *Weaving the Web*. San Francisco: Harper Collins.
- Bhatia, V. K. 2002. «Applied genre analysis: a multi-perspective model». *Ibérica*, 4: 3-19.
- Bhatia, V. K. 1997. «Genre-mixing in academic introductions». *English for Specific Purposes*, 16: 181-195.
- Bhatia, V. K. 1993. *Analysing genre. Language use in professional settings*. Londres: Longman.
- Biau Gil, J. R. y A. Pym. 2006. «Technology and translation (a pedagogical overview)». En A. Pym, A. Perekstenko y B. Starink (eds.). *Translation Technology and its Teaching*. Tarragona: Intercultural Studies Group. 5-20.
- Biber, D. 1988. *Variations across speech and writing*. Cambridge: Cambridge University Press.



- Biber, D. 1989. «A typology of English texts». *Linguistics*, 27: 3-43.
- Biber, D. 1993. «Representativeness in corpus design». *Literary and Linguistic Computing*, 8 (4): 243-257.
- Biber, D. 1995. *Dimensions of register variation*. Cambridge: Cambridge University Press.
- Biber, D., S. Conrad y R. Reppen. 1998. *Corpus Linguistics. Investigating language, structure and use*. Cambridge: Cambridge University Press.
- Biber, D. 2004. «Towards a typology of web registers: A multi-dimensional analysis». Ponencia *Corpus Linguistics: Perspectives for the Future*. Disponible en Web: <http://jan.ucc.nau.edu/~biber/Web%20text%20types.ppt>>).
- Bloomfield, L. 1935. *Language*. London: George Allen and Unwin.
- Bly, R. 2002. *The Online Copywriter's Handbook*. Nueva York: McGraw-Hill.
- Bolaños Medina, A., M. J. Rodríguez Medina y L. Losada. 2007. *Estudios sobre el discurso electrónico y sus aplicaciones para la traducción técnica*. Valencia: Aduana Vieja.
- Bolaños Medina, A. 2003. «Dificultades y estrategias de la localización de sitios web comerciales del inglés al español». En S. Posteguillo *et al.* (eds.). *Internet in linguistics, translation and literary studies*. Castellón: Universitat Jaume I. 241-252
- Bolaños, A. 2002. *Diseño y aplicación de un model didáctico innovador para la traducción de géneros digitales*. Tesis Doctoral. Universidad de las Palmas de Gran Canaria.
- Borja Albí, A. 2008. «Corpora for Translators in Spain». En G. Anderman y M. Rogers (eds.). *Incorporating Corpora: The Linguist and the Translator*. Clevedon: Multilingual Matters. 243-265.
- Borja Albí, A. 2005. «Organización del conocimiento para la traducción jurídica a través de sistemas expertos basados en el concepto de género textual». En I. García Izquierdo (ed.). *El género textual y la traducción. Reflexiones teóricas y aplicaciones pedagógicas*. Berna: Peter Lang.
- Borja Albi, A. 2003. «La investigación en traducción jurídica», en M. García Peinado y E. Ortega Arjonilla (eds.), *Panorama actual de la investigación en traducción e interpretación*. Granada: Atrio.
- Borja Albi, A. 2000. *El texto jurídico inglés y su traducción al español*. Barcelona: Ariel.

- 
- Borja Albí, A. 1998. *La traducción jurídica. Un enfoque discursivo*. Tesis doctoral sin publicar. Universidad Autònoma de Barcelona.
- Bouffard, P. y P. Caignon. 2006. «Localization et variation linguistique. Vers une geolinguistique de l'espace virtuel francophone». *Meta* 51 (4): 806-823
- Bowker, L. 2006. «Translation memory and "text"». En L. Bowker (ed.). *Lexicography, terminology and translation*. Ottawa: University of Ottawa Press.
- Bowker, L. 2001. «Towards a Methodology for a Corpus-Based Approach to Translation Evaluation». *Meta* 46 (2): 345-364.
- Bowker, L. 1998. «Using Specialised Monolingual Native-Language Corpora as a Translation Resource: A Pilot Study». *Meta* 43 (4): 331-351.
- Brinck, T. D. Gergle y S. D. Wood. 2002. *Usability for the Web*. San Francisco: Morgan Kauffman.
- Bock, J. K. 1986. «Meaning, sound, and syntax: Lexical priming in sentence production». *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12: 575-586.
- Bronckart, J. P. 1985. *Le fonctionnement des discours*. Neuchatel: Delachaux-Niestlé Editeurs.
- Brooks, D. 2000. «What Price Globalization? Managing Costs at Microsoft». En R. C. Sprung (ed.). *Translating into Success. Cutting-edge strategies for going multilingual in a global age*. Amsterdam-Filadelfia: John Benjamins. 42-59.
- Bühler, K. 1934. *Sprachtheorie*. Jena: Fischer.
- Bustos Gisbert, J. M. 1996. *La construcción de textos en español*. Salamanca: Ediciones Universidad de Salamanca.
- Bustos Gisbert, J. M., M. C. Cortés Zaborras y M. J. Hernández Guerrero. 2005. *La traducción periodística*. Cuenca: Universidad de Castilla-La Mancha.
- Byrne, J. 2006. *Technical Translation. Usability Strategies for Translating Technical Documentation*. Dordrecht, Holanda: Springer.
- Cabanillas, I. C. et al. 2007. «English loanwords in Spanish computer language». *English for Specific Purposes*, 26: 52-78.
- Cadieux, P. y Esselink, B. 2002. «GILT: Gobalization, Internationalization, localization, translation». *LISA Newsletter. Globalization Insider* 11 (1.5), 22 de marzo. Disponible en Web:  
<[http://www.lisa.org/archive\\_domain/newsletters/2002/1.5/index.html](http://www.lisa.org/archive_domain/newsletters/2002/1.5/index.html)>

- Castellà, J. M. 1992. *De la frase al text. Teories de l'ús lingüístic*. Barcelona: Empúries.
- Castro, X. 1996. «El español neutro». Presentación en el *Congreso Anual de la ATA* (Asociación Estadounidense de Traductores), Colorado Springs, 5 de noviembre. Disponible en Web: <[http://ourwold.compuserve.com/homepages/xose\\_castro/neutro.htm](http://ourwold.compuserve.com/homepages/xose_castro/neutro.htm)>.
- Chandler, H. M. 2004. *The game localization handbook*. Hingham, MA: Charles River Media.
- Chaume, M. J. 2004. *Cine y Traducción*. Madrid: Cátedra.
- Cheng, S. 2000. «Globalizing an e-Commerce Website». En R. C. Sprung (ed.), *Translating into Success. Cutting-edge strategies for going multilingual in a global age*. Amsterdam-Filadelfia: John Benjamins: 29-41.
- Chesterman, A. 1993. «From 'is' to 'ought': Laws, norms and strategies in translation studies». *Target* 5: 1-20
- Chomsky, N. 1965. *Aspects of the Theory of Syntax*. Cambridge: MIT Press.
- Codina, L. 2003. «Hiperdocumentos: composición, estructura y evaluación». En J. Díaz Noci y R. Salaverría Aliaga (coords.). *Manual de redacción ciberperiodística*. Barcelona: Ariel Comunicación. 141-194.
- Conklin, J. 1987. «Hypertext: An introduction and survey». *Computer* 20: 17-41.
- Connor, U. 1996. *Contrastive rhetoric*. Cambridge: Cambridge University Press.
- Corpas Pastor, G. 2001. «Compilación de un corpus ad hoc para la enseñanza de la traducción inversa especializada». *Trans* 5: 155-184.
- Corte, Noelia, 2000. *Web Site localization and Internationalization. A case study*. Tesis de Master (MA Thesis). City College, Reino Unido.
- Crowston, K. y M. Williams. 1999. «The effect of Linking on Genres on Web Documents». *Actas del la XXXIII Annual Hawaii International Conference on System Sciences*, Kilea, Hawaii. Los Alamitos, CA: IEEE-Computer Society.
- Crowston, K. y M. Williams. 1997. «Reproduced and emergent genres of communication on the World Wide Web». *Actas del la XXX Annual Hawaii International Conference on System Sciences*, Maui, Hawaii. Los Alamitos, CA: IEEE-Computer Society. 30-39.
- Cronin, M. 2003. *Translation and Globalization*. Londres: Routledge
- Crystal, D. 2001. *Language and the Internet*. Cambridge: Cambridge University Press.

- Crystal, D. y D. Davy. 1969. *Investigating English Style*. Londres: Longman.
- Delisle, J. 1980. *L'Analyse du discours comme méthode de traduction*. Cahiers de Traductologie, 2. Ottawa: Les Presses. Editions de l'Université d'Ottawa.
- DePalma, D. 2003. «Rage against the content machine» [en línea]. *The Multilingual Digital World, 2003, Conference del Localization Research Centre*, Limerick, Irlanda. Disponible en Web: <http://www.localization.ie/publications/presentations/2003/Conference/Presentations/DePalma%20RLC.ppt>
- Díaz Cintas, J. 2003. *Teoría y práctica de la subtitulación inglés-español*. Barcelona: Ariel.
- Días Fouces, O. 2004. «A localização de páginas da internet na formação de tradutores». En *Confluências, Revista de Tradução Científica e Técnica*, 1: 16-52. Disponible en Web: <http://www.confluencias.net/n1/fouces/html>.
- Díaz Noci, J. y R. Salaverría Aliaga (eds.). 2003. *Manual de redacción ciberperiodística*. Barcelona: Ariel Comunicación.
- Dietz, F. 2006. «Issues in Localizing computer games». En K. Dunne (ed.). *Perspectives on Localization*. Amsterdam-Filadelfia: John Benjamins. 122-134.
- Dillon, A. y B. A. Gushrowski 2000. «Genres and the Web: is the personal home page the first uniquely digital genre? ». *Journal of the American Society for Information Science* 51 (2): 202-205.
- Downing, A. 1996. «Register and/or genre?». En I. Vázquez y A. Hornero (eds.). *Current Issues in Genre Theory*. Zaragoza: Mira Editores. 11-27.
- Duff, D. 2000. *Modern Genre Theory*. Harlow: Pearson.
- Duff, A. 1981. *The third language: recurrent problems of translation into English*. Oxford: Pergamon Press.
- Dudley-Evans, T. y M. J. St. John. 1998. *Developments in English for specific purposes: A multi-disciplinary approach*. Cambridge: Cambridge University Press.
- Dunne, K. (ed.). 2006. *Perspectives on Localization*. Amsterdam-Filadelfia: John Benjamins.
- Dunne, K. 2006. «A Copernican Revolution». En K. Dunne (ed.). *Perspectives on Localization*. Amsterdam-Filadelfia: John Benjamins. 1-11.

- Duque García, M. 2000. *Manual de estilo. El arte de escribir en inglés científico-técnico*. Madrid: Paraninfo.
- Dr. International (Microsoft Corporation). 2003. *Developing International Software*. Redmond: Microsoft Corporation.
- Engberg, J. 1997. *Konventionen von Fachtextsorten. Kontrastive Analysen zu deutschen und dänischen Gerichtsurteilen*. Tübinga: Narr.
- Engelbrechtsen, M. 2000. «Hypernews and Coherence». *Journal of Digital Information*, Volume 1 (7): 12-19.
- Esselink, B. 2001. *A Practical Guide to Localization*. Amsterdam-Filadelfia: John Benjamins.
- Esselink, B. 2006. «The evolution of localization». En A. Pym, A. Perekstenko y B. Starink (eds.). *Translation Technology and its Teaching*. Tarragona: Intercultural Studies Group. 21-30.
- Ezpeleta, P. 2005. «La noción de género en la planificación de la docencia de la traducción de la primera lengua extranjera. En García Izquierdo, I. (ed.). *El género textual y la traducción. Reflexiones teóricas y aplicaciones pedagógicas*. Berna: Peter Lang.
- Faber, P., C. I. López y M. Tercedor Sánchez. 2001. «Utilización de técnicas de corpus en la representación del conocimiento médico». *Terminology* 7 (2): 167-197.
- Fernández Antolín, M. J. 2000. *La problemática de los tipos textuales inglés-español en los estudios de traducción*. Tesis Doctoral sin publicar. Universidad de Valladolid.
- Fernández, C. 2003. «E-banks web sites: an evolving genre». En S. Posteguillo *et al.* (eds.). *Internet in linguistics, translation and literary studies*. Castellón de la Plana: Publicaciones de la Universidad Jaume I. 375-390.
- Fernández Sánchez, F. 2004. *El folleto de cursos de idiomas para extranjeros: Análisis contrastivo (alemán-español) por tipos de emisor y subtextos*. Tesis doctoral. Universidad Pompeu Fabra.
- Firth, J.R. 1957. *Papers in Linguistics*. London: Oxford University Press.
- Firth, J.R. 1935. «The Technique of Semantics». *Transactions of the Philological Society*. 36-72.
- Foralón, D. 2006. «A discipline coming of age in the digital age». En K. Dunne (ed.) *Perspectives on localization*, Amsterdam-Filadelfia: John Benjamins. 195-222.

- 
- Fortanet, I., S. Posteguillo, J. C. Plamer y J. F. Coll (eds.). 1998. *Genre Studies in English for Academic Purposes*. Castellón de la Plana: Publicaciones de la Universitat Jaume I.
- Foltz, P. W. 1996. «Comprehension, Coherence and Strategies in Hypertext and Linear Text». En J. F. Rouet, J. J. Levonen, A. Dillon y Rand J. Spiro. *Hypertext and Cognition*. Nueva Jersey: Lawrence Erlbaum. 109-136.
- Frawley, W. 1984. «Prolegomenon to a theory of translation». En W. Frawley (ed.). *Translation. Literary, linguistic and philosophical perspectives*. London/Toronto: Associated University Presses. 159-175.
- Fritz, G. 1998, «Coherence in hypertext». En W. Bublitz, U. Lenk y E. Ventola (eds.). *Coherence in Spoken and Written Discourse: How to Create it and How to Describe it*. John Benjamins: Amsterdam.
- Frogg, B. J. 2002. «Stanford Guidelines for Web Credibility». A Reserch Summary from the Stanford Persuasive Technology Lab. Stanford University. Disponible en Web: <[www.webcredibility.org/guidelines](http://www.webcredibility.org/guidelines)>.
- Gamero Pérez, S. 2001. *La traducción de textos técnicos*. Barcelona: Ariel.
- Gamero Pérez, S. 2000. *La traducción de textos técnicos y la diversidad tipológica*. En *Sendebarr* 10 (11): 127-158.
- Gamero, Silvia, 1998. *La traducción de textos técnicos (alemán-español). Géneros y subgéneros*. Tesis doctoral. Unversitat Autònoma de Barcelona.
- García Izquierdo, I. y E. Monzó Nebot (2003). «Una enciclopedia para traductores. Los géneros de especialidad como herramienta privilegiada del traductor profesional». En R. Muñoz Martín (ed.). *Actas del I Congreso Internacional de la Asociación Ibérica de Estudios de Traducción*, Granada, Asociación Ibérica de Estudios de Traducción e Interpretación. 83-97.
- García Izquierdo, I. 2006. «El español neutro y la traducción de los lenguajes de especialidad». *Sendebarr* 17: 149-167.
- García Izquierdo, I. 2002. «El género: plataforma de confluencia de nociones fundamentales en didáctica de la traducción». *Discursos. Série Estudos de Tradução*, 2. Universidade Aberta: Lisboa.
- García Izquierdo, I. 2000. *Análisis textual aplicado a la traducción*. Valencia: Tirant lo Blanc.
- García Izquierdo, I. y V. F. Montalt Resurreccion. 2002. «Translating into Textual Genres». *Linguistica Antverpiensa* 1: 135-143.
- Garrand, T. 2001. *Writing for Multimedia and the Web*. Woburn, MA: Butterworth-Heinmann

- Gellerstam, M. 1996. «Translation as a source for cross-linguistic studies». En K. Aijmer, B. Altenberg y M. Johansson (eds.). *Languages in Contrast*. Lund: Lund University Press. 159-175.
- Gellerstam, M. 1986. «Translationese in Swedish novels translated from English». En Wollin, L. Y H. Lindquist (eds.), *Translation Studies in Scandinavia, Proceedings from the Scandinavian Symposium on Translation Theory (SSOTT)*. Lund: CKW Gleerup. 88-95.
- Gile, D. 1995. *Basic Concepts and Models for Interpreter and Translator Training*. Amsterdam-Filadelfia: John Benjamins.
- Giménez, J. 2000. «Business e-mail communication: some emerging tendencies in register». *English for Specific Purposes* 19: 237-251.
- Gläser, R. 1990. *Fachtextsorten im Englischen*. Tübinga: Narr.
- Gömmlich, K. 1993. «Text Typology and Translation-Oriented Text Analysis». En S. E. Wright y L. D. Wright (eds.). *Scientific and Technical Translation*. Amsterdam-Filadelfia: John Benjamins.
- González Aguilar, A., C. Ferreyros Soto y V. Carrascosa López. 2004. *Los contratos en la sociedad de la iformación. Formularios de contratos informáticos e Internet*. Granada: Comares.
- González-Herrero, A. y M. Ruiz de Valbuena. 2006. «Trends in online media relations: Web-based corporate press rooms in leading international companies». *Public Relations Review* 32: 267-275.
- Gömmlich, K. 1993. *Scientific and technical translation*. Amsterdam-Filadelfia: John Benjamins.
- Göpferich, S. 1995a. *Textsorten in Naturwissenschaften und Technik. Pragmatische Typologie- Kontrastierung-Translation*. Tübinga: Gunter Narr.
- Göpferich, S. 1995b. «A Pragmatic Classification of LSP Texts In Science and Technology». *Target* 7: 305-326.
- Gambier, Y. y H. Gottlieb. 2001. «Multimedia, Multilingua: Multiple Challenges». Gambier Y. y H. Gottlieb (eds.). *(Multi) Media Translation. Concepts, Practices, and Research*. Amsterdam: John Benjamins. 8-20.
- Gouadec, D. 2002. *Profession: Traducteur*. Paris: La Maison du Dictionnaire.
- Gouadec, D. 2003. «Le bagage specifice du locateur/localisateur. La vrai «nouvel profil» requis». *Meta* 28 (4): 526-545.

- 
- Grosse, S. y Mentrup, W. (eds.). 1980. *Bürger - Behörde - Formular. Arbeitstagung zum Kommunikationsmittel «Formular»*. Tübingen: Narr.
- Gülich, E. 1981. «Formulare als Dialoge». En Deutsche Akademie Für Sprache und Dichtung, *Der öffentliche Sprachgebrauch (Band II): Die Sprache des Rechts und der Verwaltung*. Stuttgart: Klett-Cotta. 322-356.
- Gülich, E. y R. Raible (eds.). 1972. *Textsorten*. Frankfurt, Alemania: Athenäum.
- Gutiérrez Rodilla, B. A. 2005. *El lenguaje de las ciencias*. Madrid: Gredos.
- Hall, E. 1976. *Beyond Culture*. Garden City, NY: Anchor Press.
- Halliday, M. A. K. 1973. *Explorations in the Functions of Languages*. Londres: Arnold.
- Halliday, M. A. K. y R. Hasan. 1976. *Cohesion in English*. Londres: Longman.
- Halliday, M. A. K. y R. Hasan. 1985. *Language, Context and Text. Aspects of Language in a Social-semiotic Perspective*. Victoria: Deakin University Press.
- Halliday, M. A. K. 1985a. *An introduction to functional grammar*. London: Arnold.
- Halliday, M. A. K. 1985b. «Dimensions of discourse analysis and grammar». En Van Dijk, T. (ed.). *Handbook of discourse analysis*, 2. London: Academic Press. 29-56.
- Hale, C. y J. Scanlon. 1999. *WIRED Style: Principles of English Usage in the Digital Age*. New York: Broadway Books.
- Halverson, S. L. 1999. «Conceptual Work and the “Translation” Concept». *Target* 11 (1): 1-31.
- Halverson, S. L. 1998. «Translation studies and representative corpora: establishing links between translation corpora, theoretical/descriptive categories and a conception of the object of study». *Meta* 43 (4): 494-514.
- Hammerich, I. y C. Harrison. 2001. *Developing Online Content: The Principles of Writing and Editing for the Web*. Nueva York: John Wiley & Sons.
- Hanks, W. F. 1996. *Language and Communicative Practices. Critical Essays in Anthropology*. Oxford: Westview Press.
- Hartmann, R. 1980. *Contrastive Textology. Comparative discourse analysis in Applied Linguistics*. Heidelberg: Groos.
- Hasan, R. 1989. «The identity of a text». En M. A. K. Halliday y R. Hasan (eds.), *Language, text and context*. Oxford: Oxford University Press. 97-118.



- Hatim, B. y I. Mason. 1990. *Discourse and the Translator*. Londres–Nueva York: Longman.
- Hatim, B. y I. Mason. 1997. *The Translator as Communicator*. Londres: Routledge.
- Hatim, B. y J. Munday. 2004. *Translation, an Advanced Resource Book*. Nueva York: Routledge.
- Heinze, N. y Q. Hu. 2006. «The evolution of corporate Web presence: A longitudinal study of large American companies». *International Journal of Information Management* 26 (4): 313-325.
- Herring, S., L. Scheidt, S. Bonus y E. Wright. 2004. «Bridging the gap: A genre analysis of weblogs». Actas del *XXXVIII Annual Hawaii International Conference on Systems Sciences, Digital Media Track*. IEEE Computer Society Press.
- Herrmann, A. y F. Sachse. 2005. «Internacionalización de aplicaciones de software». En Reineke, D. (ed.), *Traducción y Localización*. La Palmas de Gran Canaria: Anroart Ediciones. 45-70.
- Hinges, C. 2000. *Localization, a global manifesto*. Londres: Earthscan Publications.
- Hoffman, L. 1988. *Vom Fachwort zum Fachtext*. Beiträge zur Angewandten Linguistik. Forum für Fachsprachen-Forschung 5. Tübinga: Narr.
- Hofstede, G. 1991. *Cultures and Organizations: Software of the Mind*. Londres: McGraw-Hill.
- Holmes, J. 1988. «The Name and Nature of Translation Studies». En J. Holmes (ed.). *Translated! Papers on Literary Translation and Translation Studies*. Amsterdam: Rodopi. 67-80.
- Holzner, S. 2006. *Secrets of the RSS*. Berkeley: Peachpit Press.
- Holz-Mänttari, J. 1984. *Translatorisches Handeln. Theorie und Methode*. Helsinki: Suomalainen Tiedeakatemia.
- Horton, W. K. 1990. *Designing and Writing Web Documentation: Help files to Hypertext*. Nueva York: Wiley.
- House, J. 1977. *A Model for Translation Quality Assessment*. Tübingen: TBL Verlag Gunter Narr.
- Hunt, K. 1995. «The Design of World Wide Web Home Pages: Using Visuals to Establish Organizational Ethos». *Actas de la Society for Technical Communication (STC)* [en línea]. Disponible en Web: <<http://www.stc.org/confproceed/1995/PDFs/P364367.PDF>>

- 
- Hurtado Albir, A. 2001. *Traducción y Traductología. Introducción a la traductología*. Madrid: Cátedra.
- Hurtado Albir, A. 1999. *Enseñar a traducir. Metodología en la enseñanza de traductores e intérpretes*. Madrid: Edelsa.
- Hurtado Albir, A. 1996. «La traduction: classification et éléments d'analyse». *Meta* 41, (3): 366-377. Disponible en Web: <<http://www.erudit.org/revue/meta/1996/v41/n3>>.
- Hymes, D. 1974. *Foundations in sociolinguistics: an ethnographic approach*. Filadelfia: University of Pennsylvania Press.
- Isenberg, H. 1983. «Probleme der Texttypologie». *Linguistische Studien* 112. 303-342 («Cuestiones fundamentales de tipología textual». En E. Bernárdez (ed.). *Lingüística de texto*. Madrid: Arco Libros. 95-103.
- ISO 9241. 1998. *ISO 9241-11 DIS Ergonomic Requirements for Office Work with Visual Displays (VDTs)*, Part 11: Guidance on Usability.
- Jääskeläinen, R y S. Tirkkonen-Condit. 1991. «Automatised Processes in Professional vs. Non-Professional Translations: a think-aloud protocol study». En S. Tirkkonen-Condit (ed.). *Empirical Research in Translation and Intercultural Studies*. Tubinga: Gunter Narr.
- Jakobson, R. 1960. «Closing statement: Linguistics and Poetics». En Sebeok T. (ed.), *Style in Language*. Cambridge, Mass.: M.I.T. Press. 350-377.
- Janoschka, A. 2003. *Web Advertising*. Amsterdam-Filadelfia: John Benjamins.
- Jeney, C. 2007. *Writing for the Web: a Practical Guide*. Columbus, OH: Pearson Prentice Hall.
- Jiménez Crespo, M. A. 2008a. «Web genres in Localization: a Spanish corpus study». *Localization Focus – The International Journal of Localization*, vol. 6 (1).
- Jiménez Crepo, M. A. 2008b. «Web texts in Translation Training». En M. Garant (ed.). *Current Trends in Translation Teaching and Training*. Helsinki University Translation Studies Department Publication IV. Helsinki, Finlandia: Universidad de Helsinki.
- Jiménez Crespo, M.A. 2008c. «Caracterización del género "sitio web corporativo" español: análisis descriptivo con fines traductológicos». En M. Fernández Sánchez y R. Muñoz Martín (eds.). *Procesos de Traducción e Interpretación*. Granada: Comares.

- Jiménez Crespo, M. A. 2007a. «El hipertexto en la docencia en traductología». En C. Balbuena y A. García (eds.) *Traducción y mediación cultural. Reflexiones interdisciplinarias*. Granada: Editorial Atrio. 390-398.
- Jiménez, Crespo, M. A. 2007b. «Textual Conventions in Localization: What the Industry Can Learn from Translation Studies». Ponencia presentada en el *LISA – Lessius Forum, Teaching Localization for Global Business Readiness* en Lessius College, Amberes, Bélgica. 10 al 12 de diciembre del 2007.
- Jiménez Crespo, M. A. 2007c. «Le genre et le design expérimental d'une étude de corpus comparable de sites web localisées». Ponencia presentada en *La Langue et la Localisation: Politiques, Stratégies et Pratiques* organizada por la Université McGill, Montreal, Canada, 13 al 14 de mayo del 2007.
- Jiménez Crespo, M. A. 2007d. «La coherencia en la localización de hipertextos». Ponencia presentada en el *3rd encuentro de la AIETI* (Asociación Ibérica de Estudios de Traductología) en Barcelona, España, 22 al 25 de marzo del 2007.
- Jiménez Crespo, M. A. 2006a. *La localización de hipertextos: el género y la tipología textual en los sitios web corporativos*. Trabajo de investigación tutelada sin publicar. Universidad de Granada.
- Jiménez Crespo, M. A. 2006b. «El hipertexto en la docencia en traductología». Ponencia presentada en la *5ª Conferencia Internacional Traducción, Texto e Interferencias* de la Universidad de Córdoba, diciembre 2006.
- Jiménez Crespo, M. A. 2005a. «Las peculiaridades textuales de las páginas web localizadas al español». En Greenfield, M. (ed.). *Actas del XLVI Congreso Anual de la American Translator Association*. Seattle, EEUU.
- Jiménez Crespo, M. A. 2005b. «Translating Corporate Webpages: A corpus study». Ponencia presentada en el *46 Congreso anual de la Asociación Americana de Traductores* en Seattle, EEUU, noviembre 2005.
- Jiménez Serrano, O. 2002. *La traducción técnica inglés-español. Didáctica y mundo profesional*. Granada: Editorial Comares.
- Johansson, S. y S. Oksefiell (eds.). 1998. *Corpora and Cross-linguistic Research: Theory, Method and Case Studies*. Amsterdam-Atlanta: Rodopi.
- Jucker, A. H. 2003 «Mass Media Communication at the Beginning of the Twenty Century». *Journal of Historical Pragmatics* 4 (1): 129-148.
- Jung, L. 2000. *La escuela traductológica de Leipzig*. Granada: Comares.

- 
- Kade, O. 1968. *Zufall und Gesetzmässigkeit in der Übersetzung*. Leipzig: Verlag Enzyklopädie.
- Karsch, B. I. 2006. «Terminology workflow in the localization process». En K. Dunne (ed.). *Perspectives on Localization*. Amsterdam-Filadelfia: John Benjamins. 173-191.
- Kennedy, A. y M. Shepherd. 2005. «Automatic Identification of Home Pages on the Web». *Actas del XXXVIII Annual Hawaii International Conference on System Sciences*, Maui, Hawaii. Los Alamitos, CA: IEEE-Computer Society.
- Kenny, D. 2003. «Parallel Corpora and Translation Studies: old questions, new perspectives? Reporting that in Gepcolt. A case study». En G. Barnbrook (ed.), P. Danielsson (ed. y introd.) y M. Mahlberg (ed.), *Meaningful texts: the extraction of semantic information from monolingual and multilingual corpora*. London, England: Continuum. 141-151
- Kenny, D. 2001. *Lexis and Creativity in Translation. A corpus-based study*. Manchester: St. Jerome.
- Kenny, D. 1998. «Creatures of Habit? What translators usually do with words». *Meta* 43 (4): 515-529.
- Kersten, G. E., M. A. Kersten y W. Rakowski, 2001. «Application Software and Culture: Beyond the Surface of Software Interface». *InterNeg.org*. Disponible en Web: <http://interneg.concordia.ca/interneg/research/papers/2001/01.pdf>
- Kilian, C. 1999. *Writing for the Web*. Self-Counsel Press: North Vancouver, Canada.
- Kiraly, D. 1995. *Pathways to Translation. Pedagogy and Process*. Kent, Ohio: Kent State University Press.
- Krug, S. 2005. *Don't make me think*. Indianápolis: New Riders.
- Kuntz, Helmut, 1979. *Zur textsortenmäßigen Binnendifferenzierung des Fachs Kraftfahrzeugtechnik*. Kümmerle: Verlag Göppingen.
- Kussmal, P. 1997. «Text-Type conventions and Translating: Some Methodological Issues». En A. Trosborg (ed.). *Text Typology and Translation*. John Benjamins: Amsterdam. 67-82.
- Kussmaul, P. 1995. *Training the Translator*. Amsterdam-Filadelfia: John Benjamins.
- Lan, L. 2000. «Email: a challenge to Standard English?». *English Today* 16 (4): 23-29.

- Landow, G. 1992. *Hypertext: The convergence of contemporary Critical Theory and Technology*. Baltimore: The John Hopkins University Press.
- Larose, R. 1998. «Méthodologie de l'évaluation des traductions». *Meta* 43 (2): 163-186.
- Laviosa, S. 2002. *Corpus-based Translation Studies*. Amsterdam: Rodopi.
- Laviosa, S. 1998. «The English Comparable Corpus: a Resource and a Methodology». En L. Bowker, M. Cronin, D. Kenny y J. Pearson, (eds.). *Unity in Diversity: Current Trends in Translation Studies*. Manchester: St. Jerome Publishing. 101-112.
- Laviosa, S. 1997. «How Comparable Can 'Comparable' Corpora Be? ». *Target* 9 (2): 289-319.
- Laviosa, S. 1996. *The English Comparable Corpus (ECC): A resource and a Methodology for the Empirical Study of Translation*. Tesis Doctoral. Manchester: Centre for Translation Studies, UMIST.
- Labov, W. 1973. «The Boundaries of Words and their Meanings». En C. J. Bailey y R.W. Shuy (eds.). *New ways of Analyzing Variation in English*. Washington: Georgetown University Press. 340-73.
- Leech, G. 1992. «Corpora and Theories of Linguistic Performance». En A. Svartvik (ed.). *Directions in Corpus Linguistics*. La Haya: Mouton de Gruyter. 105-122.
- Lewis, K. D. 1969. *Convention. A Philosophical Study*. Cambridge, MA: Harvard University Press.
- Lingo Systems, 2004. *Guide to Translation and Globalization*. Portland, OR: Lingo Systems.
- Lommel, A. y R. Ray (ed.) 2007. *LISA Globalization Industry Primer*. Romainmôtier, Suiza: LISA.
- LISA. 2006. *LISA Best Practice Guide: Managing Global Content. Global Content Management and Global Translation Management Systems, 2<sup>nd</sup> edition*. Toon, A., Draheim, A., Lommel y P. Cadieux. 2006 (eds.). Ginebra: The Localization Industry Standards Association (LISA).
- LISA. 2004. *Localiztion Industry Primer, 2nd Edition*. A. Lommel (ed.). Ginebra: The Localization Industry Standards Association (LISA).
- LISA. 2003. *Localization Industry Primer*. D. Fry (ed.). Ginebra: The Localization Industry Standards Association (LISA). Disponible en Web: <<http://www.lisa.org/products/primer.html>>

- 
- Liu, C. *et al.* 1997. «Web Sites of the Fortune 500: Facing Customers through Home Pages». *Information & Management* 31 (1): 335-345.
- Liu, C. y K. P. Arnett. 2000. «An examination of Privacy Policies in Fortune 500 Web Sites». *Mid-American Journal of Business* 17 (1): 13-21.
- Lockwood, R. 2000. «Machine Translation and Controlled Authoring at Carterpillar». En R. C. Sprung (ed.). *Translating into Success. Cutting-edge strategies for going multilingual in a global age*. Amsterdam-Filadelfia: John Benjamins. 187-202.
- Lockwood, R. 1999. «You snooze, you lose». *Language International* 11 (4): 12-16.
- Longacre, R. E. 1983. *The grammar of discourse*. Nueva York: Plenum.
- Lockwood, R. y K. Scott. 1999. *A Writer's Guide to the Internet*. London: Allison and Busby.
- López Rodríguez, C. I. 2001. *Tipologías textuales y conexión en la traducción biomédica inglés*. Tesis doctoral. Universidad de Granada. Disponible en Web: < [http://www.ugr.es/~dpto\\_ti/prof/lrci-tesis.zip](http://www.ugr.es/~dpto_ti/prof/lrci-tesis.zip)>.
- López Rodríguez, Clara I. y M. Tercedor Sánchez. 2004. «Problemas, evaluación y calidad en traducción científica y técnica». *Sendebarr* 15: 29-44.
- López Rodríguez, C. I. y M. Tercedor Sánchez. 2008. «Corpora and Students' Autonomy in Scientific and Technical Translation training». *Jostrans, Journal of Specialized Translation*, 9: 2-19.
- López Quero, S. 2003. *El lenguaje de los «chats». Aspectos gramaticales*. Granada: Port-Royal Ediciones.
- Lörscher, W. 1991. *Translation Performance, Translation Process, and Translation Strategies: A Psycholinguistic Investigation*. Tübingen: Gunter Narr.
- Luna, D., L. A. Peracchio y M. D. de Juan. 2002. «Cross-Cultural and Cognitive Aspects of Web Site Navigation». *Journal of the Academy of Marketing and Science* 30 (4): 397-410
- Lynch, C., 2006. «GMS technology making the localization business case». En K. Dunne (ed.). *Perspectives on Localization*. Amsterdam-Filadelfia: John Benjamins. 37-46.
- Lynch, P. y S. Horton, 2002. *Web Style Guide*. Boston: Yale University Press. Disponible en Web : <<http://www.webstyleguide.com/>>.
- McDonough, J. 2006. «Beavers, Maple Leaves and Maple Trees. A study of National symbols on Localized and Domestic Websites». *Localization Focus* 5 (3): 7 – 14.

- Macizo, P. y M. T. Bajo. 2005. «Memoria de trabajo y traducción». *Cognitiva* 17 (1): 29-54.
- Macklovitch, E. y G. Russel, 2002. «What's been forgotten in translation memory». *Actas del Congreso AMTA 2000*. Cuernavaca: Association for Machine Translation in the Americas. 137-146.
- Macklovitch, E. y G. Russel, 2000. «Machine Translation and Controlled authoring at Caterpillar». En R. C. Sprung(ed.). *Translating into Success: Cutting-Edge Strategies for Going Multilingual in a Global Age*. Amsterdam-Filadelfia: John Benjamins. 187-202.
- Macromedia, 2004. *Using Macromedia Flash MX as a multilingual authoring tool*. Disponible en Web:  
<<http://www.adobe.com/support/flash/languages/flashlocalization/>>
- Malinowski. B. 1923. «The problem of meaning in primitive languages». En C. K. Ogden y I. A. Richards (eds.). *The Meaning of Meaning*. Londres: Kegan Paul.
- Malmkjær, K. 1998. «Love thy neighbour: will parallel corpora endear linguists to translators?». *Meta* 43 (4): 534-541.
- Martin, J. R. 1985. «Process and text: two aspects of human semiosis». En J. D. Benson y W. S. Graves (eds.). *Systemic perspectives on discourse*, vol. 1. Norwood, NJ: Ablex. 6-20.
- Martin, J. R. 1992. *English Text: System and Structure*. Amsterdam: Benjamins.
- Martínez Melis, N. y A. Hurtado Albir. 2001. «Assesment in Translation Studies: Research Needs». *Meta* 47 (2): 272-287.
- Mason, I. 2001. «Translator Behaviour and Language Usage». *Hermes* 26: 65-80.
- Mata Pastor, M. 2005. «Localización y traducción de contenido web». En D. Reineke (ed.). *Traducción y Localización*. La Palmas de Gran Canaria: Anroart Ediciones. 187-252.
- Mata Pastor, M. 2004. «Optimum Cost-Ware in Translator Trainig for the Localisation Market». *Localisation Focus* 3 (4): 11-15.
- Mauranen, A. 2008. «Universal Tendencies in Translation». En G. Anderman y M. Rogers (eds.). *Incorporating Corpora: The Linguist and the Translator*. Clevendon: Multilingual Matters. 32-48.
- Mauranen, A. 2000. «Strange strings in translated language: a study on corpora». En M. Olohan (ed.). *Intercultural Faultlines: Research Models in Translation Studies 1: Textual and Cognitive aspects*. Manchester: St. Jerome. 119-141.

- 
- Maynard, M. y Y. Tian. 2004. «Between global and glocal: Content analysis of the Chinese websites of the 100 top global brands». *Public Relations Review* 30: 285-291.
- Mayor Serrano, M. B. 2002. *Tipología textual pragmática y didáctica de la traducción en el ámbito biomédico*. Tesis doctoral. Universidad de Granada, Servicio de Publicaciones de la Universidad de Granada.
- Mayoral Asensio, R. 2004. «El lenguaje de especificidad y traducción especializada. La traducción jurídica». En C. Gonzalo García y V. García Yebra (eds.). *Manual de documentación y terminología para la traducción especializada*. Madrid: Arco Libros.
- McAlpine, R. 2001. *Web Word Wizardry. A guide for writing for the Web and Intranet*. Berkley, CA: Ten Speed Press.
- McEnery, A. y R. Z. Xiao. 2008. «Paralell and comparable corpora: what are they up to?». En *Incorporating Corpora: Translation and the Linguist. Translating Europe*. Clevedon: Multilingual Matters.
- McEnery, A. y A. Wilson. 1996. *Corpus Linguistics*. Edimburgo: Edinburgh University Press.
- McGovern, G. R. Norton y C. O'Dowd. 2002. *The Web Content Style Guide*. Londres: Prentice Hall.
- Mill, D. 2005. *Content is king*. Oxford: Elsevier.
- Miller, C. R. 1984. «Genre as Social Action». *Quarterly Journal of Speech* 70: 151-67.
- Miller, C. R. 1994. «Rhetorical community: the cultural basis of genre». En A. Freedman y P. Medway (eds.), *Genre and the New Rhetoric*. London: Taylor and Francis. 67-78.
- Miller, C. R. y D. Shepherd. 2004. «Blogging as social action: a genre analysis of the weblog». En L. G. Gurak, S. Antonijevic, L. Johnson, C. Ratliff y J. Reyman (eds.). *Into the blogosphere: rhetoric, community, and culture of weblogs*. Disponible en Web: <<http://blog.lib.umn.edu/blogosphere/>>
- Mendiluce Cabrera, G. 2005. *Estudio comparado inglés/español del discurso biomédico escrito: la secuenciación informativa, la matización asertiva y la conexión argumentativa en la Introducción y la discusión de artículos biomédicos escritos por autores nativos y no nativos*. Tesis doctoral sin publicar. Universidad de Valladolid.
- Meyer, C. F. 2002. *English Corpus Linguistics*. Cambridge: Cambridge University Press.



- Mossop, B. 2006. «How computerization has changed translation». *Meta* 51 (4): 777-793.
- Multilingual Computing and Technology, 2004. *The guide to project management*. Suplemento de la publicación *Multilingual Computing and Technology* 15 (3).
- Müller, E. 2005. «Step-by-step localization». *Multilingual Computing* 75 (16): 16-18.
- Muñoz Sánchez, P. 2008. «En torno a la localización de videojuegos clásicos mediante técnicas de romhacking: particularidades, calidad y aspectos legales». *Jostrans, Journal of Specialized Translation* 9.
- Nielsen, J. 2004. *The need for Web Design Standards*. Disponible en Web: <<http://www.useit.com/alertbox/20040913.html>>.
- Nielsen, J. 2003. *PDF: Unfit for Human*. Disponible en Web: <<http://www.useit.com/alertbox/20030714.html>>
- Nielsen, J. 2000. *Designing Web Usability: the practice of simplicity*. Indianápolis: News Riders.
- Nielsen, J. y H. Loranger. 2006. *Prioritizing Web Usability*. Indianápolis: News Riders.
- Nielsen, J. y M. Tahir. 2002. *Homepage usability: 50 Websites deconstructed*. Indianápolis: News Riders.
- Nielsen J. 1999. *Top ten mistakes revisited*. Alertbox. Disponible en Web: <<http://www.useit.com/alertbox/990502.html>>
- Neubert, A. y M. Shreve. 1992. *Translation as Text*. Kent, Ohio: Kent State University.
- Neumüller, M. 2004. «Hypertext Semiotics». *Razón y Palabra*, 38. Disponible en Web: <<http://www.razonypalabra.org.mx/anteriores/n38/mneumuller.html>>
- Nelson, T. 1993. (Primera edición 1981) *Literary Machines*. Sausalito: Mindful Press.
- Newmark, P. 1981. *Approaches to Translation*. Oxford: Pergamon Press.
- Newmark, P. 1988. *Textbook of Translation*. Oxford: Pergamon Press.
- Nielsen Netratings. 2003. «Global Internet Population Grows an Average of Four Percent Year-Over-Year» [en línea]. Comunicado de prensa del 20 de febrero del 2003. Disponible en Web: <[http://www.nielsen-netratings.com/pr/pr\\_030220\\_hk.pdf](http://www.nielsen-netratings.com/pr/pr_030220_hk.pdf)>

- 
- Nobs, M. 2006. *La traducción de folletos turísticos: ¿Qué calidad demandan los turistas?*. Granada: Comares.
- Nord, C. 2003. «El análisis contrastivo y cultural en la clase de lengua». *Quaderns, Revista de Traducció* 10: 23-39.
- Nord, C. 2001. *Lernziel: Professionelles Übersetzen Spanisch-Deutsch*. Wilhelmsfeld: Gottfried Egert.
- Nord, C. 1997a. «A functional typology of translations». En A. Trosborg (ed.). *Text Typology and Translation*. Amsterdam-Filadelfia: John Benjamins. 43-66.
- Nord, C. 1997b. *Translating as a Purposeful Activity. Functionalist Approaches Explained*. Manchester: St. Jerome.
- Nord, C. 1996. «El error en la traducción: categorías y evaluación». En A. J. Hurtado (ed.). *La enseñanza de la traducción, col. Estudis sobre la traducció* 3. Castelló: Universitat Jaume I. 91-107.
- Nord, C. 1991. *Text Analysis in Translation*. Amsterdam-Atlanta: Rodopi.
- O'Brien, S. 1998. «Practical Experience of Computer-Aided Translation Tools in the Software Localization Industry». En L. Bowker, M. Cronin, D. Kenny y J. Pearson (eds.). *Unity in diversity? Current trends in Translation Studies*. Manchester: St. Jerome Publishing. 115-122.
- O'Hagan, M. y D. Ashworth. 2003. *Translation-Mediated Communication in a digital World: facing the challenges of Globalization and Localization*. Clevedon, Inglaterra: Multilingual Matters.
- Oldenburg, H. 1992. *Angewandte Fachtextlinguistik "conclusiones" und Zusammenfassungen*. Tubinga: Narr.
- Olohan, M. 2004. *Introducing Corpora in Translation Studies*. Londres: Routledge.
- Olohan, M. 2002. «Comparable Corpora in Translation Research: Overview of recent analyses using the Translational English Corpus», *LREC Language Resources in Translation Work and Research Workshop Proceedings*. 5-9.
- Olohan, M. 2001. «Spelling out the Optionals in Translation: A Corpus Study». *UCREL Technical Papers*, 13: 423-432.
- Olohan, M. y M. Baker. 2000. «Reporting that in Translated English. Evidence for Subconscious Processes of Explication?». *Across Languages and Cultures* 1(2): 141-158.
- Orozco, M. 2002. «Revisión de investigaciones empíricas en traducción escrita». *Trans* 6: 63-85.

- PACTE. 2007. «Une recherche empirique expérimentale sur la compétence de traduction». En D. Gouadec, D. *Quelle qualification les traducteurs?*. Paris: Maison du Dictionnaire. 95-116.
- PACTE. 2005. «Investigating Translation Competence: Conceptual and Methodological Issues». *Meta* 50 (2): 609-619.
- PACTE. 2001. «La competencia traductora y su adquisición». *Quaderns, Revista de Traducció* 6: 39-45.
- PACTE. 2000. «Acquiring Translation Competence: Hypotheses and Methodological Problems of a Research Project». En A. Beeby, D. Ensinger y M. Presas (eds.). *Investigating Translation*. Amsterdam-Filadelfia: John Benjamins. 99-106.
- Palomares, R. 1999. «Evaluación de recursos de comunicación para el traductor en Internet». En M. Pinto y J. A. Cordón (eds.). *Técnicas documentales aplicadas a la traducción*. Madrid: Síntesis. 179-193.
- Paltridge, B. 1997. *Genre, Fames and Writting in Research Settings*. Amsterdam-Filadelfia: John Benjamins.
- Paolillo, J. C., J. Warren y B. Kunz. 2007. «Social network and genre emergence in amateur flash multimedia». *Actas del 40 Hawaii International Conference on System Sciences*. Los Alamitos, CA: IEEE Press.
- Pearson, J. 2000. «Surfing the Internet: Teaching Students to Choose their Texts Wisely». En L. Burnard y T. McEnery (eds.). *Rethinking Language Pedagogy from a Corpus Perspective*. Frankfurt: Peter Lang. 235-239.
- Pearson, J. 1999. *Terms in Context*. Amsterdam-Filidelfia: John Benjamins.
- Petöfi, J.S. 1979. *Text versus sentence. Basic questions of Text Linguistics*. Papiere zur Textlinguistik 20, 2. Hamburgo, Alemania: Buske.
- Pierini, P. 2007. «Quality in Web Translation: An Investigation into UK and Italian Tourism Web Sites». *Jostrans, The Journal of Specialized Translation* 8: 85-103.
- Perez Hernandez, M. C. 2002. *Explotación de los córpora textuales informatizados para la creación de bases de datos terminológicas basadas en el conocimiento*. Madrid: CSIC/Elides, volumen 18. Disponible en Web: <<http://elies.rediris.es/elies18/>>
- Peters, C., E. Picchi y L. Biagini. 1996. «Parallel and comparable bilingual corpora in language teaching and learning». E. S. Botley *et al.* (eds.). *Proceedings of Teaching and Language Corpora (UCREL Technical Papers, 9)*. Lancaster: University of Lancaster. 68-80.

- 
- Pailhé, P. 2005. «Is Quality Still fashionable?». *Annual Localization Reader*, 2: 77-80.
- Pöckl, W. 1999. «Kontrastive Textologie». En E. Eckkrammer, N. Hödl y W. Pöckl, *Kontrastive Textologie*. Viena: Edition Praesens. 13-46.
- Posteguillo, S. 2003. *Netlinguistics: languages, discourse and ideology in Internet*. Castellón: Publicaciones de la Universidad Jaume I.
- Posteguillo, S. 2002. «Netlinguistics: Towards a Comprehensive Analytical Framework to Account for Linguistic Change in the Internet». *Ibérica*, 4.
- Posteguillo, S. 1997. *Network English: English in the World of Computers*. Castellón de la Plana: Publicaciones de la UJI.
- Price, J. y Price, L. 2002. *Hot text. Web Writing that works*. Berkeley, CA: News Riders.
- Prieto Velasco J.A., M. I. Tercedor Sánchez y C. I. López Rodríguez. 2007. «Using Multimedia Materials in the Teaching of Scientific and Technical Translation». *Lingüística Antverpiensia*, 1(6): 115-135.
- Puurtinen, T. 1995. *Linguistic Acceptability in Translated Children's Literature* (University of Joensuu Publications in the Humanities N. 15), University of Joensuu: Atk.
- Pym, A. 2005. «Localization. On its nature, virtues and dangers». *Synaps* 17: 17-25.
- Pym, A. 2004a. *The Moving Text: Localization, Translation and Distribution*. Amsterdam-Filadelfia: John Benjamins.
- Pym, A. 2004b. «Localization from the Perspective of Translation Studies: Overlaps in the digital divide?». Ponencia presentada en la *Conferencia SCALLA 2004* Khatmandu. Disponible en Web: <http://www.elda.org/en/proj/scalla/SCALLA2004/Pymv2.pdf>
- Pym, A. 2003. «What localization models can learn from Translation Theory». *The LISA Newsletter. Globalization Insider*, 12 (2/4).
- Pym, A. 1992. *Translation and Text Transfer. An Essay on the Principles of Intercultural Communication*. Frankfurt/Main: Peter Lang.
- Quah, C. K. 2006. *Translation and Technology*. Hampshire, Inglaterra: Palgrave Mcmillan.
- Quirion, M. 2003. «La formation en localisation à l'université : pour quoi faire ?». *Meta* 48 (4): 546-558. Disponible en Web: <http://www.erudit.org/revue/meta/2003/v48/n4/008725ar.html>

- Rabadán, R. 1991. *Equivalencia y traducción: problemática de la equivalencia transléctica inglés-español*. León: Universidad de León.
- RAE. 2001. *Diccionario de la lengua española*. Madrid: Espasa-Calpe, XXI edición. 2001.
- RAE. 2005. *Diccionario Panhispánico de Dudas*. Madrid: Espasa-Calpe.
- Rehm, G. 2002. «Towards Automatic Web Genre Identification – A Corpus-Based Approach in the Domain of Academia by Example of the Academic's Personal Homepage». En *Actas del XXXV Hawaii International Conference on System Sciences (HICSS-35)*. Big Island, Hawaii.
- Reid, I. (ed.). 1987. *The place of Genre in Learning*. Geelong, Australia: Deakin University.
- Reineke, D. 2005. «Herramientas de Localización». En Reineke, D. (ed.). *Traducción y Localización*. La Palmas de Gran Canaria: Anroart Ediciones. 143-160.
- Reineke, D. (ed.). 2005. *Traducción y Localización*. La Palmas de Gran Canaria: Anroart Ediciones.
- Reineke, D. y E. Sánchez de León. 2005. «Perfil laboral y formación de traductores: una encuesta a proveedores de servicios». En D. Reineke (ed.). *Traducción y Localización*. La Palmas de Gran Canaria: Anroart Ediciones. 347-362.
- Reinke, U. 2005. *Selecting Text Material for eContent Localisation Training: Software Localisation Tools*. Saarbrücken: Universität des Saarlandes. Disponible en Web: <[http://ecolore.leeds.ac.uk/downloads/guidelines/selecting\\_text\\_for\\_l10n\\_en.pdf](http://ecolore.leeds.ac.uk/downloads/guidelines/selecting_text_for_l10n_en.pdf)>
- Reinke, U. 2004. *Selecting Text Material for eContent Localisation Training: Translation Memory Tools*. Saarbrücken: Universität des Saarlandes. Disponible en Web: <[http://ecolore.leeds.ac.uk/downloads/guidelines/selecting\\_text\\_for\\_tm\\_en.pdf](http://ecolore.leeds.ac.uk/downloads/guidelines/selecting_text_for_tm_en.pdf)>
- Reiss, K. 1968. «Überlegungen zu einer Theorie der Übersetzungskritik». *Linguística Antverpiensia*, 2: 369-383.
- Reiss, K. 1971. *Möglichkeiten und Grenzen der Übersetzungskritik*. München: Hueber.
- Reiss, K. 1976. *Texttyp und Übersetzungsmethode. Der operative Text*. Kronberg: Scriptor Verlag.
- Reiss, K. 1981. «Type, Kind and Individuality of Text: Decision Making in Translation». *Poetics Today* 2 (4).

- 
- Reiss, K. y H. J. Vermeer. 1984. *Grundlegung einer Allgemeinen Translationstheorie*. Tubinga: Niemeyer. (Traducción al español de S. García Reina y C. Martín de León. 1997. *Fundamentos para una teoría funcional de la traducción*. Madrid: Akal).
- Renau Renau, M. L. 2004. *La página web de la industria cerámico-azulejera: un estudio descriptivo y su aplicación a la enseñanza del inglés para la informática*. Tesis Doctoral. Universidad Jaume I. Disponible en Web:  [<www.tdx.cesca.es/TESIS\\_UJI/AVAILABLE/TDX-26105124531//renau.pdf>](http://www.tdx.cesca.es/TESIS_UJI/AVAILABLE/TDX-26105124531//renau.pdf)
- Robbins, S. S. y A. C. Stylianou. 2003. «Global corporate web sites: an empirical investigation of content and design». *Information & Management* 40: 205-212.
- Roberts G. 1998. «The Home Page as Genre: A Narrative Approach». *Proceedings of the 31st Hawaii International Conference on System Sciences (HICSS-31)*. Hawaii (USA).
- Roberts, R. P. 1996. «Text and Translation». En P. Fernández Nistral y J. M. Bravo Gonzalo (eds.). *A Spectrum of Translation Studies*. Valladolid: Universidad de Valladolid. 37-54.
- Rosch, E. 1978. «Principles of categorization». En E. Rosch y L. B. Lloyd (eds.). *Cognition and categorization*. Erlbaum: Hillsdale.
- Rosch, E. 1977. «Classification of real world objects: origins and representation in cognition». En P. N. Johnson-Laird y P. C. Wagon (eds.). *Thinking: readings in cognitive science*. Cambridge: Cambridge University Press. 212-222.
- Rosch, E. 1975. «Cognitive representation of semantic categories». *Journal of Experimental Psychology* 104: 573-605.
- Rosch, E. 1973. «Natural categories». *Cognitive Psychology* 4: 328-50.
- Rosso, M. 2005. *Using web genre to improve web search*. Tesis doctoral. University of North Carolina, Estados Unidos.
- Ruiz Madrid, N. 2003. «The University Homepage: definition and evolution from a generis perspectiva». En I. Fortanet, J. C. Palmer y S. Posteguillo (eds.). *Linguistic studies in academia and professional English*. Castellón de la Plana: Publicacions de la Universidad Jaume I. 275-298.
- Rumelhart, D. E. 1975. «Notes on a schema for stories». En D. G. Bobrow y A. M. Collins (eds.). *Representation and understanding. Studies in cognitive science*. Nueva York: Academic Press.

- Rumelhart, D. E. 1980. «On evaluation of story grammars». *Cognitive Science* 4: 313-316.
- Sager, J. C. 1997. «Text types and translation». En A. Trosborg (ed.). *Text Typology and Translation*. Amsterdam-Filadelfia: John Benjamins. 25-41.
- Sager, J. «Quality and Standards: The Evaluation of Translations». En C. Picken (ed.). *The Translator's Handbook*. London: ASLIB. 91-102.
- Salager-Meyer, F. 1998. «Rationale behind academic conflict: from outright criticism to contextual 'niche' creation». *UNESCO ALSED-LSP* 21 (2): 4-23.
- Salager-Meyer, F. 1994. «Hedges and textual communicative function in medical English written discourse». *English for Specific Purposes* 13 (2): 149-170.
- Salager-Meyer, F. 1993. «Imprecision and vagueness (hedging) in today's medical discourse: courtesy, coyness and necessity?». *The ESPecialist* 14 (1): 1-15.
- Sammons, M. 1999. *The Internet Writer's Handbook*. Boston: Allyn and Bacon.
- Sánchez Nieto, M. T. 2005. «Sinopsis en catalogo editorial: descripción de un género y comparación español-alemán». *Sendebarr* 16: 97-131.
- Sanchez Nieto, M. T. 2006. «Un formulario contrastivo del género formulario de inscripción a congreso (alemán-español) con aplicación didáctica para la clase de traducción». *Trans, Revista de Traducción* 10: 113-134.
- Sánchez Trigo, E. 2002. «Tipologías textuales y traducción». *Trans, Revista de Traducción* 6: 121-133.
- Sandig, B. 1975. «Zur Differenzierung gebrauchssprachlicher Textsorten im Deutschen». En E. Gülich y W. Raible (eds.). *Textsorten Differenzierungskriterien aus linguistischer Sicht*. Wiesbaden: Athenaion. 113-124.
- Sandrini, P. 2008. «Localization». En Gerzymisch-Arbogast *et al.* (eds.). *Key Issues in LSP Translation*. Amsterdam/Filadelfia: John Benjamins.
- Sandrini, P. 2007. «Mehrsprachigkeit im regionalen Raum: Webauftritte exportorientierter Unternehmen». En L. Zybatow (ed.). *Sprach(en)Kontakt - Mehrsprachigkeit-Translation. Innsbrucker Ringvorlesungen zur Translationswissenschaft V*. Frankfurt am Main: Peter Lang. 327-338.
- Sandrini, P. 2005a. «Website Localization and Translation». Actas de la conferencia *MuTra: Challenges of Multidimensional Translation*. Saarbrücken, Austria. 2-6 mayo 2005.

- Sandrini, P. 2005b. «Translationstechnologie. Überblick und Aussicht». En L. Zybatow (ed.). *Translatologie - Neue Ideen und Ansätze. Innsbrucker Ringvorlesungen zur Translationswissenschaft IV*. Frankfurt: Lang. 203-220.
- Santini, M. 2007a. *Automatic Identification of Genre in Webpages*. Tesis doctoral sin publicar. Universidad de Brighton.
- Santini, M. 2007b. «Characterizing Genres of Web Pages: Genre Hybridism and Individualization». Actas en CD Rom de la *40th Hawaii International Conference on Systems Science (HICSS-40 2007)*, Waikoloa, Big Island, HI, USA. IEEE Computer Society 2007, enero 2007.
- Santini, M. 2005a. «Clustering Web Pages to Identify Emerging Textual Patterns» [en línea]. *Ponencia en la Conferencia RÉCITAL 2005*, Dourdan, Francia. Disponible en Web <<http://www.nltg.brighton.ac.uk/home/Marina.Santini>>.
- Santini, M. 2005b. «Annotated corpora vs. raw web page collections. Text types, web pages, and Linguistic features: Some issues» [en línea]. *Ponencia en el Congreso AAAACL/ICAME*, Ann Arbor, MI. (Disponible en Web: <<http://www.nltg.brighton.ac.uk/home/Marina.Santini/aaacl6-1.pdf>>)
- Saussure, F. de. 1916. *Cours de Linguistique General*. Paris: Payot.
- Schäler, R. 2005. «The Irish Model in Localization» [en línea]. Ponencia en el Congreso *LISA Forum Cairo 2005: Perspectives from the Middle East and Africa*. Disponible en Web: <<http://www.lisa.org/utills/getfile.html?id=61136686>>
- Schäler, R. 2002. «The Cultural Dimension in Software Localization». *Localization Focus* 1 (2).
- Schäler, R. 2001. «Learning to Localize in Limerick». *Language International*, 13 (3): 22-6.
- Schmitt, E. 2000. *The multilingual site blueprint*. Cambridge, MA: Forrester Research. Disponible en Web: <[www.erikseninc.com/pdf/Multi-Lingual-Site-Blueprint.pdf](http://www.erikseninc.com/pdf/Multi-Lingual-Site-Blueprint.pdf)>
- Schmitz, K. D. 2005. «Gestión terminológica en la localización de software». En D. Reineke (ed.). *Traducción y Localización*. La Palmas de Gran Canaria: Anroart Ediciones.
- Scott, M. 1996. *WordSmith Tools (Program)*. Oxford: University Press.
- Searle, A. 1969. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.



- Seguinot, C. 1990. Interpreting errors in translation. *Meta* 35 (1) 68-73.
- Shepherd, M., Watters, C. y A. Kennedy. 2004. «Cybergenre: Automatic Identification of Home Pages on the Web». *Journal of Web Engineering* 3 (3-4): 236-251.
- Shepherd, M. y C. Watters, 1999. «The functionality Attribute of Cybergenres». *Actas del XXXII Hawaii International Conference on System Sciences*, Maui, Hawaii. Los Alamitos, CA: IEEE-Computer Society.
- Shepherd, M. y C. Watters, 1998. «The evolution of cybergenres». En Sprague R. (ed.) *Actas del XXXI Hawaii International Conference on System Sciences*. Los Alamitos, CA: IEEE-Computer Society. 97-109.
- Shreve, G. M. 2006. «Corpus Enhancement and localization». En K. Dunne (ed.). *Perspectives on Localization*. Amsterdam-Filadelfia: John Benjamins. 309-331.
- Singh, N. y A. Pereira. 2005. *The culturally customized Web site: customizing web sites for the global marketplace*. Oxford: Elsevier.
- Singh, N., O. Furrer y O. Massimilaino. 2004. «To localize or to Standarize on the Web: Empirical Evidence from Italy, India, Netherlands, Switzerland and Spain». *Multinational Business Review* 12 (1): 69-88.
- Sinclair, J. 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Snell-Hornby, M. 1995. *Translation Studies. An integrated approach*. Amsterdam: John Benjamins.
- Sprung, R. C. (ed.). 2000. «Translating into Success. Cutting-edge strategies for going multilingual in a global age». Amsterdam-Filadelfia: John Benjamins.
- Stewart, D. 2000. «Conventionality, creativity and translated text: The implications of electronic corpora in translation». En M. Olohan (ed.). *Intercultural faultlines: Research models in translation studies I: Textual and cognitive aspects*. Manchester: St. Jerome. 73-91.
- Storrer, A. 2002. «Coherence in text and Hypertext». *Document Design* 3 (2): 157-168.
- Storrer, A. 1999. «Kohärenz in Text und Hypertext». En Henning Lobin (ed.). *Text im digitalen Medium*. Oplanden: Westdeutscher Verlag. 33-65.
- Stubbs, M. 1996. *Text and Corpus Analysis: Computer-assisted studies of language and culture*. Londres: Blackwell Publishers.

- 
- Sullivan, J. 1999. «What are the functions of corporate home pages?». *Journal of World Business* 34 (2): 193–210.
- Swales, J. 2004. *Research Genres: Explorations and Applications*. Cambridge: Cambridge University Press.
- Swales, J. M 1990. *Genre Analysis. English in Academic and Research Settings*. Cambridge: Cambridge University Press.
- Swales, J. 1981. «Aspects of Article Introductions». *Aston ESP Research Reports*. Language Studies Unit. Birmingham: The University of Aston in Birmingham.
- Symmonds, N. 2002. *Internationalization and localization using Microsoft.Net*. Berkeley, CA: Apress.
- Tarone, E. *et al.* 1988. «On the use of the passive in two astrophysics Journal Papers». *English for Specific Purposes* 17 (1).
- Tercedor Sánchez, M. y Jiménez Crespo, M. A. 2007. «Accesibilidad, imágenes y traducción técnica: un proyecto piloto». Ponencia presentada en *Amadis '07, 2º Congreso de Accesibilidad a los Medios Audiovisuales para Personas con Discapacidad*. Granada, España. 22 de junio del 2007.
- Tercedor Sanchez, M. I. 2005. «Aspectos Culturales en la localización de productos multimedia». En *Quaderns. Revista de Traducció* 12: 151-160.
- Tercedor Sánchez, M. I. 1999. *La fraseología en el lenguaje biomédico: análisis desde las necesidades del traductor*. Tesis Doctoral, Universidad de Granada. Madrid: Csic/Elies, vol 6. Disponible en Web: <<http://elies.rediris.es/elies6/>>.
- Teubert, W. 1996. «Comparable or Parallel Corpora?». *International Journal of Lexicography* 9 (3): 238-265.
- Tirkkonen-Condit, S. 2004. «Unique items: over- and under-represented in translated language?». In A. Mauranen and P. Kujamäki (eds.). *Translation Universals – Do they exist?*. Amsterdam: John Benjamins. 177-184.
- Todorov, T. 1990. *Genres in Discourse*. Cambridge: Cambridge University Press.
- Torruella, J. y J. Llisteri. 1999. «Diseño de corpus textuales y orales». En J. M. Bleca *et al.*, (ed.). *Filología e informática. Nuevas tecnologías en los estudios filológicos*. Seminario de Filología e Informática, Departamento de Filología Española, Universidad Autónoma de Barcelona. Editorial Milenio: Barcelona. 45-77.

- Toury, G. 1998. «A Handful of Paragraphs on 'Translation' and 'Norms'» En C. Schäffner (ed.). *Translation and Norms*. Clevedon: Multilingual Matters, 1998. 10-32.
- Toury, G. 1995. *Descriptive Translation Studies and beyond*. Amsterdam-Filadelfia: John Benjamins.
- Toury, G. 1980. *In Search of a Theory of Translation*. Tel Aviv: The Porter Institute for Poetics and Semiotics.
- Trosborg, A. 2000. «Translating Hybrid Political Texts». En A. Trosborg (ed.). *Analysing Professional Genres*. Amsterdam: John Benjamins. 145-158.
- Trosborg, A. 1997. «Text Typology: Register, Genre and Text Type». En A. Trosborg (ed.). *Text Typology and Translation*. Amsterdam-Filadelfia: John Benjamins. 3-23.
- Tymoczko, M. 1998. «Computerized Corpora And The Future Of Translation Studies». *Meta* 43: 652-658.
- Uren, E., R. Howard y T. Perinotti 1993. *Software Internationalization and Localization: An Introduction*. New York: Van Nostrand-Reinhold.
- Usborne, N. 2002. *Net Words: Creating High-Impact Online Copy*. Nueva York: McGraw Hill.
- Vanderauwera, R. 1985. *Dutch novels translated into English: the transformation of a «minority» literature*. Amsterdam: Rodopi.
- Van Dijk, T. A. 1988. *News as Discourse*. Hillsdale, N.J: Lawrence Erlbaum Associates.
- Van Dijk, T. A. 1980. *Macrostructures*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Van Dijk, T. A. 1977. *Text and Context. Explorations in the Semantics and Pragmatics of Discourse*. Londres: Longman.
- Varantola, K. 2003. «Translators and Disposable Corpora». En F. Zanettin, S. Bernardini y D. Stewart (eds.). *Corpora in Translator Education*. Manchester: St Jerome. 55-70.
- Vaughan, M. W. y A. Dillon. 2006. «Why structure and genre matter for users of digital information: a longitudinal experiment with readers of a web-based newspaper». *International Journal of Human Computer Studies*, 64: 502-525.

- 
- Vázquez y del Árbol, E. 2006. *La redacción y traducción biomédica (inglés-español). Un estudio basado en 200 textos*. Granada: Editorial Universidad de Granada.
- Vázquez y del Árbol, E. 2002. *Propuesta de un análisis comparado de cien textos biomédicos (español e inglés) desde la perspectiva del género*. Tesis doctoral sin publicar. Universidad de Granada.
- Veloso, M. 2005. *Web copy that sells: the revolutionary formula for creating killer copy every time*. Nueva York: Amacom.
- Vela, J. 2005. «La localización de videojuegos». En Reineke, D. (ed.), *Traducción y Localización*. La Palmas de Gran Canaria: Anroart Ediciones. 253-282.
- Vengadasamy, R., A. Jaludin y A. Hamat. 2004. «Characteristics of Written Text in E-Commerce Websites». *Internet Journal of e-Language Learning & Teaching*. 1(2): 15-32.
- Vermeer, H. J. 1978. «Ein Rahmen für eine allgemeine Translationstheorie». *Lebende Sprachen* 23. 99-102.
- Vilamor, J. R. 2001. *Cómo escribir en Internet*. Madrid: Editorial Universitas.
- Vinay, J. P. y J. Dalbernet. 1958. *Stylistique comparée du français et de l'anglais*. Paris: Didier.
- Yates, J. y W. J. Orlikowski. 1992. «Genres of organizational communication: A structural approach to studying communications and media». *Academy of Management Review* 17 (2): 299-326.
- Yates, J. y W. J. Orlikowski 1998. «Genre systems: structuring interaction through communicative norms» [en línea]. Cambridge, MA: MIT. Disponible en Web: <<http://ccs.mit.edu/papers/CCSWP205/>>
- Yunker, J. 2003. *Beyond Borders: Web Globalization Strategies*. Indianápolis, Indiana: New Riders.
- W3C Consortium. Disponible en Web: <<http://www.w3c.es/>>
- Webb, L. E. 1999. *Advantages and disadvantages of translation memory: A cost benefit analysis*. Tesis de Master sin publicar. Monterey School of International Studies.
- Werlich, E. 1976. *A Text Grammar of English*. Heidelberg: Quelle and Meyer.
- Werlich, E. 1975. *Typologie der Texte*. Heidelberg: Quelle & Meyer.
- Wilss, W. 1977. *Übersetzungswissenschaft: Probleme und Methoden*. Stuttgart: Ernst Klett Verlag. (*The Science of Translation. Problems and Methods*. Tübingen, Narr).

- Wittgenstein, L. 1963. *Philosophical Investigations*. Oxford: Blackwell.
- Zanettin, F., S. Bernardini y D. Stewart (eds.). 2003. *Corpora in translator education*. Manchester-Northampton: St. Jerome Publishing.
- Zanettin, F. 2000. «Parallel Corpora in Translation Studies: Issues in Corpus Design and Analysis». En O. Maeve. (ed.) *Intercultural Faultlines. Research Models in Translation Studies I: Textual and Cognitive Aspects*. Manchester: St. Jerome. 105-118.
- Zanettin, F. 1998. «Bilingual Comparable Corpora and the Training of Translators». *Meta* 43 (4): 616-630.

## 14. APÉNDICES

### 14.1. Apéndice I. Composición del corpus comparable

#### Composición de la sección original del corpus

##### DIRECTORIO GOOGLE

1. Total. 248 sitios web
  2. Seleccionados entre el 6 y el 7 de marzo.
  3. Corpus final descargado el 4 de mayo del 2005.
  4. Seleccionado el 1er sitio web de cada subdirectorio.
  5. Subdirectorios excluidos: [Apicultura](#) (1) [portal], [Asociaciones](#) (9) [no son empresas], [Denominaciones de Origen](#) (142), [Asociaciones](#) (98)[quitar este punto], [Compras](#) (1554), [Noticias y publicaciones](#) (25), [Comunidad de Madrid](#) (4), ORGANIZACIONES, [Autoridades Portuarias](#) (21).
  6. SITIOS WEB QUE NO SE PUDIERON RECOPIRAR: Los sitios web marcados con un asterisco (\*) fueron seleccionados pero no se pudieron recopilar en la descarga sincrónica del corpus.
- 
1. [Agricultura y ganadería](#) (169)
    - 1.1. [Acuicultura](#) (2) [Caditec Testing](http://www.caditesting.com/) - http://www.caditesting.com/
    - 1.2. [Consultores](#) (8) [Francesc Casas](http://www.fcasas.com/) - http://www.fcasas.com/
    - 1.3. [Cultivos](#) (19) [Naranjas de Carcaixent](http://www.naranjasdecarcaixent.com/) - http://www.naranjasdecarcaixent.com/
    - 1.4. [Flores y plantas ornamentales](#) (2) [Taliflor](http://www.taliflor.com/) - http://www.taliflor.com/
    - 1.5. [Hortalizas](#) (5) [Horfres](http://www.horfres.com/) - http://www.horfres.com/
    - 1.6. [Semillas y viveros](#) (8) [Viveros Levante](http://www.viveroslevante.com/) - http://www.viveroslevante.com/
    - 1.7. [Ganadería](#) (39) [Miquel Avícola](http://www.miquelavicola.com/) - http://www.miquelavicola.com/
      - 1.7.1. [Porcina](#) (4) [GVP](http://www.acromax.net/) - http://www.acromax.net/
    - 1.8. [Maquinaria y equipo](#) (45) [Azud](http://www.azud.com/) - http://www.azud.com/
    - 1.9. [Materiales y suministros](#) (11) [Compo Agricultura y Jardín](http://www.compo.es/) - http://www.compo.es/ \*
    - 1.10. [Silvicultura](#) (3) [Infosa](http://www.infosa.org/) - http://www.infosa.org/
    - 1.11. [Software](#) (25) [Versas Consultores](http://www.versas.es/) - http://www.versas.es/ \*
  2. [Alimentos y bebidas](#) (536) [Laboratorios Ordesa](http://www.ordesa.es/) - http://www.ordesa.es/
    - 2.1. [Aceites y vinagres](#) (35) [En Origen](http://personal.telefonica.terra.es/web/enorigentienda/index.html) - http://personal.telefonica.terra.es/web/enorigentienda/index.html
    - 2.2. [Bebidas y refrescos](#) (13) [Viared](http://www.viared.es/) - http://www.viared.es/
      - 2.2.1. [Vinos y licores](#) (167) [Grupo Mahou-San Miguel](http://www.mahou-sanmiguel.com/) - http://www.mahou-sanmiguel.com/
    - 2.3. [Café e infusiones](#) (17) [Torrelsa](http://www.torrelsa.com/) - http://www.torrelsa.com/
    - 2.4. [Carnes y aves](#) (23) [Boadas 1880](http://www.boadas.com/) - http://www.boadas.com/ \*
    - 2.5. [Catering](#) (24) [Mokadifusión](http://www.mokadifusion.com/) - http://www.mokadifusion.com
    - 2.6. [Congelados](#) (12) [Fripan](http://www.fripan.com/) - http://www.fripan.com/
    - 2.7. [Conservas y encurtidos](#) (4) [Rianxeira](http://www.rianxeira.com/) - http://www.rianxeira.com/
    - 2.8. [Dulces y aperitivos](#) (34) [Barberá Snacks](http://www.barberasnacks.com/) - http://www.barberasnacks.com
    - 2.9. [Embutidos y jamones](#) (26) [Juan Pedro Domecq](http://www.jamonesdomecq.com/) - http://www.jamonesdomecq.com/
    - 2.10. [Especias y condimentos](#) (1) [Ceylan](http://www.ceylan.es/) - http://www.ceylan.es/
    - 2.11. [Frutas y verduras](#) (26) [Cherrita](http://www.cherrita.com/) - http://www.cherrita.com/
    - 2.12. [Legumbres, arroces y pastas](#) (5) [Arrocerías Herba](http://www.herba.es/) - http://www.herba.es/

- 
- 2.13. [Maquinaria y equipo](#) (27) [Risco Ingeniería e instalaciones](#) - <http://www.risco-sa.com/>
  - 2.14. [Mayoristas y distribuidores](#) (13) [Grupo Dia](#) - <http://www.dia.es/>
  - 2.15. [Pan y pasteles](#) (24) [Croexsa](#) - <http://www.croexsa.com>
  - 2.16. [Pescados y mariscos](#) (30) [Conservas Valcarcel](#) - <http://www.valcarcel-sa.es/>
  - 2.17. [Productos lácteos](#) (16) [Leche Pascual](#) - <http://www.lechepascual.es/> \*
  
  - 3. [Artes gráficas](#) (179)
    - 3.1. [Edición](#) (67)
      - 3.1.1. [Editoriales](#) (57) [Mapas murales](#) - <http://www.mapasmurales.com/>
      - 3.1.1. [Editoriales](#) (57) [Masson](#) - <http://www.masson.es/> \*
      - 3.1.1. [Editoriales](#) (57) [Grupo Anaya](#) - <http://www.anaya.es/> \*
    - 3.2. [Impresión](#) (45)
      - 3.2.1. [Imprentas](#) (4) [Sticard](#) - <http://www.sticard.com/>
      - 3.2.1. [Imprentas](#) (4) [Mawijo](#) - <http://www.mawijo.com/>
    - 3.3. [Maquinaria y equipo](#) (10) [Empress Servicios](#) - <http://www.franquicopy.com/>
    - 3.4. [Materiales y suministros](#) (6) [Milan-Factis](#) - <http://www.areamilan.com/> \*
  
  - 4. [Artes y entretenimiento](#) (265)
    - 4.1. [Fotografía](#) (29) [Aymerich Golf Management](#) - <http://www.aymerichgolf.com>
    - 4.1. [Fotografía](#) (29) [Fotolab 30](#) - <http://www.fotolab30.com/> \*
    - 4.2. [Música](#) (45) [Ars Antiqua](#) - <http://www.ars-antiqua.com/> \*
    - 4.2.1. [Producción discográfica](#) (38) [Acuarela Discos](#) - <http://www.acuareladiscos.com>
    - 4.3. [Producción audiovisual](#) (108) [Especialistas de cine - Xtrem Solutions](#) - <http://www.xtremolutions.com>
    - 4.3.1. [Madrid](#) (13) [Yllana Producciones](#) - <http://www.yllana.com/>
  
  - 5. [Comercio internacional](#) (35)
    - 5.1. [Transporte y logística](#) (3) [Consultorías](#) (13)
    - 5.1. [Transporte y logística](#) (3) [Ocre-bl](#) - <http://www.ocre-bl.com/>
    - 5.1. [Transporte y logística](#) (3) [Europamar](#) - <http://www.europamar.com/>
  
  - 6. [Construcción](#) (644)
    - 6.1. [Constructoras](#) (49) [Hansa Urbana](#) - <http://www.hansaurbana.com/>
    - 6.2. [Consultores](#) (4) [Grupo Masa](#) - <http://www.grupomasa.com>
    - 6.3. [Contratistas](#) (56) [ERS - Geotecnia](#) - <http://www.ersgeotecnia.com/>
    - 6.4. [Diseño](#) (64) [Enerpal](#) - <http://www.enerpal.com/> \*
    - 6.4.1. [Arquitectura](#) (19) [Teledetección y Gis](#) - <http://www.geosys.es/>
    - 6.4.2. [Interiores](#) [Campos y Pradel](#) - <http://www.campospradel.com>
    - 6.5. [Ingeniería Ambiental](#) (4) [Almagra y Siena](#) - <http://www.almagraysiena.com/>
    - 6.6. [Mantenimiento](#) (5) [Aisnorsa](#) - <http://www.aisnorsa.com/>
    - 6.7. [Maquinaria y equipo](#) (51) [3i Multihogar](#) - <http://www.3i-multihogar.com/> \*
    - 6.8. [Materiales y suministros](#) (319) [Recamasa](#) - <http://www.recamasa.com/> \*
    - 6.8.1. [Ascensores](#) (6) [Alkor](#) - <http://www.alkor.es/> \*
    - 6.8.2. [Climatización](#) (26) [Ascensores Imem](#) - <http://www.imem.com/>
    - 6.8.2.1. [Aire acondicionado](#) (11) [Aplicaciones Naturales del Sol, Ansol](#) - <http://www.ansol.es/>
    - 6.8.2.1. [Aire acondicionado](#) (11) [Ventilación industrial](#) - <http://www.ventilacion.com/>
    - 6.8.3. [Electricidad](#) (3) [Jung Electro Ibérica](#) - <http://www.jungiberica.es/>
    - 6.8.4. [Fontanería](#) (4) [Saniber](#) - <http://www.saniber.com/>
    - 6.8.5. [Hormigón](#) (11) [Sika](#) - <http://www.sika.es/>
    - 6.8.6. [Mayoristas y distribuidores](#) (5) [Cemento Natural](#) - <http://www.cementonatural.com/>
    - 6.8.7. [Metales](#) (11) [Europerfil](#) - <http://www.europerfil.es/> \*
    - 6.8.8. [Mobiliario](#) (3) [Duscholux](#) - <http://www.duscholux.es/>
    - 6.8.8.1. [Cocina y baño](#) (12) [Burgalia](#) - <http://www.buraglia.com/>
    - 6.8.8.2. [Oficina](#) (15) [Mobiofi - Muebles de Oficina](#) - <http://www.mobiofi.com/>
    - 6.8.9. [Pinturas y recubrimientos](#) (8) [MK2](#) - <http://www.mk2.es/>
    - 6.8.10. [Puertas y ventanas](#) (19) [Decortina](#) - <http://www.decortina.com/>
    - 6.9. [Piscinas](#) (10) [Acuarama](#) - <http://www.aquarama.org/>
    - 6.9.1. [Maquinaria y equipo](#) (8) [La Hita](#) - <http://www.lahitaalquiler.com/>
    - 6.9.2. [Materiales y suministros](#) (7) [Textura Decoración S.L.](#) - <http://www.instaladoresonline.com>
    - 6.9.3. [Mercado Inmobiliario](#) (59) [Casas Roda Golf](#) - <http://www.rodagolf.com/> \*
    - 6.10. [Plantas industriales](#) (6) [Ingeniería y Construcción de Naves, Iconsa](#) - <http://www.iconsa.es/>
    - 6.11. [Promotoras](#) (9) [Restaura](#) - <http://www.restaura.es/> \*

- 6.12. [Urbanización](#) [Nuevo Mundo](#) - <http://www.promociones-inmobiliarias-promotoras.com/> \*
- 6.12.1. [Constructoras](#) (49) [Construcciones especiales y dragados](#) - <http://www.drace.com/>
- 6.12.2. [Consultores](#) (4) [Corman Sistemas](#) - <http://www.cormansis.com/> \*
- 6.12.3. [Contratistas](#) (56) [Constructora](#) [Hispánica](#) - <http://www.constructorahispanica.com/> \*
7. [Diseño](#) (67) [Hexel Anzeigen](#) - <http://www.hexelanzeigen.com/> \*
- 7.1. [Páginas web](#) (22) [Prodisweb](#) - <http://www.prodisweb.com/>
- 7.2. [Rótulos](#) (7) [Cartabón](#) - <http://www.cartabon.net/>
8. [Empleo](#) (177) [HayGroup](#) - <http://www.haygroup.es/>
- 8.1. [Agencias de selección](#) (35) [Grupo RH Asesores](#) - <http://www.gruporh.com>
- 8.2. [Trabajo temporal](#) [Adecco](#) - <http://www.adecco.es/>
9. [Emprendedores](#) (27) [Melior](#) - <http://www.melior.es/>
10. [Finanzas](#) (381) [Grupo financiero Riva y García](#) - <http://www.rivaygarcia.es/cas/index.htm>
- 10.1. [Entidades financieras](#) (93)
- 10.1.1. [Bancos](#) (44) [BSCH](#) - <http://www.gruposantander.es/>
- 10.1.2. [Cajas de ahorros](#) (43) [La Caixa](#) - <http://portal1.lacaixa.es/Home/Inicio/> \*
- 10.2. [Financiamiento](#) (55) [Reunificación de préstamos](#) - <http://www.reunificaciondeprestamos.com>
- 10.2.1. [Hipotecas](#) (22) [Hipoteca Remunerada de Barclays Bank](#) - <http://www.hipotecaremunerada.barclays.es/> \*
- 10.2.2. [Tarjetas de crédito](#) (8) [Tarjetas de crédito de Caja Madrid](#) - <http://www.cajamadrid.es/CajaMadrid/Home/puente?pagina=3662>
- 10.3. [Inversiones](#) (83) [Banesto Broker](#) - <http://www.banestobroker.com/> \*
- 10.3.1. [Agencias de valores](#) (3) [Estubroker](#) - <http://www.estubroker.com>
- 10.3.2. [Capital de riesgo](#) (5) [Innova 31](#) - <http://www.innova31.com/>
- 10.3.3. [Gestoras de fondos](#) (2) [Gebasa](#) - <http://www.gebasa.es>
- 10.3.4. [Software](#) (2) [Software Financiero Bolsa S.A.](#) - <http://www.sfb.es>
- 10.4. [Medios de pago](#) (9) [4B](#) - <http://www.4b.es/> \*
- 10.5. [Seguros](#) [Sanitas](#) - <http://www.sanitas.es>
11. [Hostelería](#) (39) [In Blan](#) - <http://www.inblan.com/>
12. [Industrias](#) (1038) [Canon España](#) - <http://www.canon.es/> \*
- 12.1. [Administración de residuos](#) (20) [Detesa](#) - <http://www.detesa.com> \*
- 12.1.1. [Tratamiento de agua](#) (19) [Compañía de Aguas BelNature](#) - <http://www.belnature.com/>
- 12.2. [Automotriz](#) (53) [Seat España](#) - <http://www.seat.es>
- 12.3. [Calzado y complementos](#) (50) [PabloSky](#) - <http://www.pablosky.com/> \*
- 12.4. [Eléctrica y electrónica](#) (136) [Instrumentos de Medida](#) - <http://www.idm-instrumentos.es/>
- 12.4.1. [Automatismos y PLC](#) (7) [ASING](#) - <http://www.asing.es/>
- 12.4.2. [Cuadros y electricidad](#) (3) [Unión Fenosa](#) - <http://www.unionfenosa.es/>
- 12.4.3. [Energía e industrial](#) (4) [Endesa Oficina On-Line](#) - <http://www.endesaonline.com>
- 12.4.4. [Equipos e instrumentación](#) (1) [Mp Electronics](#) - <http://www.mpelectronics.info/>
- 12.4.5. [Iluminación](#) (15) [Salvi](#) - <http://www.salvi.es> \*
- 12.4.6. [Ingeniería y diseño](#) (2) [Diselec](#) - <http://www.diselec.com/>
- 12.4.7. [Materiales y suministros](#) (5) [Cebec](#) - <http://www.cebek.com/>
- 12.4.8. [Mayoristas y distribuidores](#) (8) [HT Instruments](#) - <http://www.htinstruments.com/>
- 12.4.9. [Productos de consumo](#) (7) [Bosch España](#) - <http://www.bosch-ed.com/>
- 12.4.10. [Sonido](#) (3) [Reflexion Arts](#) - <http://www.reflexion-arts.com>
- 12.5. [Empaques y embalajes](#) (60) [Saint Gobain La Granja](#) - <http://www.saintgobainlagranja.es/>
- 12.6. [Energía](#) (29) [Iberdrola](#) - <http://www.iberdrola.es/> \*
- 12.7. [Extractivas](#) (8) [Tino](#) - <http://www.tino.es> \*
- 12.8. [Maquinaria y equipo](#) (177) [Korta S.A.](#) - <http://www.korta-sa.com/> \*



- 
- 12.8.1. [Agrícola](#) (45) [Arteriego](#) - <http://www.arteriego.com/>
  - 12.8.2. [Comercial](#) (9) [Freecaad](#) - <http://www.freecaad.com/>
  - 12.8.3. [Laboratorios](#) (3) [Flores Valles](#) - <http://www.floresvalles.es/> \*
  - 12.8.3.1. [Textil](#) (12) [Jose Fuster](#) - <http://www.fuster.com>
  - 12.8.4. [Materiales](#) (115) [Codeca](#) - <http://www.cauchocodeca.com/> \*
  - 12.8.4.1. [Madera](#) (26) [Maderas Manuel Sanz](#) - <http://www.bemarnet.es/manuelsanz/>
  - 12.8.4.2. [Metales](#) (20) [Fundesa](#) - <http://www.fundesa.es/> \*
  - 12.8.4.3. [Papel y cartón](#) (46) [Fabregas](#) - <http://www.fabregas.com/>
  - 12.8.4.3.1. [Pasta, papel y cartón](#) (38)
  - 12.8.4.3.1.1. [Distribuidores](#) (4) [Kea Papel](#) - <http://www.keapapel.com/empresa.htm>
  - 12.8.4.3.1.2. [Fabricantes](#) (11) [ENCE: Empresa Nacional de Celulosas](#) - <http://www.ence.es/>
  - 12.8.4.3.1.3. [Maquinaria](#) (3) [Maquinaria Ceres](#) - <http://www.macesa.com/> \*
  - 12.8.4.3.1.4. [Transformadores](#) (14) [Cartonajes Levante](#) - <http://www.cartonajeslevante.com> \*
  - 12.8.4.4. [Plásticos](#) (42) [Grup Sigma](#) - <http://www.grupsigma.es/> \*
  - 12.8.4.5. [Vidrio](#) (16) [Vidrieras Victor](#) - <http://www.vidrieras-victor.com/>
  - 12.8.5. [Muebles](#) (128) [Colchones Relax](#) - <http://www.relax.es> \*
  - 12.8.6. [Química](#) (186) [Cepsa](#) - <http://www.cepsa.es> \*
  - 12.8.6.1. [Farmacéutica](#) (21) [Almirall Prodesfarma](#) - <http://www.almirallprodesfarma.com/> \*
  - 12.8.7. [Siderurgia](#) (36) [Olarra](#) - <http://www.olarra.com> \*
  - 12.8.8. [Suministros](#) (66) [Carbuos Metálicos](#) - <http://www.carbuos.com/>
  - 12.8.9. [Óptica](#) (1) [Grupo Prats](#) - <http://www.grupoprats.com>
  
  - 13. [Informática](#) (943)
  - 13.1.1. [Consultores](#) (49) [Indra Sistemas](#) - <http://www.indra.es/>
  - 13.1.1.1. [Internet](#) (25) [Davinci Consulting Tecnológico](#) - <http://www.dvc.es/>
  - 13.1.1.1.1. [E3](#) - <http://www.f-3.net> \*
  - 13.1.1.1.2. [Usabilidad](#) (21) [Xperience Consulting](#) - <http://www.xperienceconsulting.com>
  - 13.1.1.2. [Formación](#) (23) [Cursos Aulacrea](#) - <http://www.aulacrea.com/> \*
  - 13.1.3. [Hardware](#) (67) [Microrent](#) - <http://www.microrent.es>
  - 13.1.3.1. [Fabricantes](#) (18) [Inves](#) - <http://www.inves.es/> \*
  - 13.1.3.2. [Periféricos](#) (14) [Laboratorios Electrónicos Reset](#) - <http://www.todoimpresoras.com/>
  - 13.1.4. [Internet](#) (404) [Acceso a la Red Internet Española, ARI](#) - <http://www.ari.es/>
  - 13.1.4.1. [Comercio electrónico](#) (10) [YLos](#) - <http://www.ylos.com>
  - 13.1.4.2. [Hospedaje de sitios](#) (209) [Giralweb](#) - <http://www.giralweb.com>
  - 13.1.4.2.1. [Promoción de sitios](#) (11) [Discovery Net](#) - <http://www.discoverynet.es/> \*
  - 13.1.4.2.2. [Proveedores de acceso](#) (52) [TeleCable](#) - <http://www.telecable.es> \*
  - 13.1.4.2.3. [Registro de dominios](#) (15) [Namelia](#) - <http://www.namelia.net/>
  - 13.1.4.2.4. [WAP](#) [Terra Mobile](#) - <http://www.terra-mobile.es/>
  - 13.1.5. [Mayoristas y distribuidores](#) (33) [Diode España](#) - <http://www.diode.es>
  - 13.1.6. [Seguridad](#) (13) [Internet Security Auditors](#) - <http://www.isecauditors.com/>
  - 13.1.7. [Software](#) (207) [Panda Software](#) - <http://www.pandasoftware.es/>
  - 13.1.7.1. [ERP](#) (34) [Aitana](#) - <http://www.aitanasbs.com>
  - 13.1.8. [Tiendas de informática](#) [Tiendapc](#) - <http://www.tiendapc.com>
  
  - 14. [Marketing](#) (205)
  - 14.1.1. [Consultores](#) (14) [Sinergia: Imagen, marketing y comunicación](#) - <http://www.sinergia.es/>
  - 14.1.2. [Exposiciones y ferias](#) (25) [Altamira Congresos](#) - <http://www.altamira-congresos.com> \*
  - 14.1.2.1. [Traductores](#) [Ricard Sierra](#) - <http://www.ricardsierra.com/>
  - 14.1.3. [Investigación de mercados](#) (16) [Netquest](#) - <http://www.netquest.es/>
  - 14.1.4. [Marketing en Internet](#) (11) [Interactive Network](#) - [www.antevenio.com/](http://www.antevenio.com/) \*
  - 14.1.5. [Marketing promocional](#) (34) [Publiestilo](#) - <http://www.publiestilo.com>
  - 14.1.6. [Publicidad](#) (76) [Margen Design](#) - <http://www.margen.com>
  - 14.1.6.1. [Agencias](#) (17) [Tasmanet](#) - <http://www.tasma.net/> \*
  - 14.1.6.2. [Medios](#) (14) [Tecnodimensión](#) - <http://www.tecnodimension.com/>
  - 14.1.7. [Relaciones públicas](#) (13) [Atento Telecomunicaciones España](#) - <http://www.atento.es>

- 14.1.8. [Servicios de apoyo](#) (10) [Buzoneofacil](#) - <http://www.buzoneofacil.com>
- 14.1.9. [Servicios globales](#) (7) [BGS](#) - <http://www.bgssl.com>
15. [Mayoristas y distribuidores](#) (25) [Ortus Fitness](#) - <http://www.ortus.com/> \*
- 15.1.1. [Suministros](#) [Santiago Vargas](#) - <http://www.santiagovargas.com> \*
16. [Medio ambiente](#) (109) [Integración y Sistemas de Energías Renovables, ISER](#) - <http://www.iser.es/>
- 16.1. [Consultores](#) (17) [Ingenieros Asesores](#) - <http://www.ingenierosasesores.com/>
17. [Mercado inmobiliario](#) (221) [Agrofinca](#) - <http://www.agrofinca.com/>
- 17.1. [Administradores de fincas](#) (9) [Administrador de fincas en Valencia](#) - <http://www.joseroig.com>
- 17.2. [Agencias inmobiliarias](#) (44) [Remax España](#) - <http://www.remax.es/> \*
- 17.3. [Alquiler](#) (6) [CEVASA](#) - <http://www.cevasa.com/> \*
- 17.4. [Consultorías](#) (12) [Aguirre Newman](#) - <http://www.aguirrenewman.es/> \*
- 17.5. [Tasadoras](#) (25) [Tecnitasa](#) - <http://www.tecnitasa.es/> \*
18. [Oportunidades](#) (21)
- 18.1. [Cursos en línea](#) (5) [Nomaster](#) - <http://www.nomaster.net/>
- 18.2. [Franquicias](#) (15) [Tormo](#) - <http://www.tormo.com/>
19. [Productos de consumo](#) (269)
- 19.1.1. [Deportivos](#) (5) [Conor Sports](#) - <http://www.conor.es/>
- 19.1.2. [Oficina](#) (5) [Sie](#) - <http://www.siesl.com>
- 19.1.3. [Ropa](#) (73) [Etxe-Ondo](#) - <http://www.etxeondo.com/> \*
20. [Recursos humanos](#) (21) [Verja y Asociados](#) - <http://www.verdejayasociados.com/>
- 20.1. [Software](#) (42) [Seresco](#) - <http://www.seresco.es> \*
21. [Salud](#) (51)
- 21.1. [Discapacitados](#) (47) [Hiperion Biotech](#) - <http://www.hiperionbiotech.com/>
- [Intervención Social, IS](#) - <http://www.intervencionsocial.com/>
22. [Seguridad](#) (141)
- 22.1. [Contra incendios](#) (9) [Securitas](#) - <http://www.securitas.es/>
- [Extintores Eivar](#) - <http://www.eivar.com/> \*
- 22.2. [Industrial y laboral](#) (16) [Grupo Prevenir Consulting](#) - <http://www.grupoprevenir.es/>
- 22.3. [Investigación](#) (28) [Alpa-mk](#) - <http://www.alpa-mk.com>
- [Agencias de detectives](#) (20) [Detectives Privados Monopol](#) - <http://www.monopol.es/>
- 22.4. [Productos y equipos](#) (31) [Comercial Internacional de Seguridad, Sit](#) - <http://www.sit.es/>
- 22.5. [Vial](#) (5) [Grupo tecnología del tráfico](#) - <http://www.grupotrafico.es/>
- 22.6. [Vigilancia](#) [Ceca Seguridad](#) - <http://www.cecaseguridad.com/>
23. [Servicios](#) (844)
- 23.1. [Cerrajería](#) (5) [Tracasa](#) - <http://www.tracasa.es>
- [Arregui Box](#) - <http://www.arreguibox.com/> \*
- 23.2. [Consultoría](#) (188) [Con Papeles](#) - <http://www.conpapeles.com/>
- 23.2.1. [Contabilidad](#) (12) [Iasplus Spain](#) - <http://iasplus.deloitte.es/>
- 23.2.2. [Ingeniería](#) (24) [Abengoa](#) - <http://www.abengoa.es> \*
- 23.2.3. [Legal](#) (142) [Garrigues](#) - <http://www.garrigues.com/>
- 23.2.4. [Outsourcing](#) (17) [Adea](#) - <http://www.adea.es/>
- 23.2.5. [Turismo](#) (4) [Culture.02](#) - <http://www.culture02.com/html.html>
- 23.3. [Distribución y logística](#) (46) [Cega](#) - <http://www.cegamultidistribucion.com/> \*
- 23.4. [Eventos](#) (81) [Organización de eventos y fiestas](#) - <http://www.jptres.com/>
- 23.4.1. [Fiestas y celebraciones](#) (8) [Iluminaciones Truman](#) - <http://www.codsc.com/iluminacionstruman/>
- 23.4.1.1. [Bodas](#) (5) [Terron Bodas](#) - <http://www.terronbodas.com/>
- 23.5. [Fontanería](#) (4) [Tecnogas](#) - <http://www.tecnogas-sa.com/>
- 23.6. [Funerarios](#) (23) [Servisa](#) - <http://www.servisa.es>
- 23.7. [Limpieza](#) (112) [Dintre i Fora](#) - <http://www.dintreifora.com/>

- 23.7.1. [Control de plagas](#) (14) [Serkonten](http://www.serkonten.com/) - <http://www.serkonten.com/> \*
- 23.7.2. [Maquinaria y equipo](#) (29) [Alberto Casañal](http://www.maquinarialimpieza.com) - <http://www.maquinarialimpieza.com>
- 23.7.3. [Tintorerías y lavanderías](#) (5) [Lavapant](http://www.lavapant.com/spanish/index.asp) - <http://www.lavapant.com/spanish/index.asp>
- 23.7.4. [Viaria](#) [Prodesa Medioambiente](http://www.prodesa.net/) - <http://www.prodesa.net/> \*
- 23.8. [Mensajería](#) (12) [STU Express Service](http://www.stu.es/) - <http://www.stu.es/>
- 23.9. [Mudanzas](#) (47) [Cajas para mudanzas](http://www.cajasdemudanzas.com/) - <http://www.cajasdemudanzas.com/>
- 23.10. [Veterinarios](#) (48) [Analítica Veterinaria](http://www.analiticaveterinaria.com/) - <http://www.analiticaveterinaria.com/> \*
24. [Telecomunicaciones](#) (160) [Unitronics](http://www.comunicaciones.unitronics.es/) [Comunicaciones](http://www.comunicaciones.unitronics.es/) - <http://www.comunicaciones.unitronics.es/>
- 24.1. [Consultores](#) (24) [Satec](http://www.satec.es/) - <http://www.satec.es/>
- 24.1.1. [Telefonía](#) (3) [Ypertron](http://www.ypertron.com) - <http://www.ypertron.com>
- 24.2. [Fabricantes](#) (7) [Tecatel](http://www.tecatel.com/) - <http://www.tecatel.com/>
- 24.3. [Mayoristas y distribuidores](#) (6) [Jovitel Telecomunicaciones](http://www.jovitel.com/) - <http://www.jovitel.com/> \*
- 24.4. [Operadores](#) (13) [Jazztel Telecomunicaciones](http://www.jazztel.com/) - <http://www.jazztel.com/>
- 24.4.1. [ADSL](#) (13) [ADSL Ya](http://acceso.ya.com/adsl/) - <http://acceso.ya.com/adsl/>
- 24.4.2. [Cable](#) (4) [Ono](http://www.ono.es/) - <http://www.ono.es/>
- 24.4.3. [Telefonía](#) (8) [Retevisión](http://www.retevision.es/) - <http://www.retevision.es/> \*
- 24.4.3.1. [Móvil](#) (26) [Mobipay](http://www.mobipay.es) - <http://www.mobipay.es> \*
- 24.4.4. [Fabricantes](#) (5) [Nokia](http://www.nokia.es/) - <http://www.nokia.es/>
- 24.4.5. [Móvil](#) (20) [Movistar](http://www.movistar.es/) - <http://www.movistar.es/> \*
- 24.4.6. [Operadores](#) (8) [Telefónica](http://www.telefonica.es/) - <http://www.telefonica.es/>
25. [Textiles](#) (172) [Alfombras](http://www.alfombrasbelka.com/) [y](http://www.alfombrasbelka.com/) [Moquetas](http://www.alfombrasbelka.com/) [Belka](http://www.alfombrasbelka.com/) - <http://www.alfombrasbelka.com/>
- 25.1. [Materiales y suministros](#) (9) [Ideatint](http://www.ideatint.com/) - <http://www.ideatint.com/>
- 25.2. [Mayoristas y distribuidores](#) (6) [Carmen Rey](http://www.carmenrey.com/) - <http://www.carmenrey.com/>
26. [Transportes y logística](#) (292) [Alcoarte](http://www.alcoarte.com/) - <http://www.alcoarte.com/>
- 26.1. [Alquiler de barcos](#) (16) [Cosas de Barcos](http://www.cosasdebarcos.com/) - <http://www.cosasdebarcos.com/>
- 26.2. [Aéreos](#) (18) [Iberia](http://www.iberia.es) - <http://www.iberia.es> \*
- 26.3. [Consultores](#) (2) [Geswinner](http://www.geswinner.com/) - <http://www.geswinner.com/>
- 26.4. [Ferrovianos](#) (7) [RENFE](http://www.renfe.es/) - <http://www.renfe.es/>
- 26.5. [Marítimos y fluviales](#) (19) [Trasmediterránea](http://www.trasmediterranea.es/) - <http://www.trasmediterranea.es/>
- 26.6. [Por carretera](#) (40) [Europcar](http://www.europcar.es/) - <http://www.europcar.es/>
27. [Turismo](#) (86)
- 27.1. [Agencias de viajes](#) (62) [Viajar](http://www.viajar.com/) - <http://www.viajar.com/>
- 27.2. [Alquiler de vehículos](#) (32) [Victoria Rent a Car](http://www.victoriacars.com/) - <http://www.victoriacars.com/>

## Composición de la sección localizada del corpus

<b>Empresas Públicas</b>	<b>Posición en Lista Forbes mundial</b>	<b>Sitio web localizado hacia la <i>locale</i> «es-ES»</b>
1. Citygroup	1	<a href="http://www.citibank.com/spain/consumer/spanish/">http://www.citibank.com/spain/consumer/spanish/</a>
2. General Electri	2	<a href="http://www.gelighting.com/es/index.htm">http://www.gelighting.com/es/index.htm</a>
3. IBM	23	<a href="http://www.ibm.com/es/">http://www.ibm.com/es/</a>
4. Altria Group - Phillip Morris	27	<a href="http://www.philipmorrisoninternational.com/ES/pages/spa_ES/Default.asp">http://www.philipmorrisoninternational.com/ES/pages/spa_ES/Default.asp</a>
5. Microsoft	55	<a href="http://www.microsoft.com/spain/">http://www.microsoft.com/spain/</a>
6. Johnson & Johnson	57	<a href="http://www.es.acuvue.com/">http://www.es.acuvue.com/</a>
7. Johnson & Johnson	57	<a href="http://www.lifescanurope.com/es/">http://www.lifescanurope.com/es/</a>
8. Hewlett Packard	67	<a href="http://welcome.hp.com/country/es/es/welcome.html">http://welcome.hp.com/country/es/es/welcome.html</a>
9. Intel	74	<a href="http://www.intel.com/cd/corporate/europe/emea/spa/321188.htm">http://www.intel.com/cd/corporate/europe/emea/spa/321188.htm</a>
10. American Express	88	<a href="http://www.americanexpress.com/spain/homepage.shtml?location=globalsplash">http://www.americanexpress.com/spain/homepage.shtml?location=globalsplash</a>
11. Boeing	93	<a href="http://www.boeing.es/Home.do">http://www.boeing.es/Home.do</a>
12. Dow Chemical	106	<a href="http://www.dow.com/styrofoam/europe/es/index.htm">http://www.dow.com/styrofoam/europe/es/index.htm</a>
13. United Technologies	108	<a href="http://otis.com/otis/1,1352,CLI15_RES1,00.html">http://otis.com/otis/1,1352,CLI15_RES1,00.html</a>
14. United Parcel Service	109	<a href="http://www.ups.com/content/es/es/index.jsx">http://www.ups.com/content/es/es/index.jsx</a>
15. Cisco	117	<a href="http://www.cisco.com/global/ES/index.shtml">http://www.cisco.com/global/ES/index.shtml</a>
16. Merck & Co.	123	<a href="http://www.msd.es/content/corporate/index.html">http://www.msd.es/content/corporate/index.html</a>
17. Motorola	124	<a href="http://www.motorola.com/es/">http://www.motorola.com/es/</a>
18. Caterpillar	127	<a href="http://espana.cat.com/cda/layout?m=99205&amp;x=534">http://espana.cat.com/cda/layout?m=99205&amp;x=534</a>
19. Pepsico	128	<a href="http://www.pepsi.es/pepsihome.html">http://www.pepsi.es/pepsihome.html</a>
20. Ford Motor	133	<a href="http://www.ford.es/home.html">http://www.ford.es/home.html</a>
21. Dell	147	<a href="http://www1.euro.dell.com/content/default.aspx?c=es&amp;l=es&amp;s=gen">http://www1.euro.dell.com/content/default.aspx?c=es&amp;l=es&amp;s=gen</a>
22. Wyeth	154	<a href="http://www.wyeth.es/home.html">http://www.wyeth.es/home.html</a>
23. Abbott Laboratories	160	<a href="http://www.abbott.es">http://www.abbott.es</a>
24. EI duPont de Nemours	171	<a href="http://www2.dupont.com/Corian/es_ES/">http://www2.dupont.com/Corian/es_ES/</a>
25. Honeywell	183	<a href="http://www.honeywell.com/sites/es/">http://www.honeywell.com/sites/es/</a>
26. Bristol-Myers Squibb	192	<a href="http://www.bms.es/nportal/bms_corp/defaultpage.jsp?idtab=26009">http://www.bms.es/nportal/bms_corp/defaultpage.jsp?idtab=26009</a>
27. 3M	198	<a href="http://solutions.3m.com/wps/portal/3M/es_ES/EU/Country/">http://solutions.3m.com/wps/portal/3M/es_ES/EU/Country/</a>
28. Alcoa	212	<a href="http://www.alcoa.com/spain/es/home.asp">http://www.alcoa.com/spain/es/home.asp</a>
29. Eli Lilly & Co.	215	<a href="http://www.lilly.es/Nitro/">http://www.lilly.es/Nitro/</a>
30. FedEx	220	<a href="http://www.fedex.com/es/">http://www.fedex.com/es/</a>
31. Oracle	240	<a href="http://www.oracle.com/global/es/index.html">http://www.oracle.com/global/es/index.html</a>
32. Deere & Co.	241	<a href="http://www.deere.com/es_ES/index.html">http://www.deere.com/es_ES/index.html</a>
33. Cigna	250	<a href="http://www.cigna.es/index.asp">http://www.cigna.es/index.asp</a>
34. General Dynamics	267	<a href="http://www.gdsbs.com/web/frame.asp?lng=esp">http://www.gdsbs.com/web/frame.asp?lng=esp</a>
35. Medtronic	272	<a href="http://www.medtronic.es/ES/">http://www.medtronic.es/ES/</a>
36. Archer Daniels	277	<a href="http://www.admworld.com/eusp/">http://www.admworld.com/eusp/</a>

<b>Empresas Públicas</b>	<b>Posición en Lista Forbes mundial</b>	<b>Sitio web localizado hacia la <i>locale</i> «es-ES»</b>
37. Genworth Financial	280	<a href="http://www.gemieurope.com/es/4.html">http://www.gemieurope.com/es/4.html</a>
38. Apple Computer	283	<a href="http://www.apple.com/es/">http://www.apple.com/es/</a>
39. Kimberly-Clarck	303	<a href="http://www.kcprofessional.com/es/">http://www.kcprofessional.com/es/</a>
40. Atomatic Data	327	<a href="http://www.es-adp.com/pagina_inicio.cfm">http://www.es-adp.com/pagina_inicio.cfm</a>
41. Johnson Controls	343	<a href="http://www.johnsoncontrols.com/corp_profile/sp.htm">http://www.johnsoncontrols.com/corp_profile/sp.htm</a>
42. EMC	365	<a href="http://spain.emc.com/">http://spain.emc.com/</a>
43. Xerox	366	<a href="http://www.xerox.com/go/xrx/template/013.jsp?Xcntry=ESP&amp;Xlang=es_ES">http://www.xerox.com/go/xrx/template/013.jsp?Xcntry=ESP&amp;Xlang=es_ES</a>
44. Nike	408	<a href="http://nikeid.nike.com/nikeid/index.jhtml?_requestid=6363640#home*">http://nikeid.nike.com/nikeid/index.jhtml?_requestid=6363640#home*</a>
45. Aon	418	<a href="http://www.aon.com/es/es/">http://www.aon.com/es/es/</a>
46. Baxter International	420	<a href="http://www.baxter.es/default.html">http://www.baxter.es/default.html</a>
47. Cogate-Palmolive	425	<a href="http://www.colgate.es/index.shtml">http://www.colgate.es/index.shtml</a>
48. Google	439	<a href="http://www.google.es">http://www.google.es</a> *
49. Sara Lee	456	<a href="http://www.ambipur.es*">http://www.ambipur.es*</a>
50. Yahoo	473	<a href="http://es.yahoo.com/">http://es.yahoo.com/</a> *
51. Lucent Technologies	484	<a href="http://www.lucent.es/apps/emea/country/countrycenters.nsf">http://www.lucent.es/apps/emea/country/countrycenters.nsf</a>
52. Computer Sciences	499	<a href="http://es.country.csc.com/es/index.shtml">http://es.country.csc.com/es/index.shtml</a>
53. eBay	524	<a href="http://www.ebay.es">http://www.ebay.es</a>
54. General Motors	526	<a href="http://www.chevrolet.es/">http://www.chevrolet.es/</a>
55. Praxair	562	<a href="http://www.praxair.com/praxairSpain.nsf/">http://www.praxair.com/praxairSpain.nsf/</a>
56. Agilent Technologies	591	<a href="http://www.home.agilent.com/agilent/home.jsp?cc=ES&amp;lc=spa">http://www.home.agilent.com/agilent/home.jsp?cc=ES&amp;lc=spa</a>
57. Franklin Resources	606	<a href="http://www.franklintempleton.com.es/spain/jsp_cm/ft_home.jsp">http://www.franklintempleton.com.es/spain/jsp_cm/ft_home.jsp</a>
58. Boston Scientific	630	<a href="http://www.bostonscientific.es/">http://www.bostonscientific.es/</a>
59. McGraw-Hill Cos.	663	<a href="http://www.mcgraw-hill.es/*">http://www.mcgraw-hill.es/*</a>
60. Starwood Hotels	692	<a href="http://www.starwoodhotels.com/corporate/index.html?language=fr_FR&amp;IM=SW_CCTLD_FR&amp;localeoverride=">http://www.starwoodhotels.com/corporate/index.html?language=fr_FR&amp;IM=SW_CCTLD_FR&amp;localeoverride=</a>
61. Pepsi Bottling Group	695	<a href="http://www.pepsi.es">www.pepsi.es</a> *
62. Becton Dickinson	729	<a href="http://www.bdeurope.com/cgi-bin/bd/bdweb/eservices/content/show.bd?Program=C46A9D00341121A4C1256BD50038545C&amp;BV_UseBVCookie=yes&amp;BD_SID=UWtSRIV3PT06TUE9PT06*">http://www.bdeurope.com/cgi-bin/bd/bdweb/eservices/content/show.bd?Program=C46A9D00341121A4C1256BD50038545C&amp;BV_UseBVCookie=yes&amp;BD_SID=UWtSRIV3PT06TUE9PT06*</a>
63. Avon Products	742	<a href="http://www.avon.es/PRSuite/home/home">http://www.avon.es/PRSuite/home/home</a>
64. Harley-Davidson	750	<a href="http://www.harley-davidson.com/wcm/Content/Pages/home.jsp?locale=es_ES*">http://www.harley-davidson.com/wcm/Content/Pages/home.jsp?locale=es_ES*</a>
65. Whirlpool	775	<a href="http://www.whirlpool.es/webapp/wcs/stores/servlet/ProductDisplay?storeId=10301&amp;catalogId=10251&amp;langId=-5&amp;partNumber=sku-BB-WPHome1">http://www.whirlpool.es/webapp/wcs/stores/servlet/ProductDisplay?storeId=10301&amp;catalogId=10251&amp;langId=-5&amp;partNumber=sku-BB-WPHome1</a> *
66. Sun Microsystems	808	<a href="http://es.sun.com/">http://es.sun.com/</a>
67. Starbucks	810	<a href="http://www.starbucks.es/es-es/">http://www.starbucks.es/es-es/</a>
68. Stryker	811	<a href="http://www.stryker.es/st-index">http://www.stryker.es/st-index</a> *

<b>Empresas Públicas</b>	<b>Posición en Lista Forbes mundial</b>	<b>Sitio web localizado hacia la <i>locale</i> «es-ES»</b>
69. MBIA	817	<a href="http://www.mbia.com/sp/contact.html">http://www.mbia.com/sp/contact.html</a>
70. Office Depot	823	<a href="http://www.officedepot.es/">http://www.officedepot.es/</a>
71. Legg Mason	828	<a href="http://spain.citigroupam.com/retail/flash-home.jsp">http://spain.citigroupam.com/retail/flash-home.jsp</a>
72. Cummins	863	<a href="http://www.cummins.com/eu/pages/sp/distributors/spain/index.cfm">http://www.cummins.com/eu/pages/sp/distributors/spain/index.cfm</a>
73. Eastman Kodak	879	<a href="http://www.kodak.com/eknec/PageQuerier.jhtml?pq-locale=es_ES&amp;pq-path=2/6868&amp;_requestid=4819">http://www.kodak.com/eknec/PageQuerier.jhtml?pq-locale=es_ES&amp;pq-path=2/6868&amp;_requestid=4819</a> <i>.NO ESTA EN CORPUS*</i>
74. AMR	902	<a href="http://www.americanairlines.es/?url=www.americanairlines.es">http://www.americanairlines.es/?url=www.americanairlines.es</a>
75. Black & Decker	914	<a href="http://www.blackanddecker.es/index.asp?mktid=5">http://www.blackanddecker.es/index.asp?mktid=5</a>
76. Rockwell Automation	918	<a href="http://www.rockwellautomation.es/">http://www.rockwellautomation.es/</a>
77. Symantec	948	<a href="http://www.symantec.com/es/es/index.jsp">http://www.symantec.com/es/es/index.jsp</a>
78. Goodyear	951	<a href="http://es.goodyear.com/">http://es.goodyear.com/</a>
79. ProLogis	956	<a href="http://www.prologis.com/Spain/ES/">http://www.prologis.com/Spain/ES/</a>
80. CA	959	<a href="http://www.ca.com/es/">http://www.ca.com/es/</a>
81. Avaya	973	<a href="http://www.avaya.es/">http://www.avaya.es/</a>
82. Advanced Micro	979	<a href="http://www.amd.com/es-es/">http://www.amd.com/es-es/</a>
83. ManPower	114	<a href="http://www.manpower.es/Pub/content.asp?ref=Spain&amp;noderef=home&amp;langid=es*">http://www.manpower.es/Pub/content.asp?ref=Spain&amp;noderef=home&amp;langid=es*</a>
84. Adobe Systems	1067	<a href="http://www.adobe.com/es/">http://www.adobe.com/es/</a>
85. Arrow Electronics	1078	<a href="http://www.arrowiberia.com/">http://www.arrowiberia.com/</a>
86. Ingram Micro	1082	<a href="http://es.ingrammicro.com/corporate/corporate.jsp">http://es.ingrammicro.com/corporate/corporate.jsp</a>
87. Estee Lauder	1084	<a href="http://www.esteeauder.es/home.tml?ngextredir=1">http://www.esteeauder.es/home.tml?ngextredir=1</a>
88. Delta Airlines	1091	<a href="http://es.delta.com/delta/enes/?24;http://www.delta.com/home/index.jsp*">http://es.delta.com/delta/enes/?24;http://www.delta.com/home/index.jsp*</a>
89. Mattel	1097	<a href="http://service.mattel.com/es/default.asp">http://service.mattel.com/es/default.asp</a>
90. Moody's	1106	<a href="http://www.moody.com/espana/mdcsPage.aspx?template=about&amp;mdcsId=7">http://www.moody.com/espana/mdcsPage.aspx?template=about&amp;mdcsId=7</a>
91. Goodrich	1129	<a href="http://es.bfgoodrichtires.com/">http://es.bfgoodrichtires.com/</a>
92. Ecolab	1140	<a href="http://www.ecolab.com/es/">http://www.ecolab.com/es/</a>
93. Lexmark International	1181	<a href="http://www.lexmark.es/uncomplicate/site/home/0,,256618_0_0_es,00.html">http://www.lexmark.es/uncomplicate/site/home/0,,256618_0_0_es,00.html</a>
94. Interpublic Group	1188	<a href="http://www.webershandwick.es/">http://www.webershandwick.es/</a>
95. Electronic Arts	1237	<a href="http://www.es.ea.com/">http://www.es.ea.com/</a>
96. Allergan	1251	<a href="http://www.allergan.es/">http://www.allergan.es/</a>
97. CH Robinson Worldwide	1255	<a href="http://www.robinsoneurope.com/robinson_spain/default.asp">http://www.robinsoneurope.com/robinson_spain/default.asp</a>
98. Continental Airlines	1282	<a href="http://www.continental.com/web/es-ES/default.aspx?LangCode=es&amp;POS=ES&amp;SID=EF459FB184E341E7884F5754A24F970A">http://www.continental.com/web/es-ES/default.aspx?LangCode=es&amp;POS=ES&amp;SID=EF459FB184E341E7884F5754A24F970A</a>
99. Sealed Air	1361	<a href="http://www.sealedair.com/eu/es/index.htm">http://www.sealedair.com/eu/es/index.htm</a>
100. Polo Ralph Lauren	1377	<a href="http://ralphlauren.com.es/global/V2/default.asp?cid=16&amp;langid=8&amp;ab=fd_spain_interior*">http://ralphlauren.com.es/global/V2/default.asp?cid=16&amp;langid=8&amp;ab=fd_spain_interior*</a>
101. Brown-Forman	1394	<a href="http://www.jackdaniels.com/age.asp?language=Spanish">http://www.jackdaniels.com/age.asp?language=Spanish</a> *
102. Dana	1406	<a href="http://www.glaser.es/es/glaser_index.php">http://www.glaser.es/es/glaser_index.php</a>

<b>Empresas Públicas</b>	<b>Posición en Lista Forbes mundial</b>	<b>Sitio web localizado hacia la <i>locale</i> «es-ES»</b>
103.Network Appliance	1412	<a href="http://www-es.netapp.com/">http://www-es.netapp.com/</a>
104.Hartman International	1426	<a href="http://www.becker.de/beckerCC21/www_root/index.jsp?language=Es">http://www.becker.de/beckerCC21/www_root/index.jsp?language=Es *</a>
105.Autodesk	1462	<a href="http://www.autodesk.es/adsk/servlet/home?siteID=455755&amp;id=458320">http://www.autodesk.es/adsk/servlet/home?siteID=455755&amp;id=458320</a>
106.Nvidia	1494	<a href="http://es.nvidia.com/page/home.html">http://es.nvidia.com/page/home.html</a>
107.Verisign	1504	<a href="http://www.verisign.es/">http://www.verisign.es/</a>
108.Hospira	1621	<a href="http://www.hospira.es/spanish/default.aspx">http://www.hospira.es/spanish/default.aspx</a>
109.Stanley Works	1627	<a href="http://www.stanleyworks.com.es/">http://www.stanleyworks.com.es/</a>

<b>Empresas privadas</b>	<b>Puesto en lista en EEUU</b>	
110.Cargill	1	<a href="http://www.agribands.com/launch.aspx?SID=b487664a-7222-45ef-8417-b15fc80cee45&amp;PID=8604191c-9c2b-4585-998b-a2d0dbb071cd">http://www.agribands.com/launch.aspx?SID=b487664a-7222-45ef-8417-b15fc80cee45&amp;PID=8604191c-9c2b-4585-998b-a2d0dbb071cd</a>
111.Pricewaterhouse Co	4	<a href="http://www.pwc.com/es/esp/main/home/index.html">http://www.pwc.com/es/esp/main/home/index.html</a>
112.Ernst & Young	7	<a href="http://www.ey.com/global/content.nsf/Spain/Home">http://www.ey.com/global/content.nsf/Spain/Home</a>
113.Toys “R” Us	13	<a href="http://www.toysrus.es/home.htm">http://www.toysrus.es/home.htm*</a>
114.SC Johnson & Son	24	<a href="http://www.johnsonwax.es">http://www.johnsonwax.es</a>
115.Gardian Industries	36	<a href="http://www.guardian.com.es/index.asp">http://www.guardian.com.es/index.asp</a>

## 14.2. Apéndice II. *Stop words* usadas en el programa Wordsmith Tools

- 75 primeras palabras forma del corpus CREA y las letras del alfabeto

de, la, que, el, y, en, a, los, se, del, un, las, no, por, con, una, su, para, al, lo, como, más, mas, o, por, pero, sin, embargo, sus, le, me, ha, si, sin, ya, sobre, este, cuando, todo, entre, esta, tambien, dos, mi, qué, muy, hasta, yo, porque, desde, sólo, vez, ni, así, él, te, me, le, se, les, un, unos, una, unos, dónde, eso, a, b, c, d, e, f, g, h, i, j, k, l, m, n, ñ, o, p, q, r, s, t, u, v, w, x, y, z

- 70 primeras palabras forma de la lista de la sección original de nuestro corpus

y, en, el, o, ni, a, tras, al, mi, mis, tu, tus, su, sus, the, lo, sin, además, me, te, se, nos, le, les, el, lo, la, los, las, aquí, allí, qué, cómo, cuánto, porqué, dónde, que, como, cuanto, cuantos, cuantas, cuanta, donde, cual, cuales, quién, quien, quiénes, quienes, mi, ti, él, tú, ella, vosotros, ellos, ellas, nosotros, tan, tanto, tantas, tantos, tanta, esto, eso, aquello, también, tampoco, cada, .es

- Lista de eventos y etiquetas más comunes en html

title, moved, img, source, target, not, found, document, á, é, í, ó, ú, page, error, sans, serif, arial, tahoma, font, family, nbsp, alt, width, height, border, gif, src, web, align, html, absmiddle, interfaces, class, corp, bytes, clean, corpora, nat, tip, com, ed, chk, mem, shared, images, name, type, hidden, nh, clear, dot, style, middle, align, vertical, value, class, com, td, gensearch, ncamefrom, headline, mcfs, btng, id, input, jpg, div, onclick, return, validatesearchfor, hd, href, self

- Palabras del inglés usadas en la investigación de Pérez (2002) (del Oxford Advanced Learner's Dictionary y British National Corpus).

i'm, i've, i'd, a, an, able, about, above, across, after, again, against, age, ago, almost, along, already, also, alwaysamong, an, and, another, an, and, another, answer, any, anything, are, area, around, as, at, away, back, be, became, bacause, become, before, began, behind, better, between, body, both, but, by, can, cannot, certain, could, did, didn'tdoes, don't, down, during, each, either, enough, even, ever, every, everything, for, from, got, had, has, have, he, her, him, himself, his, how, however, i'll, if, in, into, it, it's, its, is, been, just, less, like, many, may, me, more, most, much, my, never, next, no, not, nothing, now, of, off, often, on, only, or, our, out, outside, over, own, with, or, of, at, quite, rather, same, shall, she, should, since, so, some, something, sometimes, still, such, than, that, the, their, them, then, they, these, thing, things, this, those, though, through, to, together, too, top, under, until, up, upon, us, very, was, were, we, what, when, where, whether, which, while, who, whole, with, within, wihtout, yes, yet, you, your, in, or, on, no, of



### 14.3. Apéndice III. Listas de frecuencia de ambas secciones del corpus lematizadas y sin lematizar

Sección original. Lista de Frecuencia. SIN LEMATIZAR				Sección localizada. Lista de Frecuencia. SIN LEMATIZAR			
N	word	Freq.	%	N	word	Freq.	%
1	com	17.309	0,34	1	información	40.364	0,47
2	es	16.088	0,32	2	productos	33.710	0,39
3	información	12.829	0,26	3	ford	28.873	0,34
4	servicios	11.781	0,23	4	es	27.757	0,32
5	productos	11.011	0,22	5	nuevo	27.429	0,32
6	euros	9.567	0,19	6	hp	27.125	0,32
7	tv	9.395	0,19	7	vehículos	24.790	0,29
8	precio	9.174	0,18	8	impresión	21.305	0,25
9	madrid	8.462	0,17	9	servicios	21.161	0,25
10	web	8.368	0,17	10	lexmark	20.968	0,24
11	internet	7.589	0,15	11	todos	20.495	0,24
12	grupo	7.555	0,15	12	españa	17.159	0,2
13	empresa	6.994	0,14	13	servicio	16.497	0,19
14	rec	6.079	0,12	14	web	15.618	0,18
15	ono	6.057	0,12	15	impresoras	15.065	0,18
16	españa	5.856	0,12	16	soluciones	14.886	0,17
17	inicio	5.623	0,11	17	uso	14.563	0,17
18	todos	5.560	0,11	18	configurador	14.544	0,17
19	franquicia	5.448	0,11	19	página	14.465	0,17
20	datos	5.379	0,11	20	privacidad	14.090	0,16
21	servicio	5.228	0,1	21	soporte	13.676	0,16
22	alta	5.020	0,1	22	producto	13.136	0,15
23	empresas	4.848	0,1	23	noticias	12.924	0,15
24	seguridad	4.844	0,1	24	sitio	12.740	0,15
25	red	4.787	0,1	25	condiciones	12.529	0,15
26	teléfono	4.752	0,09	26	comerciales	12.449	0,15
27	viajes	4.711	0,09	27	dell	12.349	0,14
28	noticias	4.608	0,09	28	empresas	12.338	0,14
29	legal	4.561	0,09	29	seguridad	12.073	0,14
30	catálogo	4.506	0,09	30	focus	11.881	0,14
31	contacto	4.378	0,09	31	vehículo	11.454	0,13
32	base	4.377	0,09	32	color	11.445	0,13
33	clientes	4.364	0,09	33	max	11.394	0,13
34	está	4.326	0,09	34	ver	11.260	0,13
35	ver	4.325	0,09	35	gama	10.918	0,13
36	producto	4.281	0,09	36	consumibles	10.890	0,13
37	cliente	4.147	0,08	37	datos	10.780	0,13
38	venta	4.048	0,08	38	prueba	10.633	0,12
39	tiendas	4.027	0,08	39	sistema	10.390	0,12
40	caja	4.007	0,08	40	puede	10.366	0,12
41	sistemas	3.955	0,08	41	emc	9.864	0,11
42	sistema	3.952	0,08	42	términos	9.635	0,11
43	buscar	3.925	0,08	43	software	9.456	0,11
44	digital	3.922	0,08	44	accesorios	9.340	0,11
45	electrónico	3.839	0,08	45	nueva	9.272	0,11
46	www	3.797	0,08	46	com	9.155	0,11

Sección original. Lista de Frecuencia. SIN LEMATIZAR				Sección localizada. Lista de Frecuencia. SIN LEMATIZAR			
47	personal	3.785	0,08	47	copyright	9.060	0,11
48	instalación	3.781	0,08	48	papel	8.783	0,1
49	inmobiliaria	3.728	0,07	49	fiesta	8.735	0,1
50	cm	3.711	0,07	50	derechos	8.734	0,1
51	medidas	3.696	0,07	51	catálogo	8.689	0,1
52	centros	3.680	0,07	52	ups	8.669	0,1
53	incluido	3.674	0,07	53	especificaciones	8.660	0,1
54	total	3.651	0,07	54	site	8.544	0,1
55	coste	3.618	0,07	55	comprar	8.413	0,1
56	iva	3.569	0,07	56	rendimiento	8.342	0,1
57	puede	3.556	0,07	57	red	8.117	0,09
58	fi	3.555	0,07	58	clientes	7.977	0,09
59	son	3.549	0,07	59	general	7.935	0,09
60	final	3.541	0,07	60	sistemas	7.844	0,09
61	acceso	3.516	0,07	61	tecnología	7.733	0,09
62	ampliar	3.434	0,07	62	está	7.707	0,09
63	otros	3.414	0,07	63	acerca	7.642	0,09
64	sicav	3.392	0,07	64	cliente	7.634	0,09
65	mapa	3.363	0,07	65	gran	7.604	0,09
66	inversión	3.323	0,07	66	política	7.541	0,09
67	cualquier	3.306	0,07	67	home	7.504	0,09
68	nuestro	3.250	0,06	68	calidad	7.468	0,09
69	reciclado	3.235	0,06	69	aplicaciones	7.440	0,09
70	características	3.231	0,06	70	financiación	7.436	0,09
71	formación	3.196	0,06	71	empresa	7.213	0,08
72	control	3.173	0,06	72	cualquier	7.149	0,08
73	solar	3.166	0,06	73	son	7.117	0,08
74	franquicias	3.157	0,06	74	preguntas	7.113	0,08
75	mayo	3.153	0,06	75	contactar	7.087	0,08
76	café	3.139	0,06	76	nuestros	7.079	0,08
77	prensa	3.110	0,06	77	usuarios	7.079	0,08
78	tarjeta	3.105	0,06	78	imprimir	7.073	0,08
79	gestión	3.088	0,06	79	reservados	7.042	0,08
80	nuestros	3.066	0,06	80	trabajo	6.774	0,08
81	comercio	3.036	0,06	81	versión	6.754	0,08
82	calidad	3.034	0,06	82	windows	6.723	0,08
83	salud	2.962	0,06	83	diseño	6.696	0,08
84	derechos	2.902	0,06	84	hojas	6.670	0,08
85	imagen	2.891	0,06	85	frecuentes	6.664	0,08
86	sector	2.873	0,06	86	otros	6.623	0,08
87	usuario	2.868	0,06	87	buscar	6.591	0,08
88	microondas	2.858	0,06	88	volver	6.557	0,08
89	online	2.854	0,06	89	inglés	6.535	0,08
90	soporte	2.849	0,06	90	solicita	6.506	0,08
91	somos	2.844	0,06	91	descargas	6.500	0,08
92	televisión	2.830	0,06	92	contacto	6.343	0,07
93	aviso	2.818	0,06	93	sun	6.342	0,07
94	internacional	2.813	0,06	94	páginas	6.257	0,07
95	casa	2.785	0,06	95	programa	6.257	0,07
96	tarjetas	2.773	0,06	96	especiales	6.241	0,07
97	página	2.771	0,06	97	adobe	6.215	0,07

Sección original. Lista de Frecuencia. SIN LEMATIZAR				Sección localizada. Lista de Frecuencia. SIN LEMATIZAR			
98	etc	2.740	0,05	98	impresora	6.195	0,07
99	varios	2.738	0,05	99	herramientas	6.185	0,07
100	cine	2.735	0,05	100	mundo	6.154	0,07

- Listas de frecuencias léxicas lematizadas

Sección original. Lista de Frecuencia. LISTA LEMATIZADA				Sección localizada. Lista de Frecuencia. LISTA LEMATIZADA			
N	word	Freq.	%	N	word	Freq.	%
1	com	17.309	0,34	1	ser	50.637	0,59
2	servicios	17.017	0,34	2	productos	46.846	0,55
3	productos	15.292	0,3	3	nuevos	43.745	0,51
4	ser	14.022	0,28	4	información	40.656	0,47
5	información	13.068	0,26	5	servicios	37.658	0,44
6	empresa	11.842	0,24	6	vehículos	36.244	0,42
7	precio	10.450	0,21	7	ford	28.873	0,34
8	euros	10.236	0,2	8	todos	27.472	0,32
9	nuestro	9.839	0,2	9	hp	27.125	0,32
10	tv	9.395	0,19	10	poder	23.427	0,27
11	todos	9.002	0,18	11	impresión	21.305	0,25
12	estar	8.907	0,18	12	impresoras	21.260	0,25
13	web	8.707	0,17	13	lexmark	20.968	0,24
14	franquicia	8.605	0,17	14	página	20.958	0,24
15	clientes	8.518	0,17	15	nuestros	20.289	0,24
16	grupo	8.498	0,17	16	empresas	19.551	0,23
17	madrid	8.462	0,17	17	soluciones	18.765	0,22
18	nuevo	8.178	0,16	18	sistemas	18.234	0,21
19	poder	7.978	0,16	19	españa	17.159	0,2
20	sistemas	7.907	0,16	20	estar	17.040	0,2
21	internet	7.589	0,15	21	comercial	16.004	0,19
22	otros	7.493	0,15	22	web	15.894	0,19
23	alta	7.312	0,15	23	cliente	15.611	0,18
24	tiendas	6.649	0,13	24	uso	14.886	0,17
25	rec	6.079	0,12	25	otros	14.840	0,17
26	ono	6.057	0,12	26	sitio	14.735	0,17
27	tarjeta	5.878	0,12	27	configurador	14.545	0,17
28	españa	5.856	0,12	28	color	14.367	0,17
29	centros	5.791	0,12	29	soporte	14.095	0,16
30	catálogo	5.787	0,12	30	privacidad	14.090	0,16
31	red	5.676	0,11	31	contacte	13.843	0,16
32	inicio	5.623	0,11	32	trabajar	13.212	0,15
33	viajes	5.457	0,11	33	noticias	13.108	0,15
34	datos	5.429	0,11	34	ver	12.910	0,15
35	inmobiliarias	5.408	0,11	35	prueba	12.715	0,15
36	electrónico	5.317	0,11	36	condiciones	12.703	0,15
37	medidas	5.182	0,1	37	dell	12.349	0,14

Sección original. Lista de Frecuencia. LISTA LEMATIZADA				Sección localizada. Lista de Frecuencia. LISTA LEMATIZADA			
38	teléfono	5.152	0,1	38	usuario	12.244	0,14
39	venta	5.014	0,1	39	comprar	12.102	0,14
40	año	4.977	0,1	40	seguridad	12.074	0,14
41	noticias	4.902	0,1	41	focus	11.881	0,14
42	seguridad	4.844	0,1	42	max	11.394	0,13
43	tener	4.791	0,1	43	solicita	11.066	0,13
44	base	4.789	0,1	44	gama	10.957	0,13
45	solar	4.768	0,09	45	datos	10.891	0,13
46	legal	4.736	0,09	46	consumibles	10.890	0,13
47	ver	4.713	0,09	47	catálogo	10.489	0,12
48	contacto	4.579	0,09	48	papel	10.290	0,12
49	digital	4.521	0,09	49	red	10.002	0,12
50	trabajar	4.509	0,09	50	general	9.893	0,12
51	caja	4.359	0,09	51	términos	9.871	0,12
52	ampliar	4.332	0,09	52	emc	9.864	0,11
53	haber	4.307	0,09	53	accesorios	9.771	0,11
54	buscar	4.233	0,08	54	software	9.456	0,11
55	personal	4.171	0,08	55	ofrecer	9.432	0,11
56	quién	4.123	0,08	56	aplicación	9.362	0,11
57	contacte	4.105	0,08	57	com	9.155	0,11
58	sector	4.091	0,08	58	mejorar	9.151	0,11
59	total	4.069	0,08	59	copyright	9.060	0,11
60	características	3.949	0,08	60	site	9.018	0,11
61	bajo	3.941	0,08	61	fiesta	8.735	0,1
62	coste	3.934	0,08	62	derechos	8.734	0,1
63	incluido	3.917	0,08	63	política	8.729	0,1
64	usuario	3.906	0,08	64	alta	8.707	0,1
65	www	3.797	0,08	65	ups	8.669	0,1
66	instalación	3.781	0,08	66	especificaciones	8.660	0,1
67	acceso	3.748	0,07	67	hacer	8.603	0,1
68	casa	3.729	0,07	68	rendimiento	8.381	0,1
69	final	3.726	0,07	69	haber	8.338	0,1
70	internacional	3.639	0,07	70	imágenes	8.305	0,1
71	iva	3.569	0,07	71	programa	8.274	0,1
72	cualquier	3.565	0,07	72	personal	8.240	0,1
73	oficina	3.558	0,07	73	versión	7.959	0,09
74	fi	3.555	0,07	74	cualquier	7.954	0,09
75	zona	3.543	0,07	75	imprimir	7.936	0,09
76	imágenes	3.533	0,07	76	año	7.926	0,09
77	ofertas	3.455	0,07	77	preguntas	7.889	0,09
78	página	3.444	0,07	78	herramientas	7.884	0,09
79	comprar	3.424	0,07	79	contacto	7.787	0,09
80	comercio	3.422	0,07	80	técnicas	7.761	0,09
81	valor	3.419	0,07	81	tecnología	7.733	0,09
82	mapa	3.413	0,07	82	acerca	7.642	0,09
83	sicav	3.392	0,07	83	gran	7.604	0,09
84	mejorar	3.379	0,07	84	tener	7.555	0,09
85	hacer	3.345	0,07	85	ir	7.548	0,09
86	inversión	3.323	0,07	86	profesional	7.530	0,09
87	mercado	3.307	0,07	87	home	7.504	0,09
88	control	3.301	0,07	88	calidad	7.494	0,09

<b>Sección original. Lista de Frecuencia. LISTA LEMATIZADA</b>				<b>Sección localizada. Lista de Frecuencia. LISTA LEMATIZADA</b>			
89	reciclado	3.300	0,07	89	financiación	7.436	0,09
90	hora	3.279	0,07	90	hojas	7.416	0,09
91	negocio	3.207	0,06	91	digital	7.301	0,09
92	financieros	3.196	0,06	92	permitir	7.249	0,08
93	formación	3.196	0,06	93	especiales	7.166	0,08
94	laboral	3.195	0,06	94	diseño	7.156	0,08
95	tipo	3.162	0,06	95	volver	7.147	0,08
96	mayo	3.153	0,06	96	reservados	7.112	0,08
97	café	3.139	0,06	97	grande	7.109	0,08
98	prensa	3.132	0,06	98	buscar	6.947	0,08
99	gestión	3.129	0,06	99	frecuentes	6.880	0,08
100	soluciones	3.128	0,06	100	windows	6.723	0,08

#### 14.4. Apéndice IV. Lista de co-ocurrencia léxica en subcorpus de páginas de contacto

n	Sección original	freq.	%	n	Sección localizada	freq.	%
1	e-mail	55	0,46	1	correo electrónico	59	0,27
2	con nosotros	36	0,3	2	por favor	46	0,21
3	de contacto	31	0,26	3	con nosotros	45	0,21
4	todos los	26	0,22	4	al cliente	43	0,2
5	en contacto	25	0,21	5	de privacidad	43	0,2
6	si desea	25	0,21	6	e-mail	41	0,19
7	contacto	24	0,2	7	información de	40	0,18
8	contacto con	23	0,19	8	contactar con	37	0,17
9	de datos	21	0,17	9	de contacto	36	0,17
10	club nokia	20	0,17	10	contacto con	35	0,16
11	correo electrónico	20	0,17	11	del sitio	34	0,16
12	los derechos	20	0,17	12	contacte con	33	0,15
13	más información	20	0,17	13	atención al	29	0,13
14	contactar con	19	0,16	14	en contacto	29	0,13
15	datos de	18	0,15	15	mapa del	29	0,13
16	este formulario	18	0,15	16	productos y	29	0,13
17	formulario de	18	0,15	17	en españa	28	0,13
18	de publicidad	16	0,13	18	de soporte	27	0,12
19	departamento de	16	0,13	19	y servicios	27	0,12
20	el siguiente	16	0,13	20	privacidad	26	0,12
21	fax #	16	0,13	21	de prensa	25	0,11
22	la página	16	0,13	22	más información	25	0,11
23	los datos	16	0,13	23	de correo	24	0,11
24	siguiente formulario	16	0,13	24	de productos	24	0,11
25	letreros de	14	0,12	25	declaración de	24	0,11
26	paneles electronicos	14	0,12	26	de uso	23	0,11
27	protección de	14	0,12	27	información sobre	23	0,11
28	rotulos luminosos	14	0,12	28	política de	23	0,11
29	código postal	13	0,11	29	soporte técnico	23	0,11
30	com departamento	13	0,11	30	asistencia técnica	22	0,1
31	de información	13	0,11	31	código postal	22	0,1
32	derechos reservados	13	0,11	32	más detalles	21	0,1
33	nf #	13	0,11	33	póngase en	21	0,1
34	acerca de	12	0,1	34	dirección de	20	0,09
35	la empresa	12	0,1	35	en línea	20	0,09
36	los campos	12	0,1	36	todos los	20	0,09
37	aviso legal	11	0,09	37	de atención	18	0,08
38	copyright ©	11	0,09	38	copyright ©	17	0,08
39	de carácter	11	0,09	39	mail teléfono	17	0,08
40	electrónicos reset	11	0,09	40	artes gráficas	16	0,07
41	laboratorios electrónicos	11	0,09	41	derechos reservados	16	0,07
42	ponerse en	11	0,09	42	los derechos	16	0,07
43	por favor	11	0,09	43	para imprimir	16	0,07
44	comercio electrónico	10	0,08	44	relations #	16	0,07
45	contacte con	10	0,08	45	servicio de	16	0,07
46	de privacidad	10	0,08	46	servicios de	16	0,07
47	dirección de	10	0,08	47	contacta con	15	0,07
48	la dirección	10	0,08	48	contáctenos	15	0,07
49	los servicios	10	0,08	49	departamento de	15	0,07
50	servicios de	10	0,08	50	hewlett-packard	15	0,07
51	campos obligatorios	9	0,07	51	la información	15	0,07
52	carácter personal	9	0,07	52	los productos	15	0,07
53	datos personales	9	0,07	53	un producto	15	0,07
54	de atención	9	0,07	54	versión para	15	0,07
55	del sitio	9	0,07	55	analyst relations	14	0,06
56	diseño web	9	0,07	56	corporate analyst	14	0,06
57	información sobre	9	0,07	57	de acceso	14	0,06
58	mapa del	9	0,07	58	de baterías	14	0,06
59	nokia service	9	0,07	59	de comunicación	14	0,06
60	rellene el	9	0,07	60	de información	14	0,06