

# **La localización de hipertextos: el género y la tipología textual en los sitios web corporativos**



**AUTOR: MIGUEL A. JIMÉNEZ CRESPO  
DIRECCIÓN: DRA. TERCEDOR SANCHEZ**

**DOCTORADO EN PROCESOS DE TRADUCCIÓN  
DEPARTAMENTO DE TRADUCCIÓN E INTERPRETACIÓN  
UNIVERSIDAD DE GRANADA  
PROYECTO DE INVESTIGACIÓN TUTELADA, 2005-2006**

## ÍNDICE

<b>1. INTRODUCCIÓN .....</b>	<b>5</b>
1.1. PRESENTACIÓN .....	5
1.2. HIPÓTESIS DE TRABAJO Y OBJETIVOS.....	7
<b>2. EL PROCESO DE LOCALIZACIÓN.....</b>	<b>9</b>
2.1. GLOBALIZACIÓN E INTERNACIONALIZACIÓN .....	9
2.2. LOCALIZACIÓN Y LOCALE .....	12
2.3. EMISORES Y RECEPTORES EN LOCALIZACIÓN .....	15
2.4. LOCALIZACIÓN Y SU COMPONENTE CULTURAL.....	16
2.5. DIFERENTES GRADOS DE LOCALIZACIÓN .....	18
2.6. ESTUDIOS SOBRE LA LOCALIZACIÓN .....	22
2.6.1. ESTUDIOS EN LOCALIZACIÓN DESDE LA PERSPECTIVA TRADUCTOLÓGICA .....	23
2.6.2. ESTUDIOS DESDE LA INDUSTRIA.....	24
<b>3. LA LOCALIZACIÓN DE SITIOS WEB.....</b>	<b>26</b>
3.1. INVESTIGACIÓN PREVIA SOBRE LA LOCALIZACIÓN DE SITIOS WEB .....	28
3.2. SITIOS WEB CORPORATIVOS COMO OBJETIVO DEL ESTUDIO .....	30
3.3. IMPORTANCIA DE UN ESTUDIO SOBRE LA TEXTUALIDAD DE LOS SITIOS WEB .....	32
3.4. TEXTUALIDAD DE SOFTWARE Y DE SITIOS WEB.....	34
<b>4. CONCEPTOS FUNDAMENTALES EN LA INVESTIGACIÓN: EL GÉNERO Y LA TIPOLOGÍA TEXTUAL. ....</b>	<b>38</b>
4.1. TIPOLOGÍA TEXTUAL Y GÉNERO.....	39
4.1.1. Tipologías textuales en traductología .....	41
4.2. LOS GÉNEROS DIGITALES O CIBERGÉNEROS.....	42
4.2.1. Evolución del concepto de género.....	43
4.2.2. El Modelo de Género de Silvia Gamero (1998, 2001) .....	46
4.2.2.1. Rasgos convencionales.....	47
4.2.2.2. Función textual .....	49
4.2.2.3. Elementos de la situación comunicativa .....	52
4.2.2.4. Influencia del contexto sociocultural .....	53
4.2.2.5. Elementos intratextuales .....	54

4.2.2.6.	Funcionalidad .....	54
4.2.2.6.1.	La funcionalidad como criterio clasificador de géneros digitales .....	56
4.2.2.7.	Cibergéneros en la web: una constante evolución.....	58
4.3.	PROPUESTA DE CARACTERIZACIÓN GENERAL DE GÉNERO WEB DE SITIO WEB CORPORATIVO	60
<b>5.</b>	<b>EL ENFOQUE TEXTUAL ¿QUÉ CONSTITUYE UN TEXTO EN UN SITIO WEB O EN UN PROGRAMA DE SOFTWARE? .....</b>	<b>62</b>
5.1.	DE TEXTO A HIPERTEXTO.....	62
5.2.	DEFINICIÓN DE TEXTO .....	63
5.3.	¿QUÉ CONSTITUYE UN TEXTO EN LOCALIZACIÓN?.....	64
5.4.	EL HIPERTEXTO .....	65
5.5.	TIPOS DE HIPERTEXTOS .....	67
5.6.	CARACTERÍSTICAS ESPECÍFICAS DE LOS HIPERTEXTOS .....	69
5.7.	COHERENCIA Y COHESIÓN .....	71
5.8.	EL SITIO WEB COMO TEXTO .....	73
5.9.	EL CONTEXTO COMO VARIABLE NECESARIA EN LA DEFINICIÓN DEL TEXTO EN LOCALIZACIÓN	75
5.9.1.	La dimensión comunicativa del contexto .....	77
5.9.1.1.	Registro .....	78
5.9.1.2.	Modo y tenor digitales.....	79
5.9.1.3.	Registro y género.....	80
5.9.1.4.	El registro en Internet .....	81
<b>6.</b>	<b>PRESENTACIÓN DE LA INVESTIGACIÓN EMPÍRICA .....</b>	<b>83</b>
<b>7.</b>	<b>METODOLOGÍA DE LA INVESTIGACIÓN EMPÍRICA .....</b>	<b>85</b>
7.1.	HERRAMIENTAS DE ANÁLISIS .....	86
<b>8.</b>	<b>PRINCIPIOS TEÓRICOS Y METODOLÓGICOS DE LA RECOPIACIÓN DEL CORPUS.....</b>	<b>87</b>
8.1.	EL CORPUS.....	87
8.1.1.	DIFERENTES CLASES DE CORPUS. ....	88
8.2.	RECOPIACIÓN DEL CORPUS.....	90
8.2.1.	Criterios para la recopilación del corpus.....	91
8.3.	DESCARGA DEL CORPUS.....	94

8.4.	CRITERIOS PARA ASEGURAR QUE SE TRATA DE UN TEXTO ORIGINAL .....	95
8.5.	LIMITACIONES.....	97
8.5.1.	Los marcos y la arquitectura servidor-usuario .....	98
8.5.2.	Macromedia Flash.....	99
8.5.3.	Sitios web eliminados por distorsionar las estadísticas.....	100
<b>9.</b>	<b>RESULTADOS DEL ANÁLISIS .....</b>	<b>101</b>
9.1.	COMPOSICIÓN FINAL DEL CORPUS .....	101
9.1.1.	Sitios web localizados del español a otras lenguas en el corpus .....	102
9.2.	CARACTERIZACIÓN DEL SITIO WEB CORPORATIVO DESDE UNA PERSPECTIVA TRADUCTOLÓGICA .....	103
9.2.1.	Diferentes clases de páginas.....	104
9.2.1.1.	Las páginas web de Splash.....	104
9.2.1.2.	Página principal.....	105
9.2.1.3.	Páginas interiores .....	106
9.2.1.4.	Las páginas basadas en la interacción por la arquitectura servidor-usuario 107	
9.2.2.	Elementos textuales a traducir en las páginas web. ....	107
9.2.2.1.	Los títulos .....	108
9.2.2.2.	Los elementos de navegación .....	108
9.2.2.3.	Pies de página .....	109
9.2.2.4.	Los segmentos textuales en eventos Html .....	109
9.2.2.4.1.	Los textos en etiquetas Meta.....	109
9.2.2.4.2.	Los segmentos textuales Alt.....	110
9.2.2.4.3.	Los segmentos textuales en OnMouse.....	110
9.2.2.4.4.	Segmentos textuales en scripts y <Select> .....	111
9.2.2.5.	Texto en el cuerpo de la página .....	112
9.2.2.6.	Elementos multimedia .....	113
9.3.	ANÁLISIS DE LA SUPERESTRUCTURA: BLOQUES Y SECCIONES EN LOS SITIOS WEB CORPORATIVOS.....	114
9.3.1.	DESCRIPCIÓN DEL BLOQUE «CONTACTO» .....	120
9.3.1.1.	Elementos de la situación comunicativa .....	120
9.3.1.2.	Función textual .....	120

9.3.1.3.	Rasgos convencionales .....	121
9.3.1.4.	Funcionalidad .....	124
9.4.	DESCRIPCIÓN DEL BLOQUE « LA EMPRESA».....	124
9.4.1.	Elementos de la situación comunicativa.....	124
9.4.2.	Función textual.....	124
9.4.3.	Rasgos convencionales .....	125
9.4.4.	Funcionalidad.....	129
9.4.5.	DESCRIPCIÓN DEL BLOQUE «NOTICIAS» .....	129
9.4.5.1.	Elementos de la situación comunicativa .....	129
9.4.5.2.	Función textual .....	129
9.4.5.3.	Rasgos convencionales .....	129
9.4.5.4.	Funcionalidad .....	130
9.4.6.	DESCRIPCIÓN DEL BLOQUE «PRODUCTOS-SERVICIOS» .....	131
9.4.6.1.	Elementos de la situación comunicativa .....	131
9.4.6.2.	Función textual .....	131
9.4.6.3.	Rasgos convencionales .....	131
9.4.6.4.	Funcionalidad .....	133
9.4.7.	DESCRIPCIÓN DEL BLOQUE «LEGAL» .....	134
9.4.7.1.	Elementos de la situación comunicativa .....	134
9.4.7.2.	Función textual .....	135
9.4.7.3.	Rasgos convencionales .....	135
9.4.7.4.	Funcionalidad .....	137
9.4.8.	DESCRIPCIÓN DEL BLOQUE «ZONA RESERVADA-ESPECIAL».....	137
9.4.9.	DESCRIPCIÓN DEL BLOQUE «INTERACTIVIDAD» .....	140
9.4.9.1.	Elementos de la situación comunicativa .....	140
9.4.9.2.	Función textual .....	140
9.4.9.3.	Rasgos convencionales .....	141
9.4.9.4.	Funcionalidad .....	143
<b>10.</b>	<b>CONCLUSIÓN .....</b>	<b>144</b>
<b>12.</b>	<b>BIBLIOGRAFÍA .....</b>	<b>151</b>
<b>13.</b>	<b>APÉNCIDE I. COMPOSICIÓN DE CORPUS .....</b>	<b>167</b>
<b>14.</b>	<b>APÉNDICE II. FIGURAS Y TABLAS.....</b>	<b>176</b>

# 1. INTRODUCCIÓN

*Plain text is the foundation of most web information [...]*  
(Nielsen y Tahir, 2002: 48)

## 1.1. Presentación

Durante los últimos años hemos sido testigos del desarrollo exponencial del campo de la localización, un fenómeno que en la actualidad reclama su estudio pormenorizado desde una perspectiva traductológica (Pym, 2004; Quirion, 2003). Este proceso, en el que la traducción representa un eslabón fundamental (Esselink, 2001), posee unas características propias que surgen por la existencia de un proceso de producción y distribución textual muy diferente a otros campos de la traducción como la jurídica, técnica, la literaria o el doblaje y substitulación<sup>1</sup>.

La localización web o *e-localización* (Cronin, 2003) se desarrolló adaptando y modelando procesos que ya existían en la localización de software (Yunker, 2003), habiendo sobrepasado en importancia económica a la misma (LISA, 2004). Con la creación de la web se extendió el uso de los hipertextos, que se han convertido en una forma básica organización textual en el mundo actual. El impacto de la hipertextualidad y los nuevos modelos de comunicación (Janoschka, 2003) en la labor de los traductores enorme. De hecho, la invención de Internet y su importancia como medio de distribución de información se considera la invención más importante en comunicación desde que Guttemberg creara la imprenta en el siglo XV (Lockwood y Scott, 1999:9).

El presente estudio se ha fraguado tras varios años de labor localizadora de sitios web del autor<sup>2</sup>, matizado y corregido a través de la dirección y comentarios que se produjeron en las presentaciones de un estudio preliminar en varias conferencias (Jiménez, 2005)<sup>3</sup>. En especial, el centro del presente estudio se basa

---

<sup>1</sup> La localización se suele considerar como un ramo de la traducción audiovisual, que incluye además el doblaje y substitulación.

<sup>2</sup> El autor está certificado por la American Translator Association en la combinación EN>ES

<sup>3</sup> «Las peculiaridades textuales de las páginas web coporativas localizadas al español». Presentado en la sesión sobre localización de la Conferencia MFLIC en Wake Forest University, EEUU el 12 de Octubre del 2005 y en la Conferencia Anual de la ATA (*American Translator Association*) en Seattle, 11 de Noviembre del 2005.

en la introducción de varios conceptos interdisciplinarios en el análisis traductológico de la localización web: el grado de localización (Brooks, 2000), el hipertexto (Storrer, 2002), el concepto de género (Gamero, 2001), o la representatividad en la recopilación de corpora web (Biber, 2004).

Durante los últimos años la traductología ha ido incluyendo un marco de estudio interdisciplinario con 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 en las convenciones tipológicas, textuales, discursivas y estructurales (Trosborg, 1997). El género 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 los que la mezcla de tipologías textuales suele ser la norma (Santini, 2005a; Tercedor, 2005).

En la actualidad la mayoría de los estudios en localización parten desde la industria de la localización (Quirion, 2003). Ésta se basa en que los textos localizados deben recibirse en la cultura meta como si se produjeran originalmente allí: «*with the look and feel of locally made products*» (LISA, 2003:5), según nuestra perspectiva, que se ajusten a las convenciones tipológicas, discursivas y de género de la cultura meta. Por lo tanto, una de las cuestiones que se pretende dilucidar en el presente trabajo es ¿qué características tipológicas, discursivas, estructurales poseen los textos web producidos originalmente en España? ¿Cómo ha llegado la industria de la localización a establecerlas? ¿Cómo influye la localización como proceso económico en este proceso de traducción? La revisión de las investigaciones en este campo muestra que este proceso se suele producir desde la cultura norteamericana hacia la cultura meta, y no al contrario, en análisis que partan desde cada cultura o *locale* (Schäler, 2002). La hipótesis inicial se fundamenta en que un análisis comparativo que parta del concepto de género ayudará a dilucidar qué elementos discursivos, estructurales y culturales son prototípicos en los sitios web españoles y qué convenciones se han consolidado. Por medio de técnicas de lingüística de corpus pretendemos proporcionar un estudio que sea de utilidad al traductor en su labor localizadora.

## 1.2. Hipótesis de trabajo y objetivos

El presente trabajo parte de tres hipótesis de trabajo:

- i. En primer lugar, que el género impone unas restricciones a todos los niveles textuales que pueden ser diferentes en distintas lenguas y culturas. Los cibergéneros o género web presentarían igualmente estas restricciones y diferencias entre lenguas y culturas, que en este campo supondrían diferencias entre *locales*.
- ii. En segundo lugar argumentamos que el proceso de localización impone unas restricciones en la labor traductológica, ya que el acceso al texto del traductor es diferenciado y el grado de localización de encargo influirá de igual manera en el proceso.
- 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 por cuanto a la superestructura, macroestructura, elementos intratextuales o funcionalidad que el traductor debe respetar para producir un texto plenamente funcional en la cultura meta.

Para validar estas hipótesis de estudio se presentan los siguiente objetivos que se plasman en la recopilación de un corpus de sitios web corporativos de España:

- a. Revisar la literatura sobre la localización para observar la influencia, cambios y restricciones que impone este proceso en la tarea de la traducción.
- b. Revisar la literatura sobre géneros y tipologías textuales en traductología, con especial énfasis en los factores extratextuales, el género, como comienzo de un análisis de factores intratextuales, la tipología textual, en estudios posteriores.
- c. Delimitar, desde un enfoque interdisciplinario, que constituye el género *sitio web corporativo*.
- d. Delimitar qué constituye un texto web desde una perspectiva traductológica.

- e. Asentar las bases teóricas para la recopilación de un corpus de textos web completos naturales en español, introduciendo el concepto de *hipertexto* (Storrer, 2002) o sitio web completo como unidad de análisis, dejando a un lado los estudios que se centran en las peculiaridades de páginas web como unidad.
- f. Introducir el concepto de *locale* en la recopilación de corpus en los estudios de localización. Justificamos la recopilación de un corpus *monolocale* para acercar las investigaciones en traductología a la industria de la localización y para acotar el campo de estudio.
- g. Establecer un marco de análisis de los textos web que sea de interés al traductor y que produzca por medio de estudios contrastivos una comparación de las convenciones a todos los niveles en textos web.
- h. Estudiar la superestructura del género web en España como base para futuras comparaciones con otras lenguas y producir una base de utilidad en la labor del traductor de este género web.
- i. Estudiar el grado de convencionalización de ciertos elementos intratextuales, como los términos empleados para la denominación de los bloques y secciones comunicativas, o el tenor digital en el género de estudio.

## **2. EL PROCESO DE LOCALIZACIÓN**

En la presente investigación es esencial delimitar cómo se desarrolla el proceso de localización y qué aspectos lo diferencian de otros procesos de traducción. Las diferentes definiciones de localización, provenientes en su mayor parte de la industria y no de la traductología, muestran varios denominadores comunes del proceso, como la separación entre un componente cultural y otro lingüístico (Pym, 2004; LISA, 2003; Esselink, 2000; O'Hagan y Ashworth, 2003; Quirion, 2003), la denominación de producto a la unidad de localización (LISA 2003; Yunker 2003; Esselink 2001), la dirección hacia una *locale* o «región» meta (Pym, 2004; LISA, 2003; Esselink, 2000; O'Hagan y Ashworth, 2003; Comité sectoriel de l'industrie canadienne de la traduction, 1999, cf. Quirion 2003), y a menudo, la no aparición tácita de la palabra traducción (LISA, 2003; O'Hagan y Ashworth, 2003; Yunker 2003)<sup>4</sup>. Este último aspecto, sorprendente desde el punto de vista de la traductología, reside en la importancia que otorga la industria a la compleja faceta técnica que no suele presentarse en otros tipos de traducción. Este proceso suele seguir un ciclo relativamente bien definido (LISA, 2004: 15) y que no se centra exclusivamente en el componente lingüístico/cultural, de ahí que se denomine localización y no traducción (Pym 2004: 26).

La localización, o *l10n* como la denomina la industria<sup>5</sup>, se está desarrollando a pasos agigantados durante los últimos años (LISA, 2004: 18)<sup>6</sup>. La localización comenzó en los años 80, centrada en los productos de software, y se fue extendiendo a otros campos como la localización de productos web, interfaces de productos de electrónica, móviles etc. (Esselink, 2001: 5 y ss). La importancia de la localización de material web y las diferencias en el proceso han llevado a algunos autores a utilizar el término *e-localization* (Schäler, 2001: 22-6; Cronin, 2003: 14).

### **2.1. Globalización e Internacionalización**

En primer lugar, y para acotar claramente el proceso de localización, necesitamos definir dos conceptos que están interrelacionados con éste y cuyos objetivos y

---

<sup>4</sup> Excepto en la definición de Bert Esselink en *A Practical Guide to Localization* (2001).

<sup>5</sup> L10n, se debe a que existen 10 letras entre la «l» y la «n» en la palabra localización.

<sup>6</sup> Según los cálculos de LISA el volumen anual de negocio es de 3.700 millones a 5.000 millones de dólares anuales.

procesos se suelen confundir. Además, estos dos procesos, o su ausencia, afectan al desarrollo de la tarea de traducción durante la localización.

La industria de la localización descubrió pronto la necesidad de realizar cambios y adaptaciones tanto la organización empresarial como en el desarrollo técnico de los productos de software para un ahorro considerable de tiempo y recursos (Brooks, 2000), apareciendo los procesos de globalización e internacionalización respectivamente. Según Pym (2004: 30) sólo existe un proceso general, la globalización, siendo la internacionalización y la localización partes integrantes del mismo. El proceso de traducción en localización se encuentra inmerso en estos otros dos otros procesos, aunque es la internacionalización la que facilita directamente el proceso de localización. En el siguiente gráfico de LISA (2004:15) se puede observar la interdependencia de los mismos:



Figure 1. The global product development cycle.

Figura 1. Esquema del ciclo de desarrollo de un producto global según LISA (2004: 15).

La globalización, o *g11n*, se centra sobre todo en la organización empresarial y se define por sus objetivos. Según LISA (2004: 14):

...making all the necessary technical, financial, managerial, personnel, marketing, and other enterprise decisions necessary to facilitate localization.

En el ámbito de la localización web, la globalización se entiende también como la aspiración de un sitio web determinado de alcanzar la mayor audiencia posible,

algunos autores llegan incluso a denominar «globalización» a la localización de sitios web (Yunker, 2003).

La internacionalización, o *I14n*, se centra en la parte técnica del proceso. En general, es el conjunto de procedimientos que aseguran que un documento o aplicación no esté culturalmente limitado y sea técnicamente fácil de localizar (Tercedor, 2005), sin que se requieran modificaciones de profundidad. LISA (2003) lo define como:

*Internationalization is the process of generalizing a product so that it can handle multiple languages and cultural conventions without the need for redesign. Internationalization takes place at the level of program design and document development. [énfasis propio]*

Un año más tarde, LISA (2004:14) añadió a su definición la abstracción de la funcionalidad como aspecto fundamental, y que reviste una importancia en los estudios de localización en traductología:

*...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.*

Es precisamente en este punto, en el objetivo de abstraer la funcionalidad comunicativa de un producto tanto de software como web e independizarla de cualquier lengua, en el que observamos una cierto paralelismo con las teorías funcionales y pragmáticas de la traducción (Nord, 1996, 1997; Hatim y Mason, 97, 90; Reiss y Vermeer, 1984), puesto que estos investigadores dan gran importancia a la necesidad de recrear la función comunicativa de un texto origen [TO] origen en el texto meta [TM]. En este caso, el concepto de «funcionalidad» equivaldría a la función comunicativa de cada componente o segmento y no en el sentido con el que se usa en la presente investigación, ya que «funcionalidad» se entiende como la interacción entre el usuario y la máquina o el usuario y el emisor del texto (véase 4.2.2.6). Uno de los objetivos de la internacionalización sería por tanto definir y acotar claramente la función de cada parte del programa o sitio web y abstraerla, con el fin de que se pueda expresar en cualquier lengua y cultura. En este caso en particular, los elementos de la situación comunicativa son mucho más limitados al tratarse de un entorno digital y con una interacción claramente definida por el desarrollador. Esta abstracción de la funcionalidad se ha relacionado con el tan criticado concepto de la equivalencia en traducción que, en el caso de la internacionalización, se puede tratar de una «*internationalization-based equivalence*» (Pym, 2004), diferente de la equivalencia entre dos lenguas.

En el caso específico de la localización web, un entorno más flexible que el software, se ha identificado también la internacionalización web con los cambios en el *back-end* o parte técnica sitio web, centrándose la localización en el *front-end* o estructura de superficie<sup>7</sup>, lo que el usuario ve (Singh y Pereira, 2005:7). Estos dos conceptos son importantes en localización puesto que indican la diferencia clara entre el texto con el que trabajan los desarrolladores y traductores y lo que ve el receptor final.

## 2.2. Localización y *locale*

Como vemos, el objetivo de la globalización e internacionalización es preparar las organizaciones y la parte técnica del producto para que sean independientes de la lengua de desarrollo, normalmente el inglés<sup>8</sup>, y que puedan ser localizados en cualquier *locale*. La definición de localización que encontramos en más investigaciones es la propuesta en 2003 por la Localisation Industry Standards Association (LISA):

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<sup>9</sup>

Esta definición incluye gran cantidad de etapas en las que el proceso de traducción es una de la más importantes, pero no la única. En esta definición de la industria de la localización, como mencionamos anteriormente, se encuentran los conceptos básicos que aparecen en el resto de definiciones. En primer lugar, tratamos con productos<sup>10</sup> y no con textos, postura esta última que se defenderá en el presente estudio que se centra en una visión traductológica. En segundo lugar, el proceso incluye una parte 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). Y en tercer lugar nos encontramos con el concepto de *locale*, concepto este

---

<sup>7</sup> En ciertos estudios se denomina *surface structure-deep structure* (Kersten et al, 2001) o *front-end – back-end* (Cronin, 2003), conceptos más usados.

<sup>8</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.

<sup>9</sup> «La localización supone tomar un producto y adecuarlo culturalmente y lingüísticamente a la población de destino (pais/región e idioma) donde se usará y se venderá [traducción propia]»

<sup>10</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.

que emana de la parte técnica del proceso y que se puede definir por una combinación específica de lengua, región y codificación (Esselink, 2001)<sup>11</sup>.

La propia LISA (2004:13), consciente de la insuficiencia de la definición, propuso en el *Localization Primer* del 2004 una definición ampliada de la misma:

Localization is the process of modifying products or services to account for differences in distinct markets...

Separa a continuación tres componentes: el lingüístico, el contenido/cultural, y 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, y deja a un lado el concepto de *locale* para denominarlos «mercados». Al fin y al cabo, el concepto de *locale*, que irregularmente corresponde con el concepto de cultura meta, surge de la industria y está claramente influenciado por el tamaño e importancia comercial del mercado en cuestión (Pym, 2004:40).

Los estudios de traducción necesitan reclamar la importancia del componente traductológico del proceso (Quirion, 2003), ya que se le suele restar importancia aun cuando supone el 50% de la inversión en localización (LISA 2004). Sólo en una de las definiciones, la que aparece en el trabajo de Bert Esselink (2001: 2), encontramos expresamente la mención de traducción:

Localization is the translation and adaptation of a software or web product, which includes the software application itself and all related product documentation. [énfasis propio]

Esta definición incluye expresamente además los productos web, objetivos de este estudio, de características ligeramente diferentes a los productos de software. Se sigue, no obstante, separando la traducción de la adaptación, que en este caso se refiere a la adaptación técnica, ya que la adaptación cultural, desde un punto de vista traductológico, es parte integrante del proceso de traducción global. Las teorías de la traducción no separan entre la necesidad de adaptación textual, cultural, o tipológica y un proceso lingüístico en general, sino que la adaptación es parte integrante del proceso de traducción (Gouadec, 2003), en nuestro caso, esta labor simplemente se refuerza de forma especial en el proceso de localización (Quirion, 2003: 547):

---

<sup>11</sup> De hecho, el término «localización» surgió a partir del término *locale*.

...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.

¿Deberíamos entonces hablar de localización cuando se traduce un manual de instrucciones español para Gran Bretaña y se incluye un número de teléfono diferente o una sección específica por las peculiaridades de la red eléctrica en el Reino Unido? (Gamero, 2001). ¿O se puede hablar entonces de localización cuando se substituye la palabra «*biscuit*» por la de «*cookie*» en la edición norteamericana de *Harry Potter*, o el título «*Harry Potter and the Phylosopher's Ston*» por «*Harry Potter and the Sorcerer's Stone*»? (Yunker, 2003: 432). Todos estos son casos de adaptación cultural inter o intralingüística, y no por ello denominados localización, sino adaptación o traducción en sí. A 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, 2004: 25), con la consiguiente actualización en las competencias del traductor (Quirion, 2003).

Otras definiciones en diferentes publicaciones incluyen generalmente la adaptación cultural y lingüística del producto, como la del Comité Sectorial de la Traducción del Canadá (1999, cf. Quirion, 2003), según el cual la localización es:

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: 66), 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.

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>12</sup>. Esta separación se asemeja a la separación anterior entre *front end* y *back end* en un producto de software.

---

<sup>12</sup> Los autores se basan en la separación que realiza Gile (1995) en el contexto de la interpretación.

En general, se observa en la mayoría de las definiciones un intento del sector de la localización de reducir a la traducción a un mero intercambio lingüístico, como si no fuera más que un proceso mecánico, separado del componente cultural y técnico (Pym, 2004). La industria tardó en reconocer la importancia y el esfuerzo necesario en este proceso para una localización de calidad (Brooks, 2000), y aun se puede observar el anhelo de automatizar el proceso, de manera que la internacionalización facilite la localización hasta el punto de que sea automática e instantánea por medio de la traducción automática (Pym, 2004)<sup>13</sup>.

Esta separación se observa además en las prácticas empresariales, puesto que normalmente el 80% de las labores de traducción (LISA, 2004) se envían a compañías fuera de la misma.

### **2.3. Emisores y receptores en localización**

El texto en localización posee unas características propias que influyen y condicionan de manera especial el proceso de traducción. El objetivo último de la localización es que el producto o texto sea recibido por los receptores como si se hubiera desarrollado en la región de destino (LISA, 2004: 11). De ahí surge una de las características principales del texto localizado, puesto que a diferencia de otros tipos de traducción, el texto localizado no necesita representar al texto origen, sino que necesita representar la función textual del mismo en otra cultura. Según Pym (2004: 6)<sup>14</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.

De este aspecto surge un punto de interés desde un enfoque textual y traductológico en estas definiciones, el hecho de que 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>15</sup> (*Translation Mediated*

---

<sup>13</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.

<sup>14</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>15</sup> Se considera como un caso de «Comunicación Mediada por Ordenador» que incluye la interpretación de chats, traducción de correos etc.

*Communication*), se menciona al emisor y al receptor. En general, esto se debe a que el software necesita ser anónimo, «*software es anonymous technology speaking to anonymous markets*» (Pym, 2004:6). En nuestro caso, los sitios web corporativos, la comunicación no sólo se entabla entre el sitio web como ente autónomo y el emisor, sino que también se produce entre la empresa como entidad única y el receptor en un nuevo modelo de comunicación de masas interactivo (Janoschka, 2003: 98).

Por cuanto al receptor del TM, 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, y se suele denominar «usuario» (LISA, 2004). Este matiz es esencial puesto que señala que los textos digitales no necesitan ser leídos, sino «usados», en un contexto de interactividad que marcan 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 el estudio de Pym (2004: 34) se critica que la definición implique que todos los receptores van a «usar» el texto, y de hecho, no todos los receptores van a usar el mismo, como en la lectura de una ventana emergente de publicidad web que en general se cierra sin más<sup>16</sup>.

## 2.4. Localización y su componente cultural

La adaptación cultural de un producto de software o texto web supone uno de los pilares de la localización (LISA, 2004: 13). Ésta adaptación cultural se asemeja a la necesaria en otros tipos de traducciones<sup>17</sup>, combinada con otros elementos específicos dados por el medio y el nuevo modelo de comunicación.

A pesar de los intentos de la industria de crear productos lo más neutros posibles (Cronin, 2003: 18), desde los comienzos de la localización se ha reconocido la importancia de la adaptación cultural (Esselink, 2001). Este componente cultural se ha estudiado en mayor profundidad durante los últimos años (Tercedor, 2005; Singh y Pereira, 2005; Signh et al, 2004; Schäler, 2002). La adaptación cultural a la cultura meta ha sido el centro de varios estudios empíricos que la relacionan con

---

<sup>16</sup> El usuario no usaría el mismo pero si que «interactúa» al cerrar esta ventana emergente.

<sup>17</sup> La adaptación del espacio disponible se lleva realizando desde hace mucho años en sustitución 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.

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).

Desde la traductología, Tercedor (2005: 153) señala cinco elementos clave de esta adaptación desde una doble perspectiva pragmático-textual y cognitiva: los iconos, los aspectos lingüístico-textuales como la intertextualidad, el registro o el macrotexto; los gráficos, los técnicos y finalmente los cognitivos, tales como la navegación, las metáforas, los modelos mentales y la interacción.

Singh y Pereira (2005), desde la perspectiva del marketing, realizan un estudio pormenorizado de las adaptaciones culturales en localización de los sitios web en el que señalan la percepción, el simbolismo y el comportamiento como elementos clave que definen la cultura, y que a su vez establecen los valores compartidos y los patrones estructurados de pensamiento. La contribución esencial de los autores se centra en la elaboración de un marco de estudio de valores culturales en el que basan su investigación sobre las adaptaciones culturales en la localización web. Los valores que incluyen se centran en el elemento del comportamiento y se fundamentan en los estudios de Hofstede (1991)<sup>18</sup>:

- Individualismo frente a colectivismo (*Individualism-Colectivism*)
- Distancia del poder (*Power distance*)
- Tolerancia hacia la incertidumbre (*Uncertainty Avoidance*)
- Masculinidad-Feminidad (*Masculinity-Femininity*)
- Contexto bajo o alto <sup>19</sup>(*Low-High Context*)

La importancia de estos valores en nuestro estudio se debe a las diferencias significativas que se observan entre España y los países anglosajones en todos los valores. Estas diferencias conllevan unas implicaciones claras 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, conlleva por ejemplo 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. Debido a que en España el valor «tolerancia hacia la incertidumbre» es alto, se puede observar el uso de una terminología más local, una importancia de la atención al cliente o la inclusión de

---

<sup>18</sup> Para mayor información sobre estas dimensiones se puede visitar el sitio web del autor.  
<http://www.geert-hofstede.com/>.

<sup>19</sup> Este valor se basa en los estudios de Hall (1976).

los datos de registro mercantil y código de identificación fiscal (CIF) de la empresa como garantía y seguridad al usuario.

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, la posibilidad de realizar adaptaciones culturales en el contexto de la localización depende del escopo (Reiss y Vermeer, 1984) y en este caso, del grado de localización que se encargue, que se tratará a continuación.

## 2.5. Diferentes grados de localización

El fin de la localización como proceso económico hace que en la misma nos encontremos con diferentes grados de localización dependiendo de la importancia económica del mercado o la *locale* (Brooks, 2000). Este aspecto es de importancia en nuestro estudio puesto que el proceso de localización se encuentra inevitablemente influenciado por el escopo o comisión del mismo (Reiss y Vermeer, 1984), por lo que no se podría hablar de unas características únicas del proceso al que se enfrentan los traductores en su labor localizadora, al depender éstas del grado de localización.

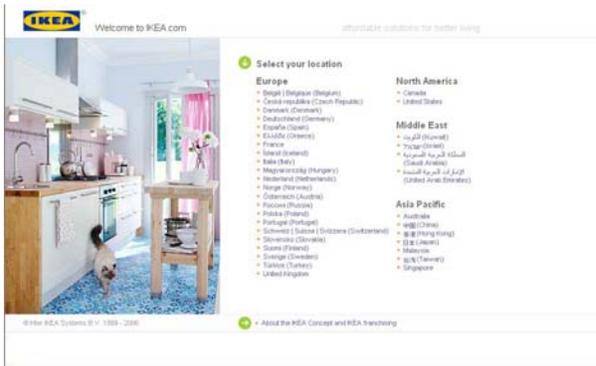
Si tomamos el ejemplo de Microsoft (Brooks, 2000:49-50), un modelo repetido en la industria y de gran impacto en la localización de sitios web, observamos que se divide en tres grados:

- Primer grado; «*enabled*». Los usuarios pueden componer textos en su propio idioma, pero el software y los archivos de ayuda están en inglés. Este grado no tendría importancia en una investigación sobre la textualidad web ya que en este caso la adaptación sería sólo técnica.
- Segundo grado; 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, 2004).
- Tercer grado; adaptado. Se revisan o se recrean las herramientas lingüísticas, el contenido y las funciones para el mercado de destino.

Este modelo se puede observar también en los sitios web, pero sólo en el segundo y tercer grado anteriores. Por ejemplo, el sitio Web de Ikea<sup>20</sup>, señalado como posiblemente 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 iguales, mientras que el francés o el español presentan diferencias notables fruto de su localización a un tercer grado.

---

<sup>20</sup> [www.ikea.com](http://www.ikea.com)



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/)



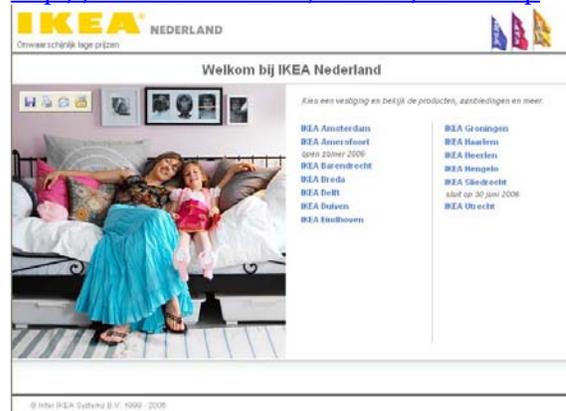
Sitio web de Turquía. El contenido y forma es similar a la web de Arabia Saudita, pero con orientación de izquierda a la derecha. <http://www.ikea.com.tr/content/index.asp>



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. El contenido es similar a la web de Holanda o República 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 2. Diferentes grados de localización en Ikea.com.

Es necesario indicar que los modelos de globalización web pueden ser centralizados o descentralizados, (Yunker, 2003; O'Hagan y Ashworth, 2003:74), por lo que en ciertos casos la localización web a un tercer grado se puede 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 3.4), como Goolge.com, y 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 web, de gran importancia para la identificación de la misma (Hunt, 1995). Un ejemplo en el 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 es una producción local<sup>21</sup>.

En varios estudios sobre la localización de sitios web ( Singh y Pereira, 2005; Yunker 2003: 128-130)<sup>22</sup> se realizan ciertas separaciones de grados ligeramente diferente dependiendo de los objetivos del estudio. Singh y Pereira (2005: 10-15) se centran en la importancia de la cultura meta en la localización y realizan la siguiente clasificación:

- Sitios web estandarizados: aquellos sitios de empresas globales cuya web es similar para todos los países, en una sólo 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 paginas web localizadas específicas para cada país, pero no se adapta el *back-end* o estructura profunda de las páginas.

---

<sup>21</sup> Determinamos que es una producción local puesto que los comentarios de programación y las variables en los script java se encuentran en español.

<sup>22</sup> Yunker diferencia entre:

- Localización incremental: localización parcial sin inversiones en el apartado técnico para observar la importancia del mercado.
- Localización exhaustiva: la localización del sitio web a partir de un sitio original en el que se debe crear una experiencia del usuario equivalente.
- 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.

- 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 (véase Figura 2. Diferentes grados de localización en Ikea.com)

Si observamos los diferentes grados que la industria aplica a la localización, llegamos a la conclusión que el proceso de localización y su producto, el texto localizado, no tendrá unas características homogéneas sino que el escopo, en este caso, el grado de localización al que se someta un sitio web, afectará al texto meta, puesto que éste presentará una estructura profunda o *back-end*, superestructura, estructura discursiva y tipología textual más o menos dependiente del texto original.

## 2.6. Estudios sobre la localización

En nuestra búsqueda bibliográfica sobre los procesos de localización, observamos que la mayoría de los estudios monográficos sobre la localización parten de la industria y son muy escasos los estudios de impacto que surgen desde la perspectiva de los estudios de traducción (Quirion, 2003). De hecho, se observa que la industria de la localización se ha desarrollado instaurando sus propios modelos y dejando a un lado el cuerpo de conocimiento de la traductología (O'Hagan y Ashworth, 2003: 130).

La gran parte de los estudios se centran mayormente en los aspectos técnicos del proceso de localización (Quirion, 2003), intentando separar el aspecto textual o parte lingüística como algo independiente y de una menor complejidad (Pym, 2004:25). Durante los últimos años, se alzan voces desde la traductología por la investigación especializada en este campo. Como nos comenta Quirion (2003: 557):

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.

Así, gradualmente este nuevo campo de la localización se irá formando a partir de estudios empíricos desde el punto de traductológico; que el autor denomina aspectos «lingüísticos» al dejarse llevar por las definiciones de la localización de la industria o una perspectiva lingüística de los estudios de traducción<sup>23</sup>.

### **2.6.1. Estudios en localización desde la perspectiva traductológica**

Desde la perspectiva de los estudios de traducción, *The moving text* de Anthony Pym (2004) es el primer monográfico que estudia exclusivamente la localización desde una perspectiva teórica. Su libro es una recreación aplicada a la localización de su libro anterior, *Translation as Text Transfer*, (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, basado en una internacionalización de la funcionalidad del producto en la que a partir de un producto internacionalizado, se desarrollarían las diferentes versiones del producto que serían a su vez TO y no TM.

Otro libro centrado en la localización es el trabajo de O'Hagan y Ashworth (2003), que consiste en un estudio general sobre la Traducción Mediada por Ordenador (*Translation Mediated Communication*) partiendo de la Comunicación Mediada por Ordenador (*Computer Mediated Communication*). Estudia los retos y estudia casos de traducción con la 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 la «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 Globalizacion*, consiste en 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 la

---

<sup>23</sup> Como observo anteriormente, la industria suele separar el componenete lingüístico y cultural del proceso. Definiciones más correctas por las que se aboga en este estudio es la ofrecida por Esselink en la que se separa la traducción de la adaptación técnica.

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 único libro en español sería el volumen coordinado por Detlef Reineke, (2005), *Traducción y Localización: mercado, gestión y tecnologías* con un fin eminentemente práctico y descriptivo del proceso de localización desde una perspectiva técnica. El coordinador conjuga en el volumen contribuciones tanto de investigadores en traducción como informáticos o gestores de proyectos. En los diferentes capítulos se revisan las diferentes vertientes de la localización como el desarrollo de productos 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 si y definen la separación entre ésta y la localización.

Por cuanto a la enseñanza 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>24</sup> en Irlanda dirigido por Schäler y que publica regularmente *The Localization Reader*.

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. Kerian Dunne<sup>25</sup> (2005) aboga por la necesidad de descomponer en proceso en aspectos específicos, puesto que hasta la fecha se ha intentado realizar sobre todo estudios que engloban la totalidad del proceso. Esta tendencia se observa en la aparición de investigaciones desde la traductología sobre los aspectos culturales (Tercedor, 2005; Schäler, 2002), educativos (Quirion, 2003; Gouadec, 2003; Días Fouces, 2004), ideológicos, (Cronin, 2003), tipológicos (Bolaños, 2002, Jiménez, 2005), gestión terminológica en localización (Schmitz, 2005), o el uso de herramientas de traducción en localización (Reineke, 2005; O'Brien, 1998).

## 2.6.2. Estudios desde la industria

---

<sup>24</sup> <http://www.localisation.ie/>

<sup>25</sup> Profesor de Localización en Kent State University, EEUU. Comunicación personal durante la Conferencia Anual de la *American Translator Association* en Seattle, 10 de noviembre del 2005

Desde la industria de la localización, con un mayor énfasis en la parte técnica del proceso, encontramos un gran número de monográficos con un objetivo eminentemente práctico, tales como *A practical guide to Localization* (2001) de Bert Esselink, en el que se estudian detalladamente todas las etapas de las que se compone el proceso global de la localización. Por cuanto a 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 pero no dejan de ser superficiales. Por la importancia del lenguaje Xml en la web es necesario mencionar el libro de Yves Savourel, *XML Internationalization and Localization*, que aparece como el trabajo de mayor impacto en los estudios posteriores en localización.

La existencia de organizaciones como *Localization Industry Standards Association*, LISA<sup>26</sup>, es capital al realizar una gran función a hora de unificar y normalizar la industria. LISA realiza una gran labor con la gran cantidad de publicaciones sobre la Localización de software, web, productos móviles, la internacionalización, la recopilación de estadísticas sobre la industria, y sobre todo, con la adopción de estándares de memorias de traducción intercambiables (TMX).

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 *American Translator Association* y el sector de la localización de hacer puentes en los que intervienen responsables de localización de las empresas más punteras como Microsoft, Canon, Lockheed Martin etc.

La desconexión entre los interesados en la localización como proceso económico y los estudios de traductología lleva a Schäler (2002, 2005) a indicar el empobrecimiento al que nos puede llevar la localización sin estudios amplios desde la traductología.

---

<sup>26</sup> [www.lisa.org](http://www.lisa.org)

### 3. LA LOCALIZACIÓN DE SITIOS WEB

El desarrollo de la localización de sitios web se produjo adaptando y modelando los procesos en localización de software a las particularidades del texto digital distribuido por Internet, con lo que se considera una rama de la localización con características propias (Yunker, 2003: 30). El vertiginoso aumento en localización de sitios web se debe a que Internet se ha convertido en un modo esencial de comunicación en un mundo globalizado. Según gran cantidad de autores, la importancia revolucionadora de Internet es comparable a la invención de la imprenta en el siglo XV (Lockwood y Scott 1999:8; Posteguillo, 2003). El análisis de las cifras muestra que el número de usuarios en todo el mundo sigue creciendo, llegando aproximadamente a mil millones en 2005 (LISA, 2004). En España el número de personas conectadas a la web aumenta de manera exponencial, de 10 millones según las cifras de la Union Española de Internautas<sup>27</sup> a 17 millones en el 2003 según Nilsen Netratings<sup>28</sup>. Esta empresa, la más reconocida en estadísticas web, indica además que España fue el país con el mayor crecimiento porcentual en número de usuarios conectados del mundo (Nilsen-Netratings, 2003). Este aumento en el número usuarios es paralelo al incremento de sitios web en la *locale* que conforma España<sup>29</sup>. Sigue igualmente creciendo el número de páginas web activas en español, que son aproximadamente 17 millones y medio<sup>30</sup> según la empresa internacional Nielsen-Netratings. El universo digital en español, aun cuando nuestro estudio está acotado al espacio de España, sería mucho mayor si se incluye el resto de *locales* en español.

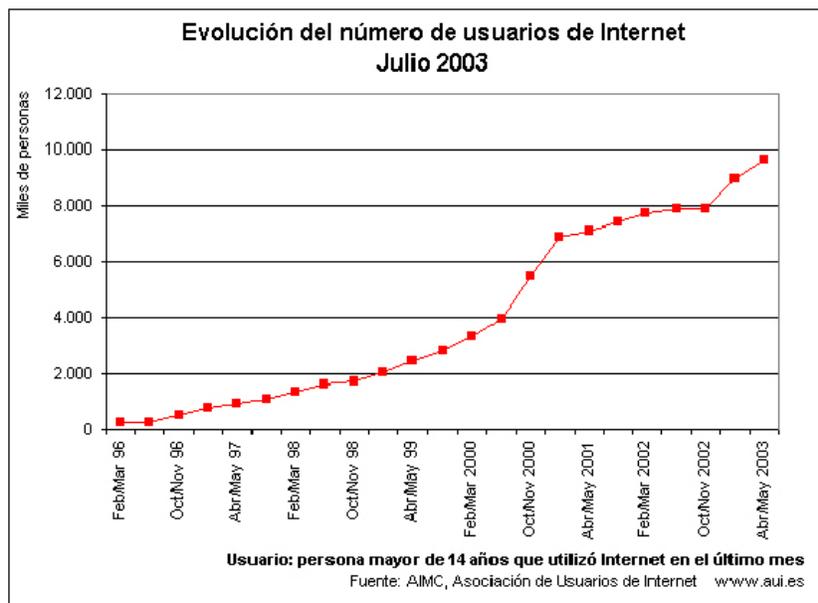
---

<sup>27</sup> [www.uai.com](http://www.uai.com)

<sup>28</sup> [www.nielsen-netratings.com](http://www.nielsen-netratings.com)

<sup>29</sup> 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>30</sup> En abril del 2006.



**Figura 3. Evolución del uso de Internet en España. Fuente: Asociación de Usuarios de Internet.**

Estas cifras muestran como se ha consolidado el uso de Internet como medio social de comunicación, objetivo del creador de la web (Berners-Lee, 1999)<sup>31</sup>, apareciendo un nuevo modelo de comunicación, que en nuestro caso específico de sitios web corporativos podemos denominar Modelo de Comunicación de Masas Interactivo (*Interactive Mass Communication*) siguiendo el trabajo de Anna Janoschka (2003: 96 y ss.).

No obstante, gran cantidad de estudios señalan que la mayoría del contenido web sigue estando en inglés. Según los cálculos de LISA (2003), del 62 al 87 % de los documentos están escritos en inglés, aun cuando el 90% de los usuarios prefieren usar una lengua diferente (Tercedor, 2005). Diversas publicaciones (LISA, 2004; Yunker, 2003) indican la primacía del inglés en las comunicaciones por Internet, en parte por el origen norteamericano de Internet que influyó en la codificación de la misma, superado con la adopción gradual de Unicode<sup>32</sup>. En nuestro caso 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>31</sup> Según menciona el creador de la web « The web is more a social medium than a technical one» (Berners-Lee, 1999:113).

<sup>32</sup> [www.unicode.org](http://www.unicode.org)

La primacía del inglés en las comunicaciones se relaciona con la adopción de terminología anglosajona (Carbolante, 2001; cf. Posteguillo, 2003) y en cierta medida de tipos textuales y géneros digitales norteamericanos entre las diferentes comunidades del mundo (Bolaños, 2002). 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 material web a las peculiaridades del nuevo medio y a un nuevo modelo de comunicación interactivo (Janoschka, 2003:96). Se debe además adaptar a las peculiaridades cognitivas de la lectura en pantalla, con gran cantidad de estudios en esta área entre los que destacan los estudios sobre usabilidad (Nielsen, 1999; Nielsen y Tahir, 2002; Brinck et al. 2002; Krug, 2005). Estos estudios de usabilidad<sup>33</sup> han desarrollado unas guías básicas basadas en la interactividad con los usuarios<sup>34</sup> desde el punto de vista cognitivo, y definen los patrones básicos del procesamiento cognitivo de páginas web, como que la lectura en pantalla es 25% más lenta que en papel o que la mayoría de usuarios escanean las páginas en vez de leerlas (Nielsen, 2000: 101), con su consiguiente efecto en el lenguaje empleado (McGovern et al. 2002; Kilian, 1999; Vilamor, 2001).

Las tipologías textuales y géneros que aparecen en la web poseen unas características diferentes en cada cultura, como se observó en el estudio del presente investigador en la variación terminológica y tipológica entre un corpus paralelo de páginas web localizadas al español y páginas producidas originalmente en España (Jiménez, 2005)<sup>35</sup>.

### 3.1. Investigación previa sobre la localización de sitios web

El centro de la investigación previa del propio autor fue la recopilación de corpus comparable (Baker, 1995) y su análisis empírico con técnicas de lingüística de corpus. El componente localizado del corpus lo compusieron 74 páginas web localizadas al español y el corpus original en español por 84 páginas. La hipótesis de trabajo inicial consistía en que en las páginas localizadas mostrarían estadísticamente unas frecuencias de unidades léxicas y colocaciones diferentes de las páginas españolas, debiéndose a una influencia clara de la lengua y cultura de origen, el inglés en nuestro caso. La traducción es un «*communicative event*

---

<sup>33</sup> «Usability» en inglés. La denominación en español de este término no está exenta de discusión.

<sup>34</sup> De nuevo, usamos la palabra «usuario» para referirnos al receptor del TM

<sup>35</sup> Estudio de investigación dentro del Programa de Doctorado en Procesos de Investigación dirigido por Maribel Tercedor y Clara Inés Lopez Rodriguez. El estudio fue publicado en las Actas de la Conferencia Anual de la Asociación Americana de Traductores, (2006) y presentado en esta conferencia y en la conferencia de la *Mountain State Foreign Language Conference* en la Universidad de Wake Forest, EEUU.

which is shaped by its own goals, pressures and context of production» (Baker, 1996:175), por lo que los textos que se producen en de esta manera deben mostrar rasgos lingüísticos distintivos que se pueden observar con técnicas de corpus (Zanettin, 2000).

El análisis empírico se centró en un elemento básico de los sitios web, la terminología empleada en los menús de navegación. En particular, se estudió los términos altamente convencionalizados en las páginas web anglosajonas, *contact us, about us, privacy policy, home y terms of use*. En los resultados del análisis se observaron diferencias claras en las características discursivas y tipológicas entre ambos componentes del corpus, que se explican por los diferentes contextos de producción textual de ambos tipos de textos. Las diferencias básicas que se encontraron se centran en la frecuencia de uso, la variabilidad terminológica, el tenor o la aparición de términos localizados inexistentes en el corpus original español. En la siguiente gráfica se observan las diferentes frecuencias de aparición de los diferentes términos que corresponden a *contact us* en inglés. Las barras azules corresponden al corpus de páginas originalmente desarrolladas en España, y las rojas al corpus de páginas localizadas al español.

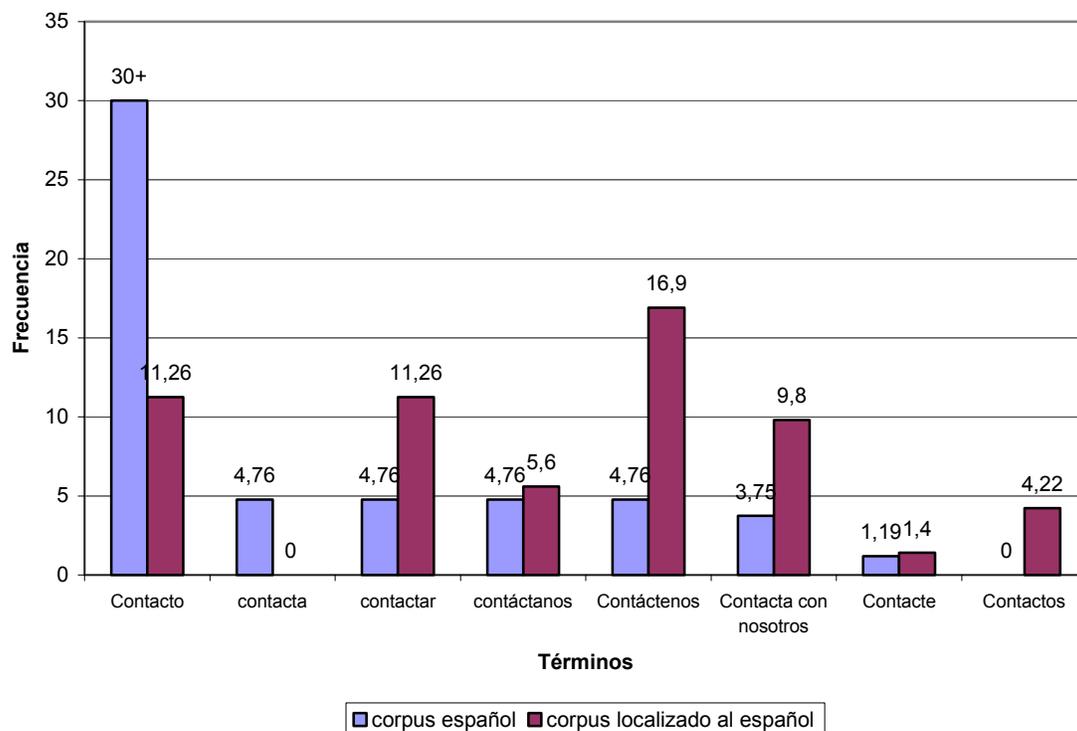


Tabla 1. Frecuencia de uso en corpus original y localizado. De Jiménez (2005).

Se observó como las frecuencias de aparición de términos eran distintas en el componente localizado del corpus, al igual que se señala en investigaciones en traductología que emplean esta metodología (Gellerstam, 1996; Laviosa-Baithwate, 1996). Se observa además como el término de mayor frecuencia en el componente traductológico, «contáctenos», refleja la influencia del término original «*contact us*», muy convencionalizado en la cultura anglosajona con un 89% de uso (Nielsen y Tahir, 2002).

Por cuanto a la variabilidad, la comparación del corpus comparable reveló como existen términos como *privacy policy* que mostraban un bajo grado de fijación en español, y por lo tanto, con una gran variabilidad en su localización. La variabilidad aparece tanto en el primer componente, *privacy*, con «privacidad» o «confidencialidad», como en *policy*, con «cláusula», «política», «declaración», o «compromiso». En este caso, justificamos la baja fijación de este término por el diferente grado de convencionalización de este bloque comunicativo en el género en diferentes culturas, la anglosajona y la española. Éste es una de los motivos por los que la presente investigación se centra en incluir el concepto de género en el estudio de la textualidad web. La inclusión del género en la investigación conlleva asimismo dejar a un lado las páginas web como unidad de estudio ya que el sitio web se constituye en la unidad de análisis.

### 3.2. Sitios web corporativos como objetivo del estudio

Para centrar el objeto del presente estudio, nos centramos en las particularidades textuales en **sitios web**, que pronto sobrepasará en importancia económica a la localización de software (LISA, 2004; Schäler, 2005)<sup>36</sup>. Los sitios web suponen una de las situaciones comunicativas que surgieron con el nuevo medio, Internet, y son uno de los posibles modos de comunicación en los que se han centrado los estudios de traductología <sup>37</sup> (O'Hagan y Ashworth, 2003).

---

<sup>36</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>37</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)

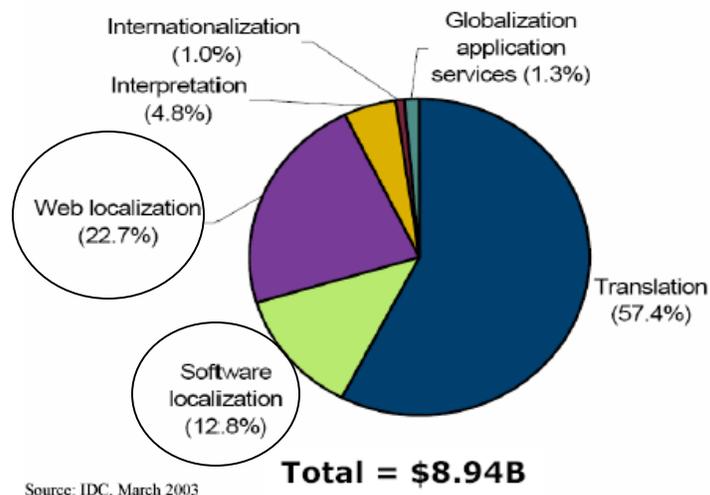


Figura 4. Predicción en el volumen de negocio en 2007. Fuente IDC Research (cf. Schäler, 2005).

La Web ha sido definida por su creador, Tim Berners-Lee (1999) como «*the universe of network-accessible information, an embodiment of human knowledge*», de donde surge el aspecto de mayor interés en este medio, la hipertextualidad (véase 5.4). En general, se suele usar la Web e Internet de manera indistinta, pero ésta es una práctica común que no se corresponde con la realidad, puesto que ésta es sólo una de las varias situaciones comunicativas en las que se puede usar Internet (Berners-Lee, 1999, cf. Crystal, 2001). Las particularidades lingüísticas de los chats, correos electrónicos, mensajes SMS o foros han sido estudiadas en gran detalle desde la lingüística<sup>38</sup> (Crystal, 2001, Posteguillo, 2003, López Quero, 2003), e incluso la traductología (O'Hagan y Ashworth, 2003), pero estas situaciones comunicativas tienen un menor impacto en la industria de la localización.

Para centrar el objetivo de estudio nos hemos guiado por dos criterios básicos. En primer lugar, los sitios web corporativos son uno de los textos web que se ha consolidado como género y por lo tanto se han convencionalizado gran cantidad de sus características (Askehave y Nielsen, 2005). Las investigaciones sobre géneros web han comprobado que los sitios web corporativos presentan un estilo y forma más únicos y definidos que en otros géneros, como los sitios web personales y sobre todo los institucionales (Kennedy y Shepherd, 2005), otro de los géneros web con mayor volumen de traducción. En segundo lugar, la elección se basa en el alto volumen de traducción en este área (Göpferich, 1995).

<sup>38</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).

La localización de sitios web presenta el mayor crecimiento en la traducción (Lockwood, 1999), y a su vez, dentro de las traducciones de sitios web, el estudio se limita a los corporativos puesto que la presencia de las empresas en la Web se ha convertido en una necesidad (Schmitt, 2000, cf. O'Hagan y Ashworth, 2003:18)<sup>39</sup>, convirtiéndose en el principal medio de información sobre productos y servicios sobre las mismas. Otra gran ventaja del nuevo medio es la facilidad para actualizar el contenido (O'Hagan y Ashworth, 2003:9). El estudio se limita por lo tanto a uno de los géneros web más convencionalizados, los «*corporate homepage*» en inglés que se puede definir preliminarmente (Kennedy y Shepherd, 2005):

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.

Este género lo definimos por lo tanto como sitios web que describen los intereses, actividades etc. de una empresa que pretende obtener un beneficio económico por medio de la venta de productos o servicios. Estas empresas pueden tener presencia exclusivamente en Internet, como [www.ebay.es](http://www.ebay.es), o puede ser el sitio web de una empresa consolidada con presencia física, [www.iberdrola.es](http://www.iberdrola.es) (Singh y Pereira, 2005: 2).

### **3.3. Importancia de un estudio sobre la textualidad de los sitios web**

El aumento de la importancia de los sitios web como plataforma de comunicación conlleva la necesidad de establecer criterios de calidad en los mismos, como se observa en la literatura sobre su diseño y producción (Brinck *et al.*, 2002). Esta necesidad de calidad afecta igualmente a su localización. Las cuestiones que se intentan discernir desde el punto de vista de esta investigación son: ¿Qué características textuales, discursivas y pragmáticas tienen los productos web desarrollados en España a los que tienen que parecerse los localizados? ¿Qué características son las que la industria de la localización aprecia como pertenecientes a la *locale* de España? ¿Cómo se ha llegado a definir las?

---

<sup>39</sup> «The June 2000 Forrester Research report (Schmitt 2000) states that multilingual web sites are no longer an option for US firms but a global business imperative».

La localización es parte de un proceso económico que está definido por el escopo (Reiss y Vermeer, 1984) y por su participación en el proceso de globalización. 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 Internet es «Atrás», y en este medio tan competitivo, si los usuarios no encuentran de manera efectiva lo que buscan, abandonan el texto web y siguen buscando (Nielsen, 2000). En España, por ejemplo, la media de permanencia por página web es de 50 segundos<sup>40</sup>. Los manuales de usabilidad y guías de estilo web remarcan la importancia de que las páginas tengan una estructura clara y concisa que permita su escaneado (Nielsen y Tahir, 2002) y que sea conforme con las convenciones discursivas, de registro, tipológicas etc. que espera el usuario (ibid). El *Localization Primer* de LISA (2004: 35) destaca la importancia de la localización de sitios web en el mundo actual:

One key benchmark of Web success is stickiness—the ability to attract new and repeat visitors and keep them on a site.

Señala así la importancia de una localización de calidad y que los productos deben parecer y sentirse como si se hubieran producido en país hacia el que se dirigen (LISA, 2004: 13)<sup>41</sup>. Esta posición no está libre de críticas (Schäler, 2002) puesto que aunque se localice de manera adecuada el sitio web con cambios de colores, iconos, etc., el contenido tiende a ser el mismo: materiales globalizados que se suponen neutros culturalmente.

Los estudios en traductología que se centran en los universales de la traducción postulan por medio de sus estudios empíricos de corpus que los textos traducidos presentan características y frecuencias que nos son típicas en textos naturales en cualquier idioma (Baker, 1995), aun cuando se aprecien como naturales por los receptores. ¿Es posible entonces el objetivo de LISA de que los productos localizados se reciban como se hubieran desarrollado en la cultura meta?

Nuestro estudio surge así por la necesidad de que la localización se estudie desde cada cultura meta o diferentes *locales* hacia la industria de la localización (Schäler, 2002), y se hace patente en nuestro caso, puesto que tradicionalmente, la localización suele ser un proceso *top-bottom*, partiendo desde empresas mayormente norteamericanas hacia cada cultura. Como ejemplo, proponemos el estudio anterior del propio autor (Jiménez, 2005), donde se encontraron términos

---

<sup>40</sup> Según Nielsen-Netratings en mayo del 2006.

<sup>41</sup> “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.” LISA (2004:11)

localizados que reflejan directamente el término anglosajón, como «sobre nosotros», «cláusula de privacidad», «declaración de privacidad» o «términos y condiciones», cuya aparición era inexistente en el corpus de 79 páginas de inicio de webs de empresas españolas, e igualmente inexistente en el corpus recopilado para la presente investigación.

De igual manera, la calidad y conformidad a las expectativas por cuanto a 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 la tipología textual y convenciones de género que esperan los usuarios españoles es esencial, puesto que de no ser así, el texto corre el riesgo de que no se perciba como verídico y creíble, y por lo tanto, no funcional en la cultura meta. Este aspecto ha sido estudiado por los analistas de captología que incluso han creado unas «*Guidelines for Web Credibility*»<sup>42</sup> de Stanford University (Frogg, 2002. Cf. Posteguillo, 2003: 154)

### **3.4. Textualidad de software y de sitios web**

Son varias las diferencias fundamentales que encontramos entre la textualidad de los programas de software y las de los sitios web, encontrándonos asimismo un gran número de características similares.

Desde el punto de vista textual, la Internet es un medio electrónico, global, e interactivo y cada uno de estas particularidades afectan en gran medida al tipo de texto y lengua que se encuentran en la misma (Crystal, 2001). La interacción e interactividad, fundamento de la funcionalidad, son esenciales y las distinguimos 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 realiza a través de correos electrónicos, formularios varios etc. La interactividad, por otro lado, se centra en 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).

En los programas 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. En los sitios web, nos encontramos tanto la interactividad entre

---

<sup>42</sup>Accesible a través de [www.webcredibility.org/guidelines](http://www.webcredibility.org/guidelines).

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 la organización o persona tras el sitio web (en formularios, contacto por correos electrónicos, compras electrónicas etc.). Es decir, que los textos digitales que se distribuyen por Internet no se ciñen a facilitar la interacción humana entre software 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>43</sup>), y por otro lado con el «emisor» del sitio, en nuestro caso, la empresa.

El medio digital influye en el proceso de traducción, puesto que los textos están inmersos en el código fuente de diverso tipo, que en el caso de software lo más normal es que esté separado<sup>44</sup> de las cadenas textuales que aparecen en pantalla, aunque pueden incluir variables de diverso tipo (Esselink 2001: 65). En el caso óptimo de localización de textos web, suelen ser directamente los programas de traducción asistida los que separan el texto del código durante el proceso de traducción. Como indicamos anteriormente, el grado de localización y de internacionalización influye en gran medida en el proceso de traducción.

El medio digital supone la visualización del texto a través de la pantalla con las restricciones de espacio que esto conlleva (Nielsen, 2000). En la localización de software, este punto suele presentar mayores problemas que se resuelven con programas de localización WYSIWYG<sup>45</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>46</sup>, es decir, adaptándolo al navegador, a la configuración de pantalla y al tamaño de letra que haya elegido el receptor (Nielsen, 2000: 318). De hecho, normalmente son las preferencias del receptor las que establecen como se visualiza la página, por lo que se menciona la necesidad de dejar a un lado el WYSIWYG (ibid).

---

<sup>43</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>44</sup> En el caso de que se haya producido una óptima internacionalización del producto de software

<sup>45</sup> WYSIWYG.- *What You See Is What You Get*

<sup>46</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.



**Figura 5. Disposición líquida en pantalla con posibilidad de variar el tamaño de la letra. Icono de [www.ono.es](http://www.ono.es).**

La localización web surgió de la localización de software, y por lo tanto, también se encuentran tipos de textos con similares procesos de traducción, como la localización de presentaciones en formato Macromedia Flash (Macromedia, 2004), o el caso de los archivos de ayuda<sup>47</sup> en localización de software. Este tipo de textos suele encontrarse en formato html o xml e interconectados por hipervínculos, por lo que su traducción es similar a la del material web. Los archivos de ayuda son 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 6. Producto de software completo.**

Por último, una de las mayores diferencias entre la textualidad de ambos casos se centra en la estabilidad del texto y en su posibilidad de actualización (Mata, 2005). El texto que compone un programa de software es más estable, puesto que las actualizaciones y cambios se producen de 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 se puede producir con

---

<sup>47</sup> De hecho, la mayoría de los archivos de ayuda en la actualidad se producen en html o xml.

gran rapidez, y por medio de la hipertextualidad, es posible añadir subtextos o elementos textuales en cualquier momento, siendo ésta una de sus características intrínsecas. Como se puede observar en el siguiente esquema de Mata (2005), el proceso de actualización del texto web se puede iniciar durante todo el ciclo vital del mismo.

<b>Ciclo Vital</b>	
<i>Producto de software</i>	<i>Contenido Web</i>
Diseño ↓	Diseño ↓
Desarrollo ↓	Desarrollo ↓↑
Producción ↓	Producción ↓↑
Lanzamiento ↓	Publicación ↓↑
Distribución ↓	Actualización ↑
Venta ↓	
Actualización [...] ↑	

Tabla 2. Ciclos de vida de productos de software *versus* contenido web (Mata, 2005: 194).

## 4. CONCEPTOS FUNDAMENTALES EN LA INVESTIGACIÓN: EL GÉNERO Y LA TIPOLOGÍA TEXTUAL.

La búsqueda de una delimitación clara de nuestro objetivo de estudio, el texto web, y en especial el texto web en sitios corporativos, está supeditada a acotar en primer lugar «género», «tipología textual», y «registro» (véase 5.9.1.1). Durante los últimos años ha aparecido en traductología un consenso más amplio sobre estos conceptos, a pesar de las diferencias que se observan en los mismos desde la introducción en el campo de la tipología textual (Reiss, 1971; Arntz, 1988; Göpferich, 1995; López Rodríguez, 2001), o el género<sup>48</sup> (Gamero, 1998, 2001; Borja Albí, 2003; García Izquierdo 2002, 2005; Bolaños, 2002). En general, los estudios que presentan una tipología textual y-o de género se pueden agrupar en aquellos que provienen de la lingüística del texto y de análisis del discurso y aquellos que provienen de la traductología. Debido a la diversidad de los mismos, y puesto que se han estudiado estos conceptos desde diferentes perspectivas, se suele observar una confusión terminológica, siendo a menudo intercambiables (Biber, 1995) o no importantes (Laviosa, 1997; Stubbs, 1996). Desde los años 60, el concepto de tipología textual es el más problemático y en el que se encuentra mayor confusión terminológica (Sánchez Trigo, 2000: 123), aunque éste suele ser el término más usado en traductología como cajón de sastre que incluye diferentes instancias de variación textual (López Rodríguez, 2001:90).

Hoy en día se acepta la separación clara de estos tres conceptos en la 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).

En nuestro trabajo, es necesario asumir una visión clara sobre qué constituye tanto el género como la tipología textual, ya que con independencia del área desde la que estudien estos conceptos, se usan de una forma diferente y con sentidos diferentes, o incluso se omiten deliberadamente como en Stubbs<sup>49</sup> (1996) o Laviosa (1998 a: 566)<sup>50</sup>. Además, varios estudios señalan que otra fuente de

---

<sup>48</sup> En Gamero (2001) se puede encontrar una excelente recopilación de los diferentes estudios en traductología que han revisado el concepto de «género».

<sup>49</sup> Stubbs (1996: 11) menciona: «*some authors distinguish between text type and genre. I will not*»

<sup>50</sup> Laviosa indica: «*I have deliberately chosed to avoid the words 'genre' and 'type'*»

confusión radica en que los términos que se empezaron a usar en traductología para «tipología textual» o *text type* equivalían a *textsorte* y *texttype* del alemán (Kussmaul, 1997: 97), es decir, «clase de tipo textual» y «tipo textual»<sup>51</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.

Pasamos a continuación a revisar los conceptos de género y de tipología textual desde una perspectiva que incluya la lingüística de texto, la traductología y la informática, dejando el registro para cuando se trate la dimensión comunicativa del contexto (véase 5.9.1.1). En esta primera fase de la investigación comenzamos introduciendo el concepto de género, dejando en un segundo plano la tipología textual, por lo que el tratamiento de este último será más breve y se centrará en su distinción clara con el concepto de género.

#### **4.1. Tipología textual y género**

La localización supone un caso de traducción de textos especializados y por lo tanto necesita ser estudiada en este contexto. Durante los últimos años se han producido en este campo un gran número de estudios que se centran en el enfoque de las tipologías textuales (Gommlich, 1993, Göpferich, 1995, Gamero, 1998, 2001; López Rodríguez, 2001; Mayor Serrano, 2002). Este enfoque se centra en la investigación sobre la identificación, clasificación y caracterización de textos especializados en dos o más lenguas, y parte de la base de que a la hora de redactar o leer 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, 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) y en la labor diaria de los traductores. Las tipologías textuales son propias de cada género textual y pueden presentar diferencias interlingüísticas e interculturales, con diferentes convenciones intratextuales en cada género. El traductor necesita reconocer todos los rasgos

---

<sup>51</sup> Traducción del alemán de Martín y Reina (1996) del libro de Reiss y Vermeer (1984).

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. En su estudio sobre las tipologías textuales Kussmal (1997) se preguntaba:

Shall we preserve the source-text-type structures and thus create a kind of alienation effect, or shall we conform to the target-text-type conventions and thus create a text which looks perfectly normal?

En nuestro caso, la respuesta es clara. El objetivo de la internacionalización es producir textos localizados que se reciban como producidos en la *locale* meta (LISA, 2004), incluyendo las convenciones tipológicas y de género web que espera el receptor en tales textos, por lo que se hace necesario su estudio en nuestra investigación.

La diferencia esencial entre el género y tipología textual se basa en que el primero se define por factores externos y socioculturales, mientras que los tipos textuales se definen por sus características lingüísticas internas. Ambos se encuentran asimismo convencionalizados en distintas lenguas: el género por cuanto a su estructura externa y su proyección sociocultural y la tipología por cuanto a su estructura lingüística (García Izquierdo y Nebot, 2003). 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, aunque no define estos dos conceptos. Los géneros además son en principio ilimitados, y la tipología textual es un conjunto cerrado con un número limitado de categorías (Trosborg 1997:15).

El presente estudio pretende conjugar las nociones de género y tipología textual, partiendo en primer lugar del concepto de género para centrarse en futuros estudios en la tipología textual típica de cada bloque comunicativo en el género analizado. Trosborg (1997: viii) señala cómo este tipo de análisis es de interés en el proceso de traducción, sobre todo en el desarrollo de estrategias que lo 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.

De ahí que en la metodología de trabajo partamos del concepto de género para desarrollar un marco comparativo sobre el que estudiar la tipología textual.

### 4.1.1. Tipologías textuales en traductología

A mediados de los sesenta empezaron a aparecer desde la lingüística y la traductología ciertas propuestas de tipologías que intentan abarcar todos los textos posibles, cubriendo todos los campos, como el técnico, jurídico, literario etc., así como los modos escritos y orales. Entre las tipologías de mayor impacto en los estudios posteriores, se encuentra la de Werlich (1975, 1983). Werlich presenta una tipología contextual que tiene en cuenta características internas y externas del texto, teniendo en cuenta el foco contextual o función del mismo. Los tipos textuales que propuso son: la narración, la descripción, la exposición, la argumentación y la instrucción<sup>52</sup>. El autor incluye la adecuación psicológica de estos tipos, con lo que pretende incluir la vertiente cognitiva en su tipología. La crítica más extendida a la clasificación de Werlich se centró en su rigidez y en su pretensión de establecer unas categorías a las que todos los textos naturales deben de ajustarse (Arntz, 1988). En estudios posteriores, se critica que los textos raramente suelen tener una sola función, y normalmente son multifuncionales (Gamero, 2001).

Si repasamos brevemente las diferentes propuestas de tipologías durante las últimas décadas, observamos que en traductología las propuestas de mayor impacto se pueden dividir entre:

- a. Tipologías funcionales, por estar basadas en la función del texto, (Reiss, 1971, 1976, 1981; Nord 1997) que parten de la función comunicativa. La primera propuesta de Reiss (1971) dividía los textos en entre los *texttypen* o tipos textuales informativos, expresivos, operativos, partiendo de las funciones del lenguaje de Bühler<sup>53</sup>. En estudios posteriores añadió los textos audiomedios (Reiss, 1976), textos escritos para ser recitados o cantados, categoría que modifica en 1980 denominándola textos multimedia para incluir elementos visuales como los comics o anuncios. La principal crítica a la tipología de Reiss

---

<sup>52</sup> La tipología de Biber desde una perspectiva lingüística distingue entre :

1. *Involved production*
2. *Informational production*
3. *Narrative concern*
4. *Explicit reference*
5. *Situation-dependent reference*
6. *Overt expression of persuasion*
7. *Abstract information*
8. *Online informational elaboration*

<sup>53</sup> Bühler (1934) identifica las funciones apelativa, expresiva y representativa

se basaba en la pretensión de que el tipo textual dicte las estrategias de traducción (Hurtado Albir, 1996).

- b. Tipologías basadas en la finalidad retórica o *rhetorical purpose* (Werlich, 1975; Baker, 1992; Hatim y Mason 1990, 1997; Trosborg 1997), que distinguen entre textos descriptivos, narrativos, expositivos, argumentativos y exhortativos. En el presente estudio aplicamos el enfoque pragmático semiótico de Hatim y Mason que engloban los textos narrativos y descriptivos dentro de la exposición.

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

Son escasos los estudios sobre tipología textual en Internet y suelen provenir de un enfoque exclusivamente lingüístico. La mezcla e hibridización tipológica aparece como el problema principal a la hora de establecer tipologías textuales en la web (Santini, 2005a, 2005b; Tercedor, 2005), dificultando la atribución de tipologías clásicas a estos textos. Aparecen por lo tanto textos que aparecen como inclasificables o multifuncionales (Santini, 2005a). Esto se debe en parte a la rápida evolución de las tipologías en el nuevo medio (Shepherd y Watters, 1998) y a la nueva organización textual que implica el hipertexto. Santini (Ibid) señala ciertos factores que dificultan la adjudicación de tipologías clásicas a los textos web: la existencia de gran número de elementos textuales codificados en imágenes con la imposibilidad de aplicar técnicas de lingüística de corpus, la gran cantidad de encabezamientos que se deben incluir en los textos web, los directorios, listados o la gran cantidad de texto tabular.

## **4.2. Los géneros digitales o cibergéneros**

A pesar de que el concepto de género tiene una larga historia como criterio clasificador, las primeras investigaciones sobre los géneros web se produjeron a partir de la década de los 90 (Crownston y Williams, 1997). La aparición de estos géneros fue paralela a la expansión de su medio, la Internet, y se suelen denominar géneros digitales o cibergéneros desde los campos de la lingüística (Crystal, 2001, Posteguillo, 2003), la Comunicación Mediada por Ordenador (O'Hagan y Ashworth, 2003), la informática ( Kennedy y Shepherd, 2005; Shepherd y Watters, 1998), o la traductología (Bolaños, 2002). Las últimas

investigaciones sobre cibergéneros<sup>54</sup> se centran en los criterios para su clasificación (Crownston and Williams, 1999; Rehm, 2002; Shepherd y Watters, 1999), descripción (Askehave y Nielsen, 2005), la identificación automática de los mismos como complemento a las tareas de búsqueda en Internet ( Kennedy y Shepherd, 2005)<sup>55</sup>, o descripción con fines traductológicos (Bolaños, 2002).

Para un estudio pormenorizado desde la traductología se hace necesario revisar la evolución de este concepto y observar los factores que influyen en la consolidación y evolución de estos nuevos géneros.

#### **4.2.1. Evolución del concepto de género**

El concepto de género apareció en los estudios literarios y no fue hasta principios de los ochenta cuando comenzó a introducir en los estudios de lenguajes de especialidad, sobre todo con el objetivo de clasificar los textos con objetivos pedagógicos (Swales, 1990; Bhatia, 1993). Ambos estudios incluyeron un estudio pormenorizado del concepto de género en la introducción a sus trabajos. Estos estudios influyeron de manera especial en la aparición del interés sobre el género en los estudios de traductología (García Izquierdo, 2004; Gamero, 2001, 1998; Hurtado Albir, 1995; Hatim y Mason, 1990, 1997; Bolaños, 2002). Al igual que sucede con el registro, a menudo se pueden encontrar gran cantidad de estudios dentro de la lingüística del texto, análisis del discurso, análisis contrastivos y traductología en los que aparecen de forma indistinta los conceptos de «género» y de «tipología textual» esenciales en nuestro trabajo. El concepto de género, no obstante, se encuentra en la actualidad más delimitado con gran cantidad de estudios e incluso grupos de investigación como el GENTT<sup>56</sup> (García Izquierdo y Nebot, 2003), exclusivamente dedicados al mismo. Son por tanto muchos los autores que coinciden en el uso del concepto de género en sus estudios (García Izquierdo y Nebot, 2003; Borja Albí 2000, 2005; Bolaños, 2002; Gamero, 2001; Trosborg, 1997; Bhatia 1993; Swales, 1990; Biber, 1989).

Son varias las definiciones de género que se mencionan en los diferentes estudios en traductología, como la de Hymes (1974), Miller (1984, 1994), Martin (1984: 286), Kress (1985), Swales (1990), Yates and Orlikowski, (1992, 1998) o Duff (2000). La definición de género que usamos en el presente estudio es la

---

<sup>54</sup> No es nuestro objetivo presentar una bibliografía exhaustiva de los estudios en cibergéneros.

<sup>55</sup> La identificación automática del género de una página web para su inclusión como criterio en las tareas de búsqueda es un campo que cuenta con gran número de investigaciones en la actualidad.

<sup>56</sup> GENTT.-Gèneres Textuals per a la Traducció/ Géneros Textuales para la Traducción

presentada por Hatim y Mason (1990), y mantenida en sus estudios posteriores (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.

Los géneros son actos de comunicación que se expresan por medio de formas textuales convencionalizadas; estos textos reflejan las funciones de una situación social determinada. Este concepto suma aspectos formales, «*conventionalized forms of texts*», cognitivos, «*purposes of the participants*» y socioculturales, «*particular social occasions*». El aspecto social de los géneros es esencial, y se observa en que su definición proviene originalmente del sociolingüista Kress (1985:19). Son varios los aspectos que encontramos en esta definición de esencial importancia en la situación comunicativa web, como la importancia de la convención textual que se establece con un fin o propósito claro (Nielsen y Tahir, 2002; Krug, 2005).

Es en el aspecto social y convencional del género, es decir, en criterios externos, dónde reside la diferencia fundamental entre los términos «género» y «tipología textual», claramente indicada por Biber (1988: 170) y mencionada en Trosborg (1997):

I use the term "genre" to refer to categorizations assigned on the basis of external criteria. I use the term "text type", on the other hand, to refer to groupings of texts that are similar with respect to their linguistic form, irrespective of genre categories. [...] In a fully developed typology of texts, genres and text types must be distinguished, and the relations among them identified and explained.

Es decir, que los géneros se dividen según criterios externos y socioculturales, mientras que los tipos textuales según las características lingüísticas internas. Los tipos textuales también se encuentran convencionalizados, pero en su estructura lingüística y no en su aspecto sociocultural (García Izquierdo y Nebot, 2003).

Los géneros presentan además subgéneros característicos. Los géneros se pueden dividir según diferentes criterios externos (véase 4.2.2): los criterios que indica Biber (1998) son el tema y el propósito del hablante y estos pueden variar, de ahí que los subgéneros de cada categoría puedan tener características lingüísticas diferentes, «*some genres have several sub-classes which are quite different from one another*» (Biber, 1988: 170). Esta división entre género y subgénero es esencial en nuestra investigación, puesto que es muy común la mezcla de géneros y subgéneros en un mismo sitio web (Shepherd y Watters, 1999). En los estudios de

género aparece también el concepto de «supragénero»<sup>57</sup> (Izquierdo y Nebot, 2003), que se refiere a una agrupación de géneros de similares características pero que no se corresponde a un género delimitado y concreto. En nuestra investigación encontramos que el género *corporate homepage*, el sitio web corporativo, se puede agrupar en el supragénero de *homepage* general, que a su vez incluye los sitios web corporativos, los institucionales y los personales (Kennedy y Shepherd, 2005), pudiendo el género *corporate homepage* dividirse en subgéneros como los sitios web corporativos clásicos, o los basados en la venta *online* entre otros.

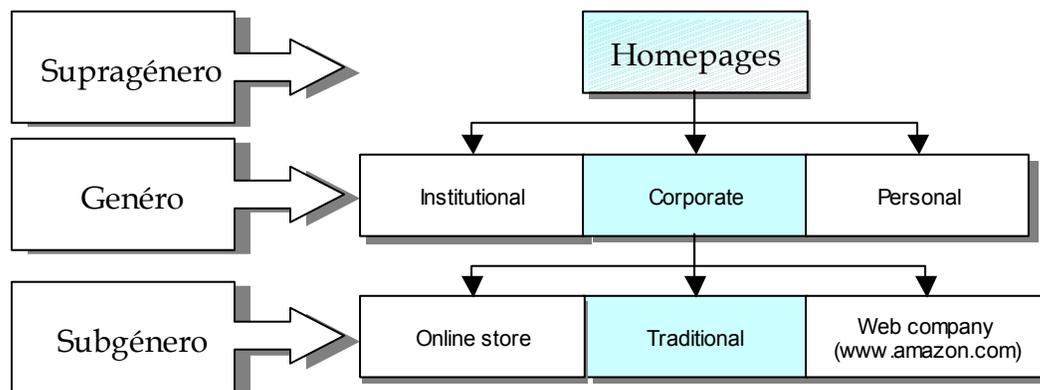


Figura 7. Propuesta de situación del género objeto de estudio.

Desde el campo de la lengua para fines específicos, Swales (1990) añade dos elementos fundamentales para entender el concepto de género y su descripción: la organización y la prototipicidad. Cada caso de texto es un ejemplo de género, que es esencial en la comunidad de discurso. A su vez cada género, resultado de una situación comunicativa convencional, se organiza según una estructura retórica formada por *etapas o movimientos*, que en el caso de los sitios web se han identificado con las diferentes páginas de las que se compone un sitio web corporativo (Askehave y Nielsen, 2005). En nuestro caso, la estructura de *movimientos o etapas* no sería lineal y fija, como puede suceder en un artículo de investigación en medicina, ya que las características del hipertexto hacen que el orden de movimientos lo dirija el usuario (véase 5.4). Tampoco es muy estricta, puesto que dependiendo de cada texto pueden existir etapas más o menos opcionales. Así un usuario, como miembro de la comunidad de discurso que

---

<sup>57</sup> Swales (1990) incluye el «pre-género» o «etiqueta supragenérica».

accede a un sitio web corporativo, esperará encontrar ciertos movimientos posibles como una página de contacto, información sobre la empresa etc. en forma de hipervínculos. Bhatia (1993), que se basa en Swales, intenta además encontrar ciertas regularidades en la organización de los géneros. Éstas se basan en características cognitivas acumuladas por los miembros de la comunidad de discurso a través de su conocimiento de la situación social y de las convenciones que potencian la eficacia comunicativa del texto.

Otra propuesta de impacto a mencionar es el modelo de Trosborg (1997:11), que para diferenciar el género utiliza también este modelo sistémico funcional de Halliday seguido por Hatim y Mason (1990, 1997) y propone definir el género a través de criterios múltiples<sup>58</sup>, usando las variables de campo, modo y tenor. En primer lugar, el campo se refiere 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 refiere al componente que representa el medio. Uno de los pilares del modelo de Trosborg es la multidimensionalidad e hibridación de los mismos (Trosborg, 1997), esenciales en los géneros digitales debido a su constante evolución que conlleva la mezcla de géneros, tipologías y registros (Tercedor, 2005).

#### **4.2.2. El Modelo de Género de Silvia Gamero (1998, 2001)**

Silvia Gamero (2001) propone un modelo de caracterización del concepto de género que, centrado en el género técnico, parte de las definiciones propuestas en el campo de las lenguas para fines específicos (Swales, 1990, Bhatia, 1993, Fortanet et al., 1998) y recoge varios de los factores que intervienen a la hora de caracterizar y definir el concepto, diferenciando claramente este concepto del de tipología textual. En su propuesta define el género técnico a través de cinco parámetros básicos que asumimos para nuestra caracterización del género de sitio web corporativo, al que añadimos la funcionalidad como característica esencial en la evolución y diferenciación de los cibergéneros (Shepherd y Watters, 1999).

En el modelo de la autora, desarrollado en la combinación de lenguas alemán-español, se tienen en cuenta los siguientes parámetros: rasgos convencionales,

---

<sup>58</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.

función textual, elementos de la situación comunicativa, influencia del contexto sociocultural y los elementos intratextuales.

#### 4.2.2.1. Rasgos convencionales

Las convenciones son de esencial importancia en nuestro estudio puesto que son una especie de guía que facilita la comunicación y la comprensión, y en el medio de Internet, la funcionalidad. Gamero, siguiendo a Göpferich (1995), toma la definición de convención desde el punto de vista de la psicología (Lewis, 1969), que la define como aquellas regularidades observadas en la conducta humana en situaciones de cooperación. Las mismas aparecen en un contexto comunicativo cuando se cumplen estos tres requisitos:

- Intervienen varias personas.
- Las situaciones se repiten.
- Las personas persiguen una intención que solo puede conseguirse si todos se comportan de la misma forma.

En el entorno web, las convenciones juegan un papel esencial no sólo desde el punto de vista de textual, sino también desde el punto de vista de la usabilidad. Nielsen y Tahir (2002: 37), en su investigación sobre páginas de inicio corporativas, indican la importancia de las convenciones en los sitios web tanto desde el punto de vista de la producción como desde el punto de vista de la recepción del texto:

[...] 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.

Los autores señalan en este caso que las expectativas del receptor no solo se centran en la estructura discursiva, tipología textual y género, «*generic mental model*», de interés en nuestro estudio sino cómo «funcionan» estos textos, por lo que se centra en un aspecto fundamental de los géneros digitales, su funcionalidad (Shepherd y Watters, 1999).

La existencia de una convención requiere que exista una alternativa posible, puesto que de otra manera no puede existir una convención (Göpferich, 1995). Es precisamente este aspecto el que nos lleva a introducir la metodología de

lingüística de corpus en nuestro estudio empírico y usar criterios estadísticos para definir los rasgos convencionales, al igual que se realizó nuestro estudio anterior sobre los menús de navegación en web corporativas (Jiménez, 2005). Por ejemplo, en inglés se encuentra muy convencionalizado por ejemplo el uso de *Contact us*. En nuestra investigación encontramos que el término más empleado sería «Contacto», con un 49,42% del total, aunque se encontraron 22 alternativas para la denominación de esta página web, todas ellas posibles aunque con distintos grados de convencionalización.

Las convenciones se adoptan al azar, y el mecanismo por el que se adoptan es totalmente arbitrario (Lewis, 1969:70). Gamero (2001) nos comenta:

Siguiendo con Lewis, 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 (Lewis, 1969: 70).

Otro factor de importancia es el reconocimiento de las convenciones tanto del emisor como del receptor, y tiene un lugar esencial en la comunicación. Desde el punto de vista cognitivo de la usabilidad del texto web, Nielsen y Tahir, (2002: 54) señalan la labor facilitadora de las convenciones en la comunicación:

The less you have to wonder about the meaning of shared features, the more you can think about the new features.

En los géneros digitales se pueden reconocer gran cantidad de convenciones, y se espera que más convenciones sigan surgiendo sobre todo en géneros específicos como los web corporativos o institucionales (ibid). La evolución de los cibergéneros es un proceso en continua expansión y evolución (Kennedy y Shepherd, 2005), y de ahí surge la necesidad de realizar nuestro estudio de forma sincrónica: a medida que cambien las funcionalidades de la web, seguirán surgiendo nuevas convenciones, como la aparición de formularios interactivos en las páginas de contacto, que aparecen en la actualidad en el 31,4% de los sitios web del corpus.

La importancia de las convenciones en traductología se basa en parte en la diferenciación entre competencia pasiva y activa de las mismas (Gläser, 1990:72; cf. Gamero, 2001). Una de las premisas de los estudios sobre el género es la necesidad de que los traductores adquieran una competencia activa de géneros de especialidad (García izquierdo, 2002). En nuestro caso, se pueden encontrar claras diferencias entre los diferentes géneros digitales; el conocimiento de las convenciones textuales por la mayor parte de usuarios de Internet puede ser

activo, como el del género *blog* o el correo electrónico, o pasivo, como en los sitios web corporativos, institucionales o portales<sup>59</sup>. Este aspecto se suma a las razones que justifican un estudio pormenorizado de las convenciones de género y tipológicas. De hecho, al igual que en el género, el conocimiento de estas convenciones es esencial en la formación de los futuros traductores (García Izquierdo, 2002; Ezpeleta, 2005), que sin una formación clara en las particularidades discursivas y textuales del texto web, presentan, siguiendo la terminología de Hatim y Mason (1997: 133), un déficit de tipología textual en este campo:

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. [énfasis del autor]

Otra de las razones por la que abogamos por que se incluya el concepto de género en los estudios sobre la textualidad web es que las convenciones de género imponen una serie de restricciones a todos los niveles (Gamero, 2001: 38). En primer lugar, las convenciones textuales varían dependiendo del género o subgénero del que se trate y por lo tanto, no se puede generalizar los resultados prácticos de investigaciones que pretendan generalizar sobre la textualidad web global, y en segundo lugar, este aspecto justifica la utilidad limitada de investigaciones que usan páginas web sacadas fuera de contexto, una metodología que aparece en varios estudios de género web ( Santini, 2005a; Rennau, 2004; Bolaños, 2002).

#### **4.2.2.2. Función textual**

La importancia de las convenciones se conjuga también con el hecho que el emisor o emisores<sup>60</sup> del texto tienen una intención tras la elaboración de la misma (Beaudegrande y Dressler, 1981: 7). Una página de inicio tiene la función expositiva-descriptiva al mostrarnos qué se puede encontrar en el sitio web y a la vez exhortativa o fática, para asegurar que dentro de la marabunta de la web, nos quedemos en la misma. Es esencial por tanto la identificación de la función o funciones textuales, siendo una de la característica principal de muchos géneros

---

<sup>59</sup> Las guías de desarrollo web y de estilo web indican la importancia de contar con un profesional de la edición y escritura web (Nielsen, 2000; Brink et al, 2002), que se justifica por la necesidad de adquirir una competencia activa en este tipo de producción textual.

<sup>60</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.

la multifuncionalidad que presentan los textos (Gamero, 2001: 55, Nord, 1997: 49<sup>61</sup>), amplificada en los géneros digitales. En los estudios de traductología se suele asumir que los textos pueden tener dos funciones, una primaria y otra secundaria (ibid), y no es de extrañar encontrar dos géneros o subgéneros en los que se encuentran dos funciones pero éstas están invertidas. Como ejemplo, los sitios web corporativos se pueden dividir en 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 sería tanto expositiva como exhortativa; la primera para la descripción de productos y procesos de compra etc. y la segunda para el usuario realice una compra. Esta diferencia en la función textual supone la existencia de dos géneros web diferenciados.

En el presente estudio se usan las funciones textuales desde el enfoque pragmático-semiótico de Hatim y Mason (1990, 1997), que los autores denominan focos contextuales. Los autores distinguen entre la función *argumentativa*, la *expositiva* y la *exhortativa* (Hatim y Mason, 1990:153-158).

La función *argumentativa* se centra en 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*. En la primera se presenta el cambio entre la tesis y la contrargumentación de manera implícita o explícita sin conjunción alguna, y el argumento *lopsided* en el que se anticipa la contraargumentación con una concesiva explícita (aunque, a pesar de, no obstante, etc).

La función textual *expositiva* surge del análisis de conceptos, o a partir de la síntesis de ciertos elementos que lo constituyen. Los autores proponen tres variantes:

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

---

<sup>61</sup> «When I refer to “function”, in this context [el texto], I mean “function or set of funciones” because texts are rarely intended for one function only. Various functions usually form an hierarchy of functions, subfunctions etc.»

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. Los autores recurren a explicar las características de cada una para su reconocimiento. Por un lado, la finalidad o foco del autor en la argumentación es conducir la situación de manera que sea favorable para el mismo. Por otro lado, en la exposición conceptual el fin de la misma es presentar un «seguimiento» de manera subjetiva de los aspectos de la escena, sin valoraciones del mismo.

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 esta función los autores señalan la similitud entre la exhortación con alternativa y la argumentación, e indican que Reiss (1976) los engloba en los textos operativos. Señalan 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. Indican 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.

La función textual es una cualidad pragmática que se asigna al texto por parte del receptor en una situación particular y no es algo que este incrustada, o inherente en el texto (Nord, 1997: 49). En localización esto se refleja en la importancia del proceso de Control de Calidad (QA), puesto que se tiene que validar que la «función» del programa que en parte se realiza a través del texto es correcta.

Además el texto web se caracteriza por la constante mezcla de funciones presentes en todo momento en una página web, siendo la multifunción la regla más que la norma (Hatim y Mason, 1990: 129). Los textos por lo tanto suelen presentar una función primaria y otra secundaria (Gamero, 2001: 37), como en el caso de las páginas de inicio mencionada anteriormente. La división entre función primaria y secundaria se puede tomar también en el sentido de Göpferich (1995), que en su clasificación de textos técnicos subraya que la función secundaria se produce cuando un receptor usa el texto para un fin distinto al que fue escrito.

En el contexto de los sitios web, cobra gran importancia la relación entre las convenciones de género y las funciones o focos contextuales que aparecen. La propia naturaleza del medio, basadas en la característica cognitiva por la que los

usuarios escanean y no leer una página entera (Nielsen y Tahir, 2002), 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. En este medio, si el usuario no encuentra lo que espera o no lo entiende, intentará buscarlo en otro sitio web (Krug, 2005), y por lo tanto, el texto no cumplirá su función. De esta manera, un texto web mal localizado corre el riesgo de aparecer no funcional, algo primordial puesto que el texto en cuestión necesita recrear la función o funciones que se le asignan, como nos indica Pym (2004):

the localized document is not called to represent any OT translation but a functional text in the target locale or culture.

#### 4.2.2.3. Elementos de la situación comunicativa

Los géneros digitales se encuentran en una situación comunicativa determinada que está enmarcada tanto por el emisor y receptor, ya mencionados anteriormente (véase 2.3) como por el campo, el modo y el tenor digitales (ver 5.9.1.2). Cada uno de estos factores está relacionados con la producción y la recepción del mensaje, y a su vez son de esencial importancia el proceso de traducción.

Los sitios web corporativos se enmarcan dentro de un nuevo modelo de comunicación que denominamos Modelo de Comunicación de Masas Interactivo (*Interactive Mass Communication*) siguiendo el trabajo de Anna Janoschka (2003: 96 y ss.). La autora señala que una de las particularidades de esta situación comunicativa es la combinación de elementos de comunicaciones de masas (MC) y de comunicación interpersonal (IPC) que se desarrollan en un nuevo medio. El mayor flujo comunicativo del mensaje *online* (MO) se produce unidireccionalmente desde el emisor al receptor en esta comunicación de masas, que en la Figura 8 se representa en azul. Los elementos blancos en el esquema se refieren a la comunicación interpersonal, que se puede realizar entre grupos o entre usuarios, por lo que la autora los denomina usuarios (U). Las flechas blancas recíprocas 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.

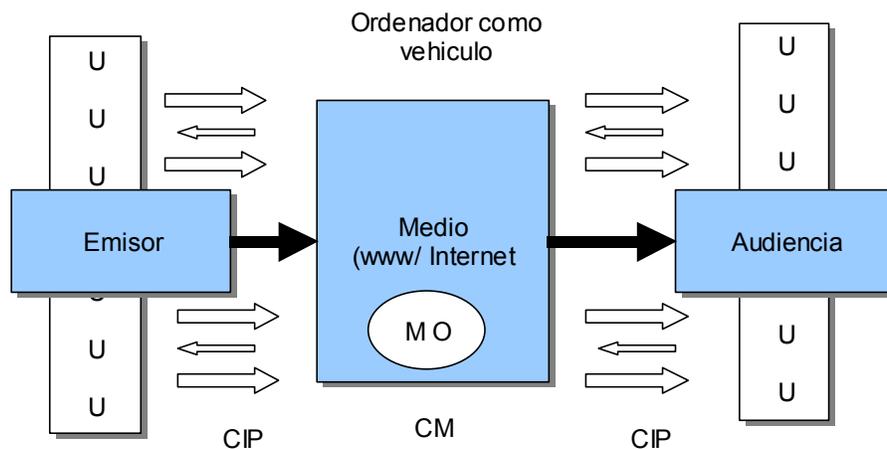


Figura 8. Modelo de comunicación de masas interactivo. Adaptado de Janoschka (2003)

#### 4.2.2.4. Influencia del contexto sociocultural

Los géneros se encuentran inmersos en un entorno sociocultural concreto y se desarrollan por la necesidad comunicativa que se repite dentro de ese entorno (Hatim y Mason, 1990). No obstante, los género digitales empezaron a desarrollarse desde sus inicios en el entorno sociocultural estadounidense (Schäler, 2002). Se ha indicado a menudo que la cultura norteamericana ha influenciado en sobremanera a estos géneros (Bolaños, 2002; Schäler, 2002), pero con la evolución y el desarrollo de Internet en todo el mundo éstos se han convertido en un elemento de diferenciación entre la localización de software y de contenido web, con mayores diferencias en géneros digitales web entre culturas diferentes (Tercedor, 2005). Las diferencias interculturales son así menores en la macroestructura de programas de software, todo ello debido a que se intenta «internacionalizar» la funcionalidad al abstraerla y separarla de lengua alguna (Pym, 2004). En los sitios web, por otro lado, se pueden observar características de los géneros que sólo existen en una cultura y no en el resto, una de las hipótesis de partida del presente estudio y principio básico que da importancia de este factor en los estudios de género (Gamero, 2001:55).

Como ejemplo, en los sitios web anglosajones nos encontramos con la convencionalización de las *FAQs* que llegan a constituirse en género digital propio (Santini, 2005a). Sin embargo, en España sólo aparece este género anglosajón en el 3,4% de los sitios web del corpus, apareciendo otras opciones para asistir al usuario como la ayuda o atención al cliente. Observamos como culturalmente en los web españoles es más común presentar una opción de

ayuda al cliente, con lo que la responsabilidad de ayudar recae sobre el emisor del texto, en vez de las «Preguntas Más Frecuentes» que corresponden a una opción más individualista de ayudarse por si mismo (Singh y Pereira, 2005).

#### **4.2.2.5. Elementos intratextuales**

El presente estudio surgió como continuación de nuestro estudio anterior (Jiménez, 2005) (véase 3.1), en el que se observaron claras diferencias intratextuales en un mismo género, las páginas web corporativas, en un corpus comparable (Baker, 1995) de páginas traducidas y originales en España. Las significantes diferencias que se encontraron por cuanto a frecuencia, variación terminológica, o uso de colocaciones inexistentes en los textos españoles se justifican en parte a las peculiaridades del proceso cognitivo de traducción de este tipo de textos en el que en gran medida, los traductores no incluyeron la superestructura típica, convenciones de textuales y discursivas del género web diferenciado que ha consolidado en España.

Es precisamente en las diferencias intratextuales donde se pueden observar diferencias y se pueden estudiar las convenciones que han surgido en detrimento de las posibles alternativas que se pudieran haber instaurado, puesto que el resto de elementos que caracterizan al género no admiten alterativa (Gamero 2001: 57). Como ejemplo, en el corpus se observa que a pesar de la variabilidad en la terminología empleada para la denominación del bloque «contacto», éste es el término con un mayor grado de convencionalización, un 49% de uso.

#### **4.2.2.6. Funcionalidad**

La funcionalidad es uno de los elementos que caracteriza claramente a los géneros digitales y la base de clasificación en el primer estudio de impacto que se realizó sobre géneros digitales (Crownston y Williams, 1997).

La funcionalidad del texto web se refiere a las posibilidades de interacción que ofrece el nuevo medio, la Internet (Shepherd y Watters, 1999) y se ha caracterizado como el principal motor que genera nuevos géneros (Shepherd y Watters, 1998). La funcionalidad la definimos por medio de la *interacción* y la *interactividad* propia del medio, Internet. 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 de 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. La interactividad, por otro lado, se centra en 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 interacción y la interactividad propias de la funcionalidad son esenciales en el nuevo modelo de comunicación que supone la Internet.

Desde que se comenzó a crear textos digitales se estudió en profundidad la interacción máquina-usuario por medio de la funcionalidad textual, en el que destacó los trabajos de Apple con el libro «*Machintosh Human Interface Guidelines*» (Apple Computers Inc., 1992). En la web, se intentó en un principio recrear la textualidad del texto impreso, pero las ciertas particularidades del medio digital como la lectura en pantalla o la funcionalidad hicieron que los cambios a los géneros existentes, como periódicos o enciclopedias, fueran de tal magnitud que aparecieron nuevos géneros (Yates and Orlikowski, 1992). Así, la funcionalidad fue uno de los elementos que contribuyó a la evolución y distinción de los géneros digitales de los impresos (Crownston y Williams, 1999).

La necesidad de incluir la funcionalidad en nuestra caracterización del género digital surge por la convencionalización de diferentes características de interacción por parte de los usuarios (Shepherd y Watters, 1999). Por ejemplo, los usuarios esperan la inclusión de vínculos en cada resultado de una búsqueda o la inclusión de un enlace de correo electrónico<sup>62</sup> que se active directamente en una página de contacto. Shepherd y Watters (1999) realizaron una investigación sobre la funcionalidad de páginas de inicio<sup>63</sup> en la que encontraron 6 géneros digitales. En líneas generales mencionan que las páginas corporativas presentan como funcionalidad la posibilidad de navegar por medio de enlaces el sitio web y la interacción por medio del correo electrónico.

Cybergenre	Content	Form	<b>Functionality</b>
Homepage	Information about	Introduction	Browsing

---

<sup>62</sup> Del tipo <a mailto:>

<sup>63</sup> De objetivos limitados puesto que incluyeron solamente 97 páginas de inicio en su estudio.

	company-person-institution	Hierarchical images Animated images	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 9. Resumen de contenido forma y funcionalidad de los 6 cibergéneros descritos por Shepherd y Watters (1999).

Dentro de esta clasificación, los sitios web corporativos se encontrarían en los *Homepage*, un supragénero que a su vez se puede dividir en las personales, las institucionales y las corporativas.

#### 4.2.2.6.1. La funcionalidad como criterio clasificador de géneros digitales

Shepherd y Watters (1998) realizaron una clasificación de géneros digitales partiendo de la existencia de un «género digital» inicial y partiendo de la funcionalidad como característica esencial en la evolución constante de éstos géneros. Su clasificación se menciona y apoya en gran número de estudios (Santini, 2005a; Rennau, 2004; Posteguillo, 2003; Bolaños, 2002).

El universo textual digital lo dividen entre los géneros impresos que dieron el salto a la web sin adaptarlos, los géneros *extant*, y aquellos géneros que se crearon específicamente en la web, los *novel* que no existían en el mundo impreso.

Los géneros *extant* se dividen en géneros *replicated*, aquellos en los que textos existentes simplemente se «cuelgan» en la red, tales como los artículos de investigación. Éstos géneros se adaptan a la red puesto que necesitan pasarse a formato .html o .pdf, pero textualmente no presentan características diferenciadas, el medio es simplemente diferente, por lo que desde el punto de vista de la traductología no necesitarían estudiarse independientemente. Éstos géneros no suelen funcionar en la web puesto que no se adaptan a las particularidades del medio (Nielsen y Tahir, 2002). Por otro lado, los géneros *variant* se basan en los géneros existentes, como los periódicos, pero incluyen ciertas funcionalidades del nuevo medio. En la actualidad, la funcionalidad de la mayoría de los periódicos web los convierten en un género *novel*.

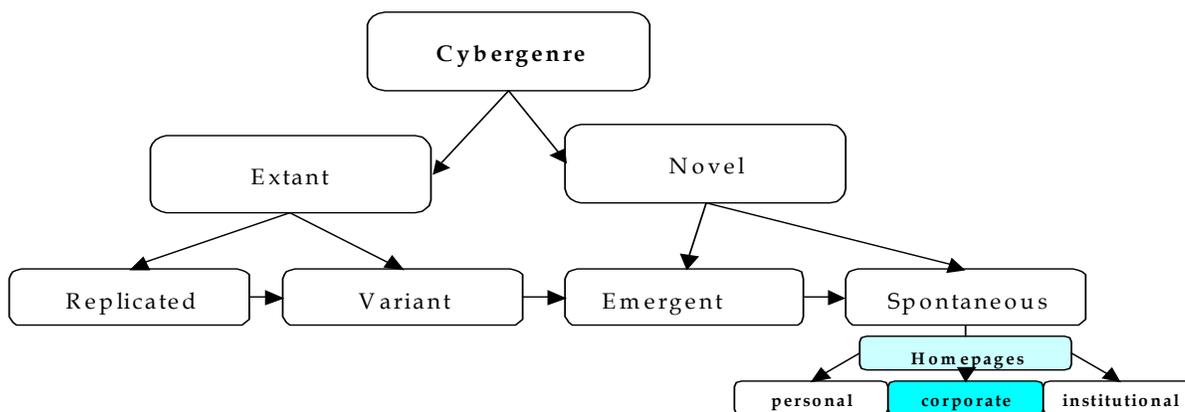


Figura 10. Los cibergéneros. Adaptado de Shepherd y Watters, (1998), y ubicación del objetivo de estudio, combinado con Kennedy y Shepherd, (2005).

Los géneros *novel* son aquellos que no existen en otro medio, según Shepherd y Watters es el nivel de funcionalidad es 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.

Estos géneros se pueden dividir entre los *emergent* y los *spontaneous*. Los emergentes aparecen con la evolución de los géneros *variant* para convertirse en un género diferenciado. Así, los periódicos *online* muestran una estructura y una funcionalidad que en constante evolución se ha ajustado a las particularidades del texto web en pantalla, con la inclusión de elementos multimedia como vídeos o mp3, gráficos Flash animados, la inclusión de foros de discusión, chats etc.,

convirtiéndose en un género diferenciado e independiente del periódico impreso. Por su parte, los géneros *spontaneous* son aquellos 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.

En su clasificación, añaden un paso más de esencial importancia en la comprensión y el estudio de los géneros web, la persistencia o dinamicidad de la información. En su terminología, los géneros pueden ser *Persistent* o *Virtual*. La diferencia principal se basa en que en los géneros *persistent*, las páginas están almacenadas y en cualquier momento pueden ser accedidas por diferentes usuarios de la misma manera. Los *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. que producen páginas web con extensiones .jsp, .aspx o .do<sup>64</sup> etc., por lo que diferentes usuarios pueden obtener textos web diferentes al descargar en el mismo momento el mismo sitio web, o recuperar páginas con resultados de búsqueda diferentes dependiendo de la sesión.

El género de sitio web corporativo objetivo de estudio será por tanto un género novel y espontáneo, no existente en el mundo impreso, y a su vez no constituye la evolución de otro género anterior. El género tendrá además subgéneros que pueden ser *extant*, como los catálogos de productos en formato .pdf, o noveles, como los formularios de contacto, y sus realizaciones pueden ser persistentes o virtuales.

En la comparación necesaria con otros estudios anteriores desde una perspectiva lingüística señalamos que este género va más allá de la simple evolución de ciertos géneros impresos como el folleto publicitario (Fernández, 2003) o el catálogo comercial (Rennau, 2004:43) con los que se ha comparado. La funcionalidad y textualidad de este nuevo género van más allá de una simple evolución de estos géneros impresos apareciendo un nuevo género.

Los propios autores admiten que la clasificación no es exacta y en nuestra opinión, la constante evolución de la tecnología web hace que se desarrollen nuevos géneros web y que se sigan adaptando.

#### **4.2.2.7. Cibergéneros en la web: una constante evolución**

---

<sup>64</sup> Para una lista exhaustiva de los diferentes formatos web, se puede visitar [www.fileinfo.net](http://www.fileinfo.net)

En varios estudios se encuentra un consenso más o menos claro sobre los diferentes géneros que han aparecido en la web, siempre teniendo en cuenta que no existen líneas claras que dividan unos de otros (Shepherd y Watters, 1999), y que los mismos se encuentran en constante evolución (Santini, 2005a). Como se menciona 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, las institucionales (Reim, 2002) y las corporativas, objetivo del estudio.
- Los folletos
- Los sitios de recursos
- Los catálogos web
- Los motores de búsqueda
- Los juegos
- Los contactos (Santini, 2005 a)
- *Quick information delivery* (Santini, 2005a)
- Las *Hotlists*

Para una revisión de los diferentes estudios en géneros digitales desde la perspectiva lingüística, como los web de informática, de universidades, etc. se puede revisar el estudio de Rennau (2004: 39-44).

En la industria de la Internet se encuentra otra clasificación de géneros web desarrollada según el campo del sitio web. La empresa Nielsen-Netratings presenta una categorización cuyo propósito básico es el muestreo de visitantes en cada género web. Esta empresa es la más importante en la investigación de mercado en Internet.

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 11. Clasificación de géneros según la empresa Nielsen-Netratins<sup>65</sup>.

Los géneros propuestos en esta clasificación, que se basa en el campo temático, incluyen portales y motores de búsqueda, deportes, negocios, servicios de noticias, correos electrónicos etc. Esta clasificación se asemeja a ciertas tipologías propuestas que se han criticado por incluir el campo temático como único criterio en su clasificación, y que por lo tanto no son de interés desde el punto de vista traductológico (Hatim y Mason, 1997: 128).

### 4.3. Propuesta de caracterización general de género web de sitio web corporativo

Si tenemos en cuenta y aplicamos los factores de caracterización y definición adaptados de Gamero (2001), podemos definir el género de sitio web corporativo como un prototipo textual convencionalizado que se utiliza en ciertas situaciones comunicativas que se repiten 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 o funciones, 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, la Internet, con la funcionalidad que esta proporciona, y presenta un modo escrito que se acerca al

---

<sup>65</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)

hablado; este género posee unos bloques o secciones y unas características textuales, tipológicas, tipográficas, y funcionales que se encuentran más o menos convencionalizadas. Los receptores son capaces de reconocer estas características convencionalizadas a partir de la forma externa, el medio y la situación de uso, pero para poder tener una competencia activa en el género, por ejemplo como traductores, es necesario realizar un esfuerzo consciente y tener conocimientos básicos de los lenguajes de programación en los que se encuentran inmersos los textos.



## 5. EL ENFOQUE TEXTUAL ¿QUÉ CONSTITUYE UN TEXTO EN UN SITIO WEB O EN UN PROGRAMA DE SOFTWARE?

Hasta la fecha, la mayoría de los estudios sobre localización en traductología han dejado a un lado uno de los pilares fundamentales de nuestra investigación empírica, qué constituye un texto en localización y cómo caracterizarlo desde el punto de vista del proceso de traducción. Los estudios en localización suelen producirse desde la parte técnica del mismo (Quirion, 2003), por lo que se suele caracterizar al texto como «material» (Esselink, 2001), «parte lingüística» (LISA, 2004), «contenido» (Mata, 2005; O'Hagan y Ashworth, 2003), o *information elements* (Lockwood, 2000). La falta de límites que acoten un texto en la web o en un programa de software, la autoría múltiple del texto (Pym, 2004), las peculiaridades de acceso al texto por parte del traductor o la falta de aplicación de un modelo de hipertexto, ha llevado a menudo a esquivar este tema. Existen gran cantidad de estudios traductológicos en nuestro país sobre las particularidades de los textos jurídicos (Borja Albí, 2000), textos técnicos (Gamero, 1998, 2001), cinematográficos (Díaz Cintas, Jorge, 2003), científicos (López Rodríguez, 2001, Tercedor, 1999)<sup>66</sup>, pero no existe ninguno que defina claramente qué constituye el texto web en sí con fines traductológicos. En una investigación en traductología que pretende usar un corpus textual, un paso previo esencial es la delimitación clara de qué constituye un texto, cómo delimitar su contexto, y cómo afectan las peculiaridades de la textualidad web al proceso de traducción.

### 5.1. De texto a hipertexto

El desarrollo de las definiciones de texto y sus características varían en consonancia con la perspectiva de análisis. Estas diferencias hacen que sea bastante problemático definirlo (Janoschka, 2003:159). De hecho, no es solamente el área desde la que se estudie, como la semántica, la pragmática, la lingüística textual o la traductología lo que nos lleva a encontrar diferentes definiciones, el medio también puede contribuir a su clasificación (Trosborg, 1997: xiv). La televisión, por ejemplo, combina textos hablados y escritos, en la web nos encontramos con diferentes modos de textos y elementos multimedia.

---

<sup>66</sup> No pretendemos incluir un listado extensivo de los diferentes estudios traductológicos en diferentes campos.

En traductología se pueden observar diferencias en las características e importancia dependiendo del paradigma desde el que se haya estudiado, como el lingüístico, el sociolingüístico, el lingüístico, el lingüístico textual, el psicolingüístico, el computacional o el crítico (Shreve y Neubert, 1992:12-32). La importancia de texto en ciertas corrientes es central, como en el modelo textual que señala la importancia del texto como punto de partida en su teoría de la traductología (Neubert, 1996; cf. Jung, 2000)<sup>67</sup>.

Los diferentes modelos, teorías y estudios empíricos en traductología suelen asumir la existencia de un texto completo y estable, en el que el proceso de lectura del texto origen por parte del traductor se asemeja en cierta manera a la de un receptor del mismo<sup>68</sup>. De igual manera, los textos usados en experimentos empíricos que intentan explorar ciertos aspectos del proceso de traducción<sup>69</sup> (véase Orozco, 2002) se suelen basar en textos completos, coherentes y cohesivos.

Para estudiar la noción de texto web o hipertexto, se necesita aclarar qué constituye un texto en general, incluyendo los hipertextos impresos como las enciclopedias o diccionarios (Jucker, 2003) y de ahí transferir ciertos aspectos necesarios para delimitar qué constituye un texto en localización o un hipertexto, partiendo de una perspectiva traductológica y sin olvidar el medio, Internet.

## **5.2. Definición de texto**

La lingüística textual ha descrito qué constituye un texto desde diferentes perspectivas. En los años 60 los estructuralistas se centran en las características estructurales de los textos basándose en investigadores como Saussure (1916); Bloomfield (1935) o Chomsky (1965). Las críticas como la falta de atención a las relaciones semánticas o al contexto hizo que surgiera otra corriente centrada en los aspectos comunicativos. La introducción de aspectos pragmáticos supone una descripción del texto en su situación comunicativa y sus aspectos funcionales, incluyendo al emisor y al receptor como elementos cruciales. Desde esta perspectiva surge la definición de texto que ha influido en los investigadores en traductología en los que se apoya el presente trabajo. Beaugrande y Dressler (1981:3) definen texto como:

---

<sup>67</sup> Representante de la Escuela de Leipzig de traductología. Jung (2000) presenta un tratamiento más detallado de esta perspectiva.

<sup>68</sup> Estudios empíricos han demostrado no obstante que la lectura es más lenta y se activan mecanismos diferentes (Macizo y Bajo, 2005)

<sup>69</sup> Ver Orozco (2002) para una relación de investigaciones empíricas en traducción escrita.

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.

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 se considera como tal. En el caso de los hipertextos, son de esencial importancia la coherencia y la cohesión, debido a ciertas particularidades del procesamiento de hipertextos (Janoschka, 2003: 165-167) como la falta de una secuencia textual fija, el procesamiento textual discontinuo o la falta de límites claros en un texto (Storrer, 2002).

En este caso, la coherencia es de especial importancia y cuenta con un gran número de estudios (Foltz, 1996, Fritz, 1999, Storrer, 2002), por lo que se estudiará en más profundidad a continuación.

### 5.3. ¿Qué constituye un texto en localización?

En localización, siguiendo los estudios que definen al texto como unidad mínima de traducción (Neubert y Shreve, 1992), se puede caracterizar al texto digital como a una unidad que se desarrolla o que se presenta al receptor de modo unitario. Pym (2004:17) lo define como: «*a text is quite simply whatever unit is distributed as a unit*». 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.

Los hipertextos difieren en cierta manera del texto en programas de software, puesto que éstos no se puede definir como unidad de distribución, 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. Esta particularidad, dentro del contexto de producción de sitios web en constante actualización (O'Hagan y Ashworth, 2003), se asemeja a la que encontramos en el contexto de la *Information Management*, en la que se la información se distribuye en forma de Objetos de Información, o *IO* (Hofmann, C y T. Mehnert, 2000). Esta concepción se acerca a nuestra propia aproximación al texto web, puesto que el

IO es una colección de información que se identifica como una unidad, y se define por su fin comunicativo, el usuario 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).

De manera preliminar, señalamos que una característica principal del texto en localización reside en que durante el proceso de traducción, los textos y segmentos están organizados (o no) según criterios informáticos y no comunicativos en la estructura interna del programa de software o sitio web (Jiménez, 2004)<sup>70</sup> y podemos definir el texto web como una unidad que se sirve de modo unitario y a la que se accede por medio de un solo dominio, con lo que correspondería con un sitio web, como por ejemplo [www.iberdola.es](http://www.iberdola.es).

## **5.4. El hipertexto**

La forma típica de acceder a información en la WWW es a través de hipertextos enlazados por vínculos. 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 60 (Landow, 1992), se puede observar que el concepto de hipertexto es heterogéneo, ya que no se clarifica si hipertexto se refiere a páginas web, a los diferentes componentes de una página web como puede ser 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>71</sup> (Landow 1992: 4), lo define a mediados de los 60 como:

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

La falta de linealidad, la interactividad y la no impresión en papel aparecen como las primeras características claras del hipertexto. Para Landow (1992, 1997) el hipertexto representa una textualidad diferenciada, organizada como una red

---

<sup>70</sup> Trabajo de investigación no publicado, dirigido por Ricardo Muñoz Marín en el programa de doctorado en Procesos de Traducción de la Universidad de Granada.

<sup>71</sup> Dentro de su investigación en a mediados de los 60, *Xanadú*, que no se llegó a implementar en ordenadores.

abierta, que no es lineal y sin un centro claro, en contraposición con la textualidad lineal del texto impreso.

Tanto la estructura de los hipertextos como su efecto en la recepción y procesamiento es un aspecto esencial en el proceso de traducción. Una de las características pragmáticas y estructurales básicas que separan los hipertextos de los textos impresos es la linealidad y sus implicaciones en la producción y recepción textual (Fritz 1999: 222). En los hipertextos, la aparente falta de linealidad es en realidad una multilinealidad (ibid), puesto que el emisor del texto configura de antemano la estructura posible del sitio web. En este aspecto, la labor del diseñador de un sitio web se asemeja más a la de un documentalista que a la de un escritor (Storrer, 2002), ya que su labor se basa en la organización de una estructura que permita su actualización y adición de material.

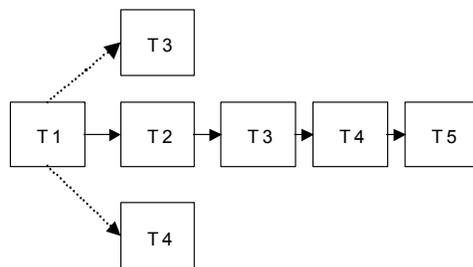


Figura 12. Modelo de texto lineal. Adaptado de Fritz (1997: 222).

El hipertexto, a diferencia del texto lineal, es una red de elementos textuales que suele tener una estructura jerárquica, siendo 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 la ruptura con la estructura que representa la linealidad de los textos. Anna Janoschka (2003: 172) revisa las propuestas anteriores y presenta un esquema de hipertexto en el que se observa la estructura de navegación de un sitio web en el contexto de la WWW, en el que señala la importancia de la multilinealidad desde la perspectiva del usuario, la *open-endedness* de los hipertextos, la inclusión de ciertos elementos textuales<sup>72</sup> independientes, o la disminución de la importancia de las páginas de inicio, puesto que se puede llegar a un sitio web por medio del *deep linking* (Nielsen, 2000:179), es decir llegando directamente a una página interna del sitio web. Así,

---

<sup>72</sup> En el caso de Janoschka, la publicidad web.

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.

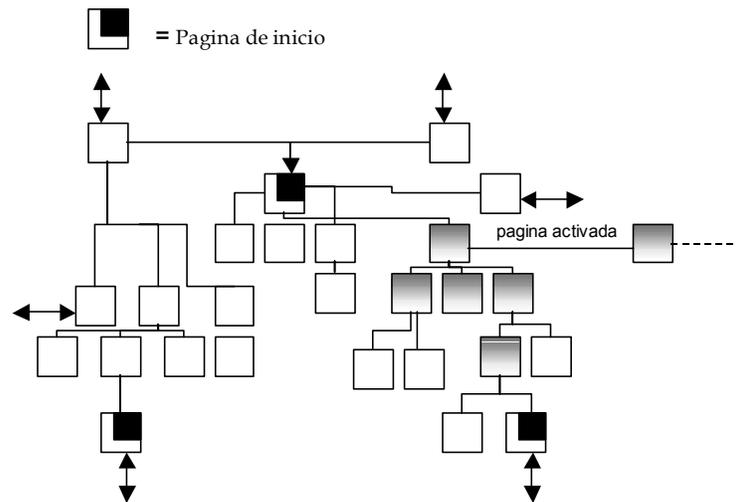


Figura 13. Modelo de la estructura de un hipertexto en el contexto de la WWW. Adaptado de Janoschka (2003: 173).

## 5.5. Tipos de hipertextos

Nuestra investigación se basa en la recopilación de un corpus de textos completos, por lo que se hace necesario acotar esta aparente falta de límites de los hipertextos. Storrer (1999, 2002) 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>73</sup>:

- En primer lugar presenta los *E-textos*<sup>74</sup>, textos organizados de forma secuencial disponibles en la WWW. Estos textos suelen ser copias de textos impresos y por tanto con semejantes mecanismos de cohesión como pueden ser tesis, artículos de periódico etc. y se suelen publicar en formato.pdf o .html por ser una forma económica y eficiente de distribuir

---

<sup>73</sup> En este caso sobre la coherencia en hipertextos (Storrer, 2002)

<sup>74</sup> Janoschka (2003: 171) indica que otros autores (Fritz, 1999) denominan los E-textos como *Textual Elements*.

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 mismo proceso de traducción de estos mismos textos impresos<sup>75</sup>. Estos textos formarían así parte del género *extant* (ver 4.2.2.6.1) presentados anteriormente.

- Los *hipertextos* son textos publicados electrónicamente con una estructura no lineal, con una función textual reconocible y una consistencia temática, como puede ser un sitio web, son además abiertos ya que los emisores pueden añadir más nodos. Estos hipertextos se diferencian de los anteriores en que su contenido se encuentra estructurado en nodos<sup>76</sup> 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 de sistema de hipertexto. Es precisamente su medio, la Internet, lo que diferencia a estos hipertextos digitales de los textos no secuenciales que se denominan hipertextos impresos, tales como diccionarios, enciclopedias o periódicos (Jucker, 2003).
- Por último, la *hiperweb* interrelaciona a los hipertextos y E-textos por medios de vínculos. La totalidad de la WWW se puede denominar a gran escala una hiperweb. La autora indica que a menor escala una hiperweb se puede caracterizar por el tema, o institución, como el caso de la hiperweb de la Universidad de Granada, [www.ugr.es](http://www.ugr.es), que se constituye de gran cantidad de hipertextos, los sitios web de diferentes departamentos, facultades, asociaciones de estudiantes etc. enlazados entre sí. En este particular, 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). Estos sitios web corporativos representan uno de los mejores ejemplos de hipertextos, puesto que casi todos los hipervínculos son internos<sup>77</sup>, hacia nodos en el mismo

---

<sup>75</sup> Aun cuando argumentamos que la presencia de código de cualquier tipo en el documento puede impactar el proceso de traducción.

<sup>76</sup> Los investigadores en hipertexto denominan «nodos» a los diferentes sub-textos enlazados por medio de hipervínculos.

<sup>77</sup> Los hipervínculos pueden ser internos, hacia el mismo sitio web, o externos, hacia otros sitios web. A su vez, los hipervínculos se componen de un *Trigger* o iniciador, de un enlace y de un *Target* o 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.

hipertexto. Los hipervínculos externos, los que se dirigen hacia otros hipertextos en la hiperweb global, son casi inexistentes puesto que éstos conllevan el abandono del hipertexto por los usuarios, algo contrario al objetivo del mismo (Nielsen y Tahir, 2002).

La clasificación de Storrer surge desde la ciencia de la información y sólo incluye como funcionalidad de los hipertextos los vínculos y la opción de búsqueda. Desde nuestra perspectiva traductológica, es necesario incluir la existencia de segmentos textuales basados en la interactividad, como pueden ser los mensajes de error o los que aparecen al posar el puntero del ratón sobre un elemento<sup>78</sup>.

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

Son varias las características generales que delimitan claramente al hipertexto en su situación comunicativa. A continuación esbozamos éstas siguiendo a O'Hagan y Ashworth (2003: 13):

- El texto se lee en la pantalla y no en papel, por lo que se dificulta la lectura (Pym, 2004), en general un 25% más lenta (Nielsen y Tahir, 2002).
- 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 la adaptación cultural de los mismos es parte integrante del proceso de localización (Tercedor, 2005).
- Los posibles receptores no son específicos y sino una población de hablantes 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).
- El hipertexto<sup>79</sup> se puede leer en un orden diferente, y por lo tanto, en contextos diferentes dependiendo de los hipervínculos que se siga. De

---

<sup>78</sup> Por medio de la etiqueta <onmouseclick>

<sup>79</sup> El estudio de O'Hagan y Ashworth (2003) no define que constituye un texto o su contexto en localización.

igual manera, durante el proceso de traducción el traductor accede al texto de manera no lineal.

- 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), como por ejemplo, una promoción que caduca a los tres días. La industria de la localización, respondiendo a los retos que supuso la organización de 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>80</sup> (LISA, 2006; Yunker, 2003: 355; Brooks, 2000). Según LISA (2006:4), 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 organizations 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.

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 los 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<sup>81</sup> (LISA, 2004: 25). Estos sistemas se asemejan a los *Content Management Systems*<sup>82</sup> en que no están diseñados específicamente para el proceso de traducción (LISA, 2006: 2), aunque poseen varias funcionalidades dirigidas a la localización, como almacenar el contenido o el texto en segmentos en una o varias lenguas y producir páginas web diferentes según las preferencias del

---

<sup>80</sup>Entre otros [www.glides.com](http://www.glides.com), [www.sdlint.com](http://www.sdlint.com), [www.globalsight.com](http://www.globalsight.com)

<sup>81</sup> Incluso dentro de una misma lengua, es necesario tener en cuenta la locale. Lisa (2004:25) presenta como ejemplo en un texto en automoción el caso: «look under the hood» (EEUU) versus «look under the bonn» (Reino Unido).

<sup>82</sup> Entre otros, [www.microsoft.com/cmsserver](http://www.microsoft.com/cmsserver), [www.broadvision.com](http://www.broadvision.com), [www.documentum.com](http://www.documentum.com).

usuario (Yunker, 2003:354). Estos CMC separan además el contenido de la forma, es decir, el texto del código y lo insertan de forma dinámica cuando es necesario.

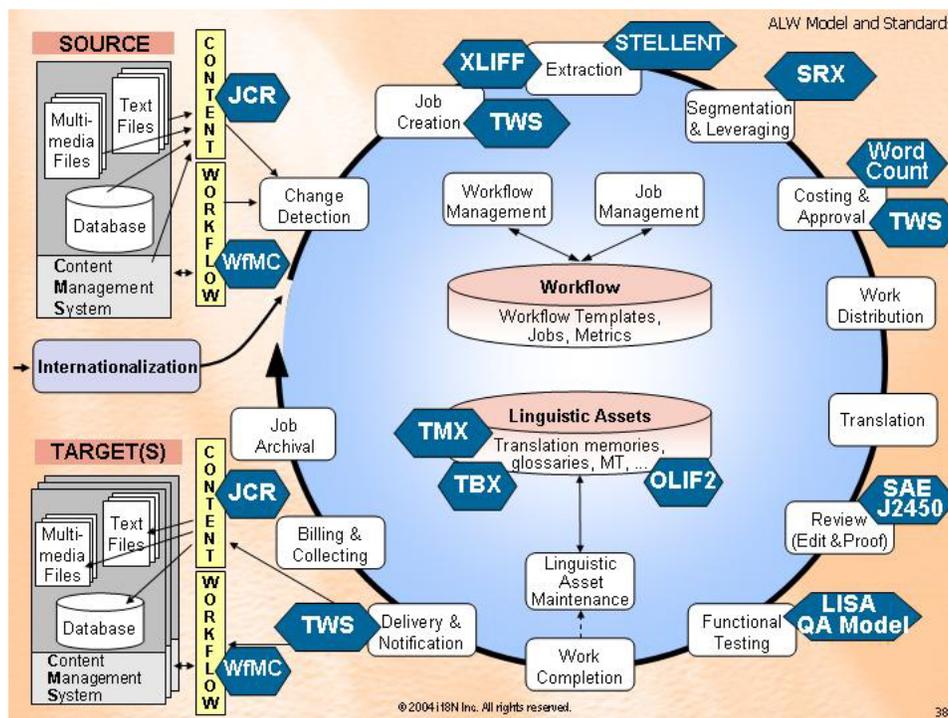


Figura 14. Esquema de Global Content Management System<sup>83</sup>.

- El texto se puede «adaptar» a los receptores de la lengua Meta incluyendo cambios de contenido. Esto entroncaría con la teoría del escopo (Reiss y Vermeer, 1984), ya que en ciertos casos no es solo el contenido sino la función la que se necesita adaptar. En este caso, el escopo determinaría el grado de localización al que se llega (véase 2.5), influyendo de esta manera en el proceso de traducción.

## 5.7. Coherencia y cohesión

Son varias las condiciones que se deben cumplir para que un texto sea definido como tal (Beaugrande y Dressler, 1981:3). En el presente estudio nos centramos

<sup>83</sup> <http://www.i18n.ca/publications/GlobalizingContentModelStandards.jpg>

en el texto como una unidad coherente y cohesiva (Hatim y Mason, 1990: 178) que como unidad tiene una función comunicativa (ibid).

La cohesión la definen Beaugrande y Dressler (1981:3) 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.

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 (1981:4):

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.

La coherencia se apoya en las relaciones formales y sintácticas que nos 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: 84) un texto «tiene sentido» por la «*continuity of senses*», la base de la coherencia, entre los conocimientos que se activan por las expresiones del texto. Durante el proceso de traducción, esta «continuidad del sentido» puede no producirse, ya que los segmentos textuales se encuentran organizados o no según criterios informáticos y no textuales, con lo que la coherencia textual se produciría de forma diferenciada entre un receptor final y el traductor durante su tarea de traducción. Varios autores que han estudiado la textualidad web señalan a la coherencia como el criterio más importante para establecer un texto ( Fritz 1999; Janoschka 2003:167), y el criterio principal en su producción y comprensión (Fritz, 1999:221).

La coherencia y la cohesión del texto web están caracterizadas por la falta de linealidad propia del lenguaje con hipervínculos (Crystal, 2001), que como hemos comentado, es más apropiado denominar multilinealidad (Janoschka, 2003: 172). La importancia del receptor en la construcción de la coherencia es esencial: la cohesión aparece de manera explícita en el texto, al ser una característica objetiva del 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). En el caso de los hipertextos interactivos la coherencia se produce desde el receptor que elige su propio camino y es *process-oriented* (Janoschka, 2003).

En los hipertextos, la coherencia y cohesión se caracterizan por la necesidad de dividir los textos en subtextos para ajustarlos al medio que los presenta en una pantalla de espacio limitado (O'Hagan y Ashworth, 2003: 13). Según la perspectiva de estudio estos subtextos se suelen denominar nodos (Landow, 1992), lexia (Barthes, 1977), elementos textuales (Fritz, 1999), hiperdocumentos (Crownston and Williams, 1999) o hipertextos/ E-textos (Storrer, 2002). La coherencia por lo tanto en un sitio web global surge por medio del uso de hipervínculos internos relacionados semánticamente, siendo esencial que el receptor comprenda la relación entre el hipervínculo y el nodo que aparezca para que siga la continuación del sentido (Beaudegrande y Dressler, 1980: 84). Las relaciones semánticas que se producen en los hipertextos se estudian en profundidad por la semiótica del hipervínculo (Neumüller, 2004).

La cohesión global de un sitio web se consigue además por la inclusión en la totalidad de las páginas de menús de navegación, y por una presentación tipográfica y visual que forma parte intrínseca del texto (Göpferich, 1995). Así, las páginas web individuales constituyen la unidad de presentación en pantalla debido a las peculiaridades cognitivas de la lectura, y pueden ser leídas por sí mismas y accedidas por medio de *deep linking*. Las páginas constituyen por tanto un subtexto que depende de las relaciones estructurales del sitio web donde se encuentren para adquirir la coherencia y cohesión global. En varios estudios anteriores las páginas web aisladas han sido usadas como unidad textual (Biber, 2005, Santini, 2005a; 2005b; Bolaños, 2002; Rennau, 2004), pero desde un punto de vista traductológico no se pueden estudiar independientemente, puesto que carecen de la coherencia y cohesión dadas por su pertenencia a un hipertexto con su superestructura y macroestructura global. En el contexto de la WWW, este aspecto es esencial puesto que un hipertexto no cohesionado por un menú de navegación o vínculos puede ser recibido como no verídico, (Frogg, 2002), y por lo tanto, no funcional.

## **5.8. El sitio web como texto**

El estudio de las particularidades y mecanismos de coherencia y cohesión en este caso nos lleva a considerar al **sitio web o hipertexto completo como texto unitario**, incluyendo tanto los aspectos gráficos, visuales y tipográficos, así como

los textos acompañantes (Göpferich, 1995). En este caso los entendemos como aquellos textos que el receptor no lee, puesto que aunque cada lectura será única debido a la estructura de hipervínculos (Janoschka, 2003). Según Göpferich (1995), los textos que lo acompañan son esenciales para llevar a cabo la función comunicativa del mismo. 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 lector inició su lectura. No obstante, en ciertos tipos de géneros digitales, como los periódicos o los *blogs*, los subtextos se pueden considerar *E-textos* (Storrer, 2002), y por lo tanto pueden tener una función pragmático-comunicativa por sí mismos.

En los albores de Internet la unidad textual la constituían las páginas web, ya que el contenido web se concentraba en una página o pantalla única (Mata, 2005). Estos inicios siguen estado presentes en ciertas investigaciones cuyo objetivo de estudio es la «página web» como unidad (Santini, 2005a; Biber, 2004; Rennau, 2004; Bolaños, 2002), aunque en nuestra opinión esto no se corresponde con desarrollo actual de los textos web. La evolución constante de los mismos hace que hoy en día, por muy pequeño que sea un sitio web, éste conste de varias páginas enlazadas.

Por otro lado, los aspectos gráficos y tipográficos unifican el texto y hacen que cada página del mismo adquiera cohesión global. Una de las principales indicaciones en globalización es la necesidad de una identidad corporativa gráfica y tipológica que dé unidad a todas las páginas localizadas de los sitios web (Yunker, 2003; Hunt, 1995). Además, al igual que los textos técnicos, los gráficos, vídeos, tablas, diagramas, iconos, y colores son parte integrante del hipertexto (Göpferich, 1995), de importancia durante el proceso de localización (O'Hagan y Ashworth, 2003). Como ejemplo, es muy corriente localizar la parte textual de un icono sin tener en cuenta que la metáfora que da lugar al mismo puede partir de un contexto anglosajón<sup>84</sup>.

De igual manera, la importancia de la funcionalidad como elemento esencial de los hipertextos (Shepherd y Watters, 1999) hace necesaria la inclusión como parte integrante del texto de las etiquetas html o xml. (Santini, 2005a), las hojas de

---

<sup>84</sup> Tal como la casa de Internet Explorer que indica la pagina principal, «home» en inglés , que en terminología de la semiótica de Pierce pasaría de ser un signo icónico a un signo simbólico en el que la relación *representamen* (el icono en sí) con el *objeto*, lo que el signo representa en realidad, es puramente convencional.

estilo CSS o XSL<sup>85</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, indican la función de segmentos textuales como la importancia del mismo con de las etiquetas como <H1>, la posición o función del segmento en pantalla con <alt>, <title> o <select>, o la presencia de un *script* que indique interacción (véase 9.2.2).

### **5.9. El contexto como variable necesaria en la definición del texto en localización**

Durante el proceso de traducción en localización, el contexto se presenta como variable esencial en la descripción del texto web, ya que el contexto al que tiene acceso el traductor como lector del texto original es totalmente diferente al del receptor de un sitio web.

En la presente investigación, usamos un enfoque comunicativo e integrador de la traducción (Hatim y Mason, 1990, 1997; Hurtado, 2001), que indica la traducción como operación textual que va más allá del plano de la lengua, y nos permite visualizar a la traducción en sus tres perspectivas: como operación cognitiva, como operación textual, y como evento comunicativo. El enfoque textual de Hatim y Mason (1990, 1997, 2004) define al contexto en tres dimensiones interrelacionadas: la comunicativa, la pragmática y la semiótica. Todas son de gran ayuda para delimitar qué constituye un texto en localización.

---

<sup>85</sup> *Cascading Style Sheets* o *XML StyleSheet Language* que presentan la configuración tipográfica de una página o sitio web.

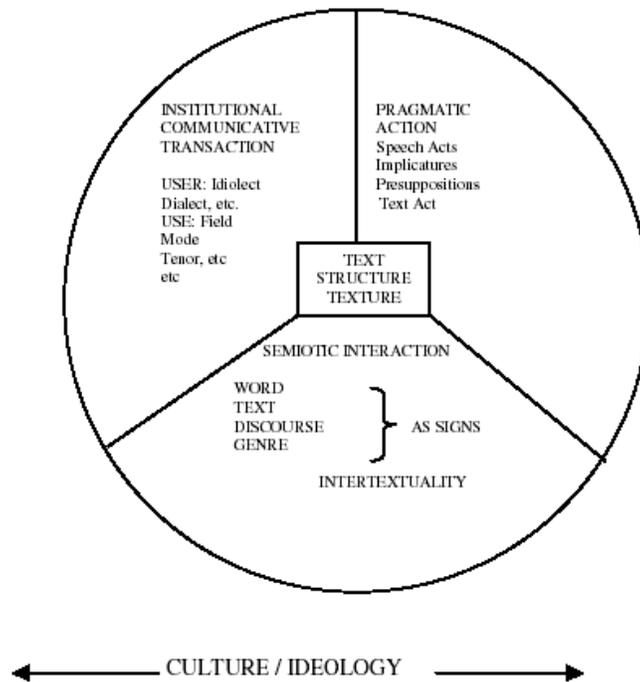


Figura 15. Esquema de las tres dimensiones del contexto ( Hatim y Mason 1990: 58)

La dimensión pragmática del contexto tiene que ver con posibilidad de «*do things with word*» (ibid: 57), es decir, en la intención que persigue el emisor del texto. En localización, la intención que pretende el emisor o emisores del texto es de esencial importancia. De esta manera, durante el proceso de producción y recepción de un texto, surge una función o conjunto de funciones<sup>86</sup>, que a un nivel muy general se puede identificar por cuanto a una tendencia en *monitor* o en *manage* una situación comunicativa dada<sup>87</sup> (Hatim y Mason, 1997: 29-30). El encontrar la función permite determinar qué se pretende al emitir un texto, y los autores identifican tres funciones básicas mencionadas anteriormente, la argumentativa, la expositiva y la exhortativa (véase 4.2.2.2).

En el tratamiento de la dimensión pragmática en un hipertexto es necesario resumir ciertos aspectos mencionados con anterioridad en el trabajo. Por ejemplo, en localización los segmentos del texto meta necesitan mantener la

---

<sup>86</sup> O foco contextual como lo denominan Hatim y Mason.

<sup>87</sup> « Situation monitoring will be performed “if the dominant function of the text is to provide a reasonably unmediated account”. Situation managing, on the other hand, takes place when “the dominant function of the text is to guide the situation in a manner favourable to the text producer’s goals» (Beaudegrande y Dressler, 1981: 162).

función del texto origen, algo que en la traducción general no es algo automático (Nord, 1997). Pero este caso en cuestión, la internacionalización pretende precisamente que el texto origen tenga una función que se pueda abstraer y que se separe del texto a traducir. La cuestión que surge es si es posible el mantenimiento de la función textual en localización. Pym (2004 : 63) habla del concepto de funcionalidad en su estudio de la equivalencia funcional como algo que se puede separar del lenguaje y se puede recrear en otros idiomas, separando la equivalencia general de la «*internationalization-based equivalence*». La diferencia principal es que en los estudios clásicos la equivalencia se estudiaba de lengua a lengua, y ésta última es un movimiento de una lengua hacia muchas más. Según el autor, esta equivalencia basada en la internacionalización se fundamenta en que no se trata de culturas en general sino una situación profesional muy delimitada por los límites del programa de software o sitio web:

internationalization, the whole localization is concerned with narrow professional locales. Translation equivalence is traditionally concerned with large scale social entities. (ibid: 63).

La internacionalización pretende que independientemente del receptor o texto, se produzca esta equivalencia lo más económicamente posible, y que la traducción se reduzca a una labor menor dentro de la localización, con el dudoso objetivo de que el producto sea igual que los producidos en el país: “*the look and feel of locally made products*” (LISA, 2003:5).

La dimensión semiótica del contexto se refiere a la capacidad de los textos de actuar como signos y de relacionarse, tanto con los interlocutores como con otros textos, a través de la intertextualidad. En el esquema general de Hatim y Mason una ocasión social se realiza a través del género, que a su vez se muestra en el discurso para realizarse finalmente en el texto. El texto presenta consecuentemente restricciones que son dadas por su pertenencia a un discurso, a un género y a una situación social dada. En esta dimensión pertenecen los elementos intratextuales que son de interés en nuestro estudio: la coherencia y la cohesión de los textos web y otros elementos como la intertextualidad.

La dimensión comunicativa se basa en las características del intercambio comunicativo entre los emisores y receptores del mismo, y por su importancia en la textualidad web se tratará en más detalle a continuación:

### **5.9.1. La dimensión comunicativa del contexto**

La dimensión comunicativa se centra en las características de la transacción entre un emisor y un receptor que son hablantes de una lengua y que se expresan en un dialecto determinado. Así, dentro de esta dimensión se distinguen aspectos relacionados con los usuarios de la lengua, como es el dialecto, y otros con los usos de la misma, 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 tendría una importancia doble en los sitios web: por un lado, las empresas suelen decidir sobre la terminología 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 globalización (Yunker, 2003), la industria intenta escribir en un lenguaje lo más neutro posible<sup>88</sup>, alejado del ideolecto<sup>89</sup>.

### 5.9.1.1. Registro

El registro reclama su importancia en la localización de páginas web puesto que como indica Tercedor (2005), «el texto multimedia en todas sus dimensiones muestra claras diferencias interlingüísticas en lo que a registro se refiere». El propio autor de este trabajo observó diferencias claras en entre el registro de las páginas web traducidas al español y las producidas originalmente en España, con una mayor formalidad en el lenguaje de los menús de navegación en las producciones originalmente producidas en español (Jiménez, 2005).

Para Hatim y Mason, que parten de Halliday, el registro se identifica con los usos de la lengua (1997, 1990: 55). Éste concepto trata de explicar por que varía la lengua que usamos en situaciones diferentes (Halliday, 1985), y de ahí que se defina por sus situaciones de uso a través de la interacción de tres factores: el campo, el modo y el tenor:

...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)

---

<sup>88</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>89</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 ideolectales puede ralentizar la navegación (Krug, 2005)

Los autores incluyen en su esquema de registro la pragmática y la semiótica, estableciendo una relación entre campo y género, de especial importancia en nuestro estudio. El registro impone unas restricciones por cuanto al léxico y a la sintaxis del texto (Trosborg 1997: 10).

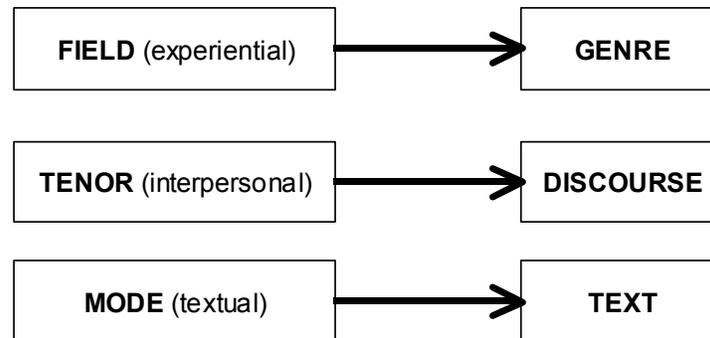


Figura 16. Fuente, Hatim y Mason (1990:75). Pragmática y Semiótica del contexto.

El campo, en general, se asocia con el campo temático, el legal, científico, técnico (Hatim y Mason, 1990), en sus palabras es «*what is going on*» (ibid: 48). 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 o publicidad de sus propios productos, es decir, el género web se suele caracterizar por la presencia de varios tipos textuales y géneros discursivos, incluso en la misma página. Son muchos los estudios sobre el modo y tenor digitales, por lo que requieren un estudio en mayor profundidad.

### 5.9.1.2. Modo y tenor digitales

La Internet es un medio de comunicación especial. Desde los años 90 se ha observado en la lingüística el desarrollo de unas características diferenciadas en el *ciberlenguaje* o *netlenguaje* (Crystal, 2001). Es de especial interés cómo el modo y el tenor se producen forma distintiva en la web, apareciendo lo que se ha denominado el «tenor» y «modo» digitales (Posteguillo, 2003). El «tenor digital» está caracterizado por una reducción en el nivel de formalidad a la hora de interactuar por medio de la red, tanto en correo electrónico (Giménez, 2000; Lan, 2000; cf. Posteguillo 2003) como en otros géneros como las páginas web

personales, la publicidad en internet (Fortanet, Palmer y Posteguillo, 1999) o las páginas web corporativas (Jiménez, 2005).

Hatim y Mason (1997) indican al tenor como el elemento más importante, y en el estudio del tenor digital es necesario indicar el anonimato general que caracteriza los textos en localización (Pym 2004: 8), «*software es anonymous technology speaking to anonymous markets*». En el caso de los sitios web, la necesidad de identificar un emisor es criterio necesario en ciertos géneros <sup>90</sup>, y en nuestro caso, la empresa se presenta como emisor y se comunicaría con el receptor como tal. Así, debido a las particularidades del registro en español, el tenor en estos sitios web se caracteriza por un grado de formalidad mayor que el que se encuentra en páginas anglosajonas (Jiménez, 2005; Rennau, 2004; Posteguillo, 2003: 101-102).

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 la 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). Ciertas características del texto web que lo acercan al oral emanan de los estudios cognitivos de lectura en pantalla, que señalan aspectos como brevedad, claridad y la necesidad de ser conciso para que el texto sea efectivo (McGovern et al. 2002). Es necesario indicar que el modo oral puede aparecer de igual manera 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).

### 5.9.1.3. Registro y género

La relación entre género digital y registro es esencial puesto que en la hibridización característica de este género se presenta una relación de interés entre ambos. Un registro puede estar presente en gran cantidad de géneros, e igualmente, un género se puede expresar a través de varios registros. Siguiendo el estudio de Trosborg (1997: 10-11):

a particular genre may cut across a number of registers. [...] Genres are subordinated to registers only in the sense that one register may be realized through various genres. Conversely, one genre may be realized through a number of registers

---

<sup>90</sup> Dependiendo del género, muchos requieren que el autor se identifique, tales como las páginas personales, las páginas de contactos, los blogs, emails.

Estamos de acuerdo con la autora, puesto que un género o subgénero, como el formulario web, se puede llevar a cabo por medio diferentes registros, y un registro, como el jurídico especializado, se puede presentar en diferentes géneros o subgéneros como los avisos legales, los contratos de acceso a sitios web etc.

Los registros a su vez son en principio ilimitados, al igual que los géneros:

registers comprise an open-ended set of varieties (or styles) of language typical of occupational fields. [...] Taking register in its narrow sense of occupational field.

En esta definición de registro de Trosborg (1997), que se aproxima a la de Hatim y Mason, se establece además la relación entre registro y la prototipicidad que aparece en las áreas profesionales.

#### **5.9.1.4. El registro en Internet**

Son pocos los estudios monográficos que se centran en el estudio del registro en la web. Biber (2004) nos presenta una tipología de registros por medio de un estudio de páginas web que realiza con análisis de *cluster* de parámetros establecidos siguiendo sus estudios en lingüística de corpus. En el estudio se usan dimensiones similares a las que estableció en su tipología de textos en inglés (Biber, 1989). En sus resultados preliminares presenta una propuesta con cuatro posibles registros (Biber, 2004):

- *Personal, involved narration*
- *Persuasive, argumentative*
- *Advice??<sup>91</sup>*
- *Abstract-Technical discourse*

No obstante, es necesario indicar que Biber (1995:9) identifica registro en 1995 con lo que denominaba género en 1988 y lo justifica con que la diferencia entre ambos es muy sutil y abstracta. De esta manera, identifica el registro como término general asociado con los aspectos de la variación en el uso, en sus palabras «*the varieties of the language the we use in different situations are referered to as registers*» (Biber, 1998:2). Indica además la importancia de la descripción de los

---

<sup>91</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, 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.

mismos «*and describing the characteristics of these registers is an important area of study*» (ibid).

Otra de las características esenciales que señala el autor es la multidimensionalidad de los registros, que se hace obvio puesto que durante el transcurso de un día, todos hablamos y escribimos unan gran variedad de registros (ibid: 135). Esta multidimensionalidad se refiere a las dimensiones de análisis que propuso en su investigación sobre la identificación de tipos textuales (1988, 1989, 1995, 2004).

## **6. PRESENTACIÓN DE LA INVESTIGACIÓN EMPÍRICA**

El presente estudio tiene como objetivo el estudio de la textualidad web en los sitios web producidos en España desde una perspectiva traductológica. En la revisión de la literatura sobre localización encontramos que uno de los objetivos del sector es que los productos localizados se reciban como producidos en cada *locale*, (LISA, 2003:5), pero la industria no ha realizado estudios amplios sobre qué particularidades textuales, tipológicas, pragmáticas y de género tienen los sitios web españoles, objetivo del presente estudio. Así, en un primer lugar nos centramos en la inclusión del concepto de género en la localización de sitios web como elemento determinante de las restricciones textuales que se presentan en los diferentes niveles en cada cultura (Hatim y Mason, 1990) y como potenciador de convenciones textuales que son claves en la eficacia comunicativa de los textos meta en su cultura, acentuada en este caso especial debido a las particularidades de la situación comunicativa.

El componente empírico de la investigación se centra en:

1. Recopilar un corpus de *hipertextos* representativos del género textual objetivo del estudio, investigando las limitaciones en la recopilación de corpus web.
2. Análisis cuantitativo del corpus que sirva para la caracterización prototípica de texto objetivo de estudio.
3. Análisis del corpus desde el punto de vista del género para describir los elementos convencionales y los bloques comunicativos que los componen. A su vez, el análisis servirá como base de «control» para estudios posteriores de distinta índole dependiendo del futuro del estudio.
4. Análisis empírico del género web para obtener la superestructura prototípica del mismo.
5. Caracterización de los subgéneros y bloques comunicativos de este género web determinado.
6. Análisis estadístico de la terminología empleada para en la denominación de los bloques y secciones del género.

7. Análisis del tenor digital o el grado de formalidad en los sitios web corporativos españoles.

## 7. METODOLOGÍA DE LA INVESTIGACIÓN EMPÍRICA

El centro de la investigación se basa pues en la recopilación y análisis del corpus *monocale* de sitios web siguiendo la metodología de la lingüística de corpus<sup>92</sup>, procediendo a su recopilación de manera sincrónica. Nos centramos además en la recopilación de textos web que pertenecen a un solo cibergénero, los sitios web corporativos, tal y como se caracterizan en el apartado 4.3.

A continuación se realizará un análisis cuantitativo del corpus utilizando la herramienta de análisis Webbudget, que proporciona estadísticas sobre los elementos textuales de interés desde el punto de vista traductológico.

Una vez analizado el corpus cuantitativamente, se procederá a caracterizar los distintos tipos de páginas web y elementos textuales que componen los hipertextos del sitio web corporativo.

Se procederá a realizar un análisis manual de la totalidad de los enlaces hacia páginas web independientes que se han identificado con los movimientos o bloques característicos de los géneros web (Askehave y Nielsen, 2005), identificando la terminología empleada en la denominación de tal bloque. La necesidad de realizar el análisis manual viene dado por la codificación de ciertos elementos textuales en gráficos o en formato Flash en los menús de navegación, con lo que no se podrían obtener de manera automática (Santini, 2005b). De esta manera, se podrá elaborar el grado de prototipicidad de los diferentes bloques y secciones por medio de las estadísticas de aparición en el corpus en general, y de igual manera se estudiará estadísticamente el grado de convencionalización de los términos empleados en la denominación de bloques y secciones.

A partir del corpus general se separarán en futuras investigaciones varios subcorpus: el análisis del género proporcionará el elemento de subdivisión del corpus en sus bloques y secciones para un estudio comparado de la tipología textual empleada en diferentes lenguas. En este caso, se podrá realizar un análisis en los subcorpus con programas como *Worthsmith Tools*, ya que el diferente tamaño de los sitios web completos y la repetición interna de diferentes segmentos textuales en los textos web no distorsionarán las estadísticas.

---

<sup>92</sup> Para una revisión de los estudios en traductología que usan Lingüística de Corpus, ver Kenny (2001)

## 7.1. Herramientas de Análisis

En la metodología empleada en la investigación se emplean las distintas herramientas de análisis:

1. *Httrack*. Un programa de descarga de sitios web completos que se establecen con unos parámetros predefinidos. Este programa puede descargar de manera sincrónica un gran número de sitios web, basándose en el análisis sintáctico de las diferentes páginas web y la descarga de la totalidad de los enlaces internos que encuentra.
2. *Wordsmith tools*. 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.
3. *Webbudget*. Es un paquete de software destinado a la localización de sitios web. El programa puede realizar estadísticas desde un punto de vista traductológico 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.
4. *Avral Tramigo*.-Paquete de software dirigido a la extracción y localización de los elementos textuales en los archivos multimedia Flash.
5. *Jdprint*.- Herramienta que proporciona un listado completo estructurado de la totalidad de las carpetas y archivos de todo tipo dentro de un corpus.

## 8. PRINCIPIOS TEÓRICOS Y METODOLÓGICOS DE LA RECOPIACIÓN DEL CORPUS

### 8.1. El corpus

El centro del presente estudio lo constituye la recopilación y análisis de un corpus monolingüe (Laviosa, 1997: 292) de sitios web desarrollados en España y su posterior análisis.

En primer lugar necesitamos definir este concepto ya que constituye el centro de nuestra investigación. Durante los últimos años se han propuesto varias definiciones. Una de las definiciones más aceptadas de corpus es la propuesta por Sinclair (1991:171), «*a corpus is a collection of naturally-occurring language text, chosen to characterize a state or variety of a language*». Esta definición remarca la necesidad esencial de que los textos sean naturales (Stubbs, 1996:4), aunque fuera precisamente en traductología donde se comenzó a crear corpus de textos traducidos que no se consideran naturales<sup>93</sup> (Baker, 1995). La definición propuesta por Biber (1998:4) nos indica la necesidad de que la colección de textos sea extensa y se recopile según unos criterios previos claramente establecidos, «*a large and principled collection of natural texts*». Con el advenimiento de la informática desde finales de los años 80 se introdujo claramente la necesidad del formato electrónico en los corpus para analizarlos de forma automática o semiautomática (Baker, 1995: 226), incrementando exponencialmente el tamaño de los mismos<sup>94</sup>. 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*». Otras definiciones propuestas ahondan en diferentes características esenciales de los corpóra, como que el mismo debe de ser construido de acuerdo a una serie de criterios específicos (Atkins et al. 1992:1; Pérez, 2002), que los textos incluidos en el corpus deben ser seleccionados según unos criterios lingüísticos explícitos que aseguren que los mismos sean representativos de la variedad de lengua que se consideran (McEnery y Wilson, 1996:24), o que los corpus deben tener un tamaño limitado,

---

<sup>93</sup> El objeto de estas investigaciones de corpus se centra en los textos traducidos que poseen características diferenciadas al ser producto de proceso de producción diferente. Según Baker un texto traducido constituye «*a communicative event which is shaped by its own goals, pressures and context of production*» (Baker, 1996:175).

<sup>94</sup> El British National Corpus cuenta con unos 100 millones de palabras de inglés escrito y oral.

formato electrónico y debe de formar una referencia estándar para la variedad de lengua que se pretende representar (McEnery y Wilson, 1996).

### 8.1.1. Diferentes clases de corpus.

Una vez definido un corpus como tal, es necesario revisar los diferentes tipos de corpus para situar el nuestro dentro de los estudios de lingüística de corpus y traductología. En éste último campo, la división más aceptada y aplicada es la propuesta por Baker (1995) que divide los corpus entre *monolingües*, *paralelos* y *comparables*.

Un *corpus monolingüe* es aquel que contiene una sola lengua (Kenny, 2001: 58) y en traductología se distingue entre los corpus monolingües traductológicos o naturales (Laviosa, 2002: 35). El corpus para la presente investigación constituiría un corpus monolingüe<sup>95</sup> no traductológico, que en nuestro caso podríamos denominar *monocale*<sup>96</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, 2004), con lo que definir nuestro corpus como monolingüe no sería suficiente. La importancia del concepto de *locale* en la recopilación del corpus es esencial y se estudiara a continuación (véase 8.2.1).

En traductología se han usado corpus no traductológicos de textos especializados 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) y para la extracción de terminología (Pearson, 1998; Faber et al, 2001). Además, en traductología, debido a la necesidad de comparar los datos obtenidos (Kenny, 2001: 125)<sup>97</sup>, se usan varios corpus monolingües de control como el *British National Corpus* o *Cobuild Bank of English* en inglés o el CREA en español<sup>98</sup>. Nuestro estudio preliminar con un corpus *monocale* surge además con la intención de crear un corpus de «control» claramente delimitado para estudios empíricos futuros sobre el proceso de traducción en localización.

En el estudio previo del presente autor, germen de la presente investigación (Jiménez, 2005), se estudiaron las diferencias tipológicas en textos web a partir de un **corpus comparable** según la definición de Baker (1995), «*a structured*

---

<sup>95</sup> Para el establecimiento de una base de un investigación más compleja sobre el proceso de traducción en localización.

<sup>96</sup> Definimos nuestra locale según la W3C en es-Es, es decir, español de España.

<sup>97</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>98</sup> Disponible en <http://corpus.rae.es/creanet.html>

*electronic collection of texts originally written in a particular language, alongside texts translated into that same language»*<sup>99</sup>. Son muchos los estudios en traductología que han usado este tipo de corpus, que en general estudian las características de los textos traducidos comparados con los textos originalmente producidos en la lengua idioma meta, destacando los estudios de Baker (1995, 2001) sobre los «universales de la traducción» en textos traducidos al inglés, y continuado por otros investigadores que han estudiado las peculiaridades textuales de las traducciones en otros idiomas (Toury, 1980; Puurtinen, 1995; Laviosa, 1996, 1997, 1998; Kenny, 2001). En estas investigaciones con corpus comparables se ha demostrado que en los textos traducidos suelen aparecer estructuras y frecuencias que no son típicas si los comparamos con textos que se han producido originalmente en la misma lengua (Gellerstain, 1996; Laviosa-Baithwate, 1996). Estos resultados se observaron igualmente en nuestra investigación previa (Jiménez, 2005); 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>100</sup>, y por otro lado en la frecuencia atípica de varias colocaciones como «contacta con nosotros» o «sobre nosotros»<sup>101</sup>. En nuestro caso, los resultados obtenidos se interpretan además por los efectos y presiones del entorno del proceso de traducción en localización que según nuestra hipótesis incide de manera diferenciada a la traducción literaria (Kenny, 2001) o periodística (Baker, 1995).

Por último, 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 traductología, aun cuando varios autores toman la definición proveniente de análisis contrastivo que equivaldría al corpus comparable (véase Laviosa, 2002: 37).

Son varias las características y aspectos adicionales que se consideran en la descripción de corpus usados en estudios de traductología. En los estudios de género realizados por el grupo GENTT (García Izquierdo y Nebot, 2003) se usa la propuesta de Corpas, (2001: 3 y ss.), que partiendo de una propuesta de Torruella y Llisterri (1999), la combina con las propuestas en traductología de

---

<sup>99</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.

<sup>100</sup> Reflejo de la colocación conveccionalizada en los géneros web de páginas corporativas norteamricanas «Privacy Policy», «Terms of use» y «home» respectivamente.

<sup>101</sup> «Contact us» y «About us» respectivamente en inglés.

Baker (1995) mencionada anteriormente y Johansson (1998). Esta propuesta clasifica los corpus según los siguientes criterios:

- 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.
- Según la especificidad de los documentos contenidos en el corpus: corpus general, corpus especializado, corpus canónico, corpus periódico o cronológico y corpus diacrónico.
- 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.
- Según la codificación y anotación: corpus no anotado, corpus anotado
- Según la documentación que acompaña al corpus: Corpus documentado y corpus no documentado.

Si combinamos nuestra propuesta en localización con la clasificación de Corpos, nuestro corpus sería un corpus *monolocale*, extenso, especializado puesto que se centra en un campo técnico como la localización, textual, puesto que está compuesto por textos completos, no anotado, y documentado.

## 8.2. Recopilación del corpus

Las consideraciones teóricas a la hora de construir un corpus deben ser claras y preceder a la creación del mismo (Biber, 1993: 256):

[...] theoretical research should always precede the initial corpus design and actual compilation of text .

En nuestra opinión, éste es uno de los puntos de los que adolecen varias de las investigaciones sobre textos web hasta la fecha, en las que se usa el concepto de corpus sin definir claramente que entienden por el mismo y el análisis claro de los criterios y limitaciones que han llevado a su recopilación.

En los estudios en traductología se establece claramente que el diseño del corpus es de esencial importancia (Kenny, 2001; Zanettin, 2000), y puesto que la

compilación de un corpus conlleva gran esfuerzo y tiempo, es necesario que se cree con unas bases teóricas claras para que se obtenga el mayor beneficio del mismo (Zanettin, 2000), siendo además los resultados que se obtienen tan válidos como lo es el mismo corpus (Sinclair 1991:13).

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.

La recopilación del corpus se llevará a cabo claramente según los criterios expuestos en los estudios de traductología (Pérez Hernández, 2002) y procediendo de manera cíclica, revisando nuestros principios de recopilación en caso de que sea necesario durante la fase de análisis<sup>102</sup> (Biber, 1993: 256). Por ejemplo, tras los errores en la descarga del corpus, se tuvieron que eliminar las páginas de presentación (*Splash pages*) en Macromedia Flash al impedir la descarga del sitio web con la herramienta usada, Htrack (véase 8.5.2). Así, aun cuando estas páginas forman parte del sitio web completo, fue preciso eliminarlas para garantizar una descarga sincrónica en un solo día de los sitios web seleccionados.

### **8.2.1. Criterios para la recopilación del corpus**

Nos embarcamos en la recopilación de este corpus, que se ha guiado por diversos criterios. Los criterios de representatividad, estandarización y tipología de los corpórea han sido tres conceptos que más se han debatido en el seno de la comunidad científica (Pérez Hernández, 2002). El criterio más importante en nuestro caso es la representatividad, concepto esencial en la recopilación. 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.

La representatividad se establece incluyendo toda la posible variación en la población «*the full range of variability in a population*» (Biber 1993: 243). En la presente investigación, basada en los estudios de localización en traductología y con el fin de unir esfuerzos con la industria de la localización, la

---

<sup>102</sup> «Regardless of the initial design the compilation of a representative corpus should proceed in a cyclical fashion.» (Biber, 1993: 256).

población a la que hacemos referencia sería la representada por la *locale* es-ES<sup>103</sup>, *Spanish -Spain*, que define la W3C, el consorcio que regula Internet. Este código viene de la combinación de la ISO 639 que representa los códigos de lenguas<sup>104</sup> con la ISO 3166, que representa los códigos de países<sup>105</sup>. En nuestra investigación, no se podría establecer el uso del dominio *.es*<sup>106</sup> como representativo de la población total, puesto que muchas empresas españolas usan el dominio *.com*, o *.net* como primario.

El uso de *locale* y no lengua en las investigaciones en localización se debe, como comenta Pym (2004: 22), a que el concepto de lengua es muy amplio, «*due to many countries, communities, dialects, etc.*». *Locale* se entiende como un «prototipo» (Rosch, 1978, 1977) sin unos límites claros que establezcan la separación, con unos casos más centrales que otros. En la acotación necesaria de una *locale* (Pym, 2004:22), se indica 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.

Según el autor, si un texto se puede distribuir sin localización, no existirá una *locale* nueva<sup>107</sup>. La dependencia de un mercado en la constitución de una *locale* sería la diferencia esencial entre ésta y un dialecto ya existente. A nuestro parecer, la resistencia en textos web no sería suficiente puesto que varios estudios han concluido que, en el contexto económico, los usuarios prefieren realizar transacciones en sitios web en su lengua, (Singh y Pereira, 2005: 5) y a pesar de que los usuarios internacionales puedan usar sitios web en inglés, los beneficios desde el punto de vista económico de la localización son muy importantes (Yunker, 2003: 12-13).

Para asegurar la representatividad de sitios web corporativos en la *locale* es-ES, nuestro corpus se ha recopilado usando el directorio de [www.google.com](http://www.google.com), siguiendo la investigación sobre registros web de Biber (2004). En este directorio, se usó la categoría [World>Español>Regional>Europa>España>Economía y Negocios](#)<sup>108</sup>. A continuación se eligió el primer sitio web de cada subdirectorio siempre que el mismo fuera una creación natural en

---

<sup>103</sup> Los códigos de países se pueden obtener en <http://www.w3c.org/International/O-HTML-tags.html>

<sup>104</sup> ISO 639 *Codes for the representation of names of languages*

<sup>105</sup> ISO 3166 *Codes for the representation of names of countries*

<sup>106</sup> Los códigos de país de dominios se asigna por la [Internet Assigned Numbers Authority \(IANA\)](http://www.iana.org)

<sup>107</sup> Como podría ser el caso de las lenguas indígenas en Norteamérica

<sup>108</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.

español y no el producto de la localización de un sitio web (véase 8.4). No se incluyó el subdirectorio «Compras», ya los sitios web corporativos centrados en el comercio electrónico representan otro género diferenciado. De igual manera no se incluyeron aquellos subdirectorios que no incluyeran el género web de estudio, tales como «Asociaciones», «Noticias y publicaciones», «Organizaciones» o «Autoridades».

En este caso, y puesto que se pretende realizar un estudio del género web como forma textual convencionalizada, se eligió el primer sitio en cada subdirectorio con el objetivo claro de recopilar los sitios web más usados en España en todas las áreas económicas. El directorio Google se basa en el directorio *Open Source Directory*<sup>109</sup> de Netscape, y a su vez Google ordena los sitios web siguiendo su tecnología *Pagerank*<sup>110</sup>, según el uso y relevancia de cada sitio, con lo que la representatividad se garantiza seleccionando los sitios web más usados de la totalidad de las áreas económicas. Los sitios web menos usados suelen tener unas características no convencionalizadas y con lenguaje más ideolectal por lo que no se incluyen en el estudio (Göpferich, 1995)<sup>111</sup>.

Las investigaciones en lingüística de corpus indican la necesidad de incluir **textos completos**, puesto que hay características lingüísticas que raramente se distribuyen de manera igualitaria a lo largo del texto y puede tener baja o alta representación (Kenny 2000; Baker 1995:240, Biber 1994: 389; Sinclair, 1991: 19). En el presente estudio se concluye que el texto web, o hipertexto, lo constituye el sitio web completo (ver 5.3). Varias investigaciones incluyen exclusivamente páginas web aisladas (Bolaños, 2002; Fernández, 2003; Ruiz, 2003; ; Biber, 2004; Rennau, 2004; Santini, 2005a), o por motivos metodológicos se excluyen páginas de menos de 200 palabras (Biber, 2004), cuando uno de las principales características de las páginas web bien escritas es la brevedad y concisión<sup>112</sup> (Krug, 2005, McGovern et al., 2002). La media de palabras por página web en nuestro estudio es de 205,99 palabras en cuerpo de página web o 356,50 palabras incluyendo los elementos textuales en el código. Desde el punto de vista del proceso de traducción, y siguiendo nuestra definición de texto en localización, las páginas web no constituirían un texto completo sino un subtexto. Una página web no se sirve de modo unitario puesto que como se ha indicado depende de

---

<sup>109</sup> <http://dmoz.org/>

<sup>110</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>111</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.

<sup>112</sup> Nielsen (2000) y Kilian (1999) comentan reducir el texto al 50%, y de ahí, reducirlo el 25% más.

las relaciones contextuales que les proporciona el sitio web para completar su textualidad<sup>113</sup>.

La estandarización (Pérez Hernández, 2002) de los sitios web es otro criterio de importancia, y por lo tanto no se incluirán sitios web que estén en construcción o sitios que no cumplan con las convenciones básicas de las páginas web como la inclusión de un menú de navegación<sup>114</sup>. Como hemos mencionado, la exclusión de textos no convencionalizados es común en los estudios de diferentes tipologías textuales (Göpferich, 1995, Werlich, 1975).

Estos sitios web representan una misma situación comunicativa, que se basa en la relación comercial entre una empresa multinacional y los clientes de la misma. El esquema general que representa esta situación comunicativa (Janoschka, 2003) se expuso con anterioridad (véase 4.2.2.3).

La tipología textual y de género que al que se ajustan los sitios web seleccionados está definida por la función comunicativa y pragmática de estos sitios web. Éstas están muy influenciadas por el márketing y la imagen, puesto que la misma representa la imagen corporativa. Los sitios web incluidos en el corpus deben pertenecer al género web «*corporate homepage*» establecido anteriormente.

### 8.3. Descarga del corpus

La selección del corpus se realizó entre los días 6 y 7 de marzo del 2006. La descarga final del corpus de Internet se realizó en un solo día, el 4 de mayo del 2006<sup>115</sup>, por medio del programa Httrack<sup>116</sup> para asegurar el carácter sincrónico del mismo. En la descarga 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 en general suelen ser sincrónicos para una caracterización en un momento temporal determinado (Kenny, 2001: 106). La constante actualización a la que se suelen someter los sitios web (O'Hagan y

---

<sup>113</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>114</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

<sup>115</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.

<sup>116</sup> El programa se puede obtener a través de [www.httrack.com](http://www.httrack.com)

Ashworth, 2003) supuso la descarga de todos los sitios web en un día, con lo que se consiguió almacenar una representación sincrónica de la población web objeto de estudio. De hecho, la naturaleza inestable del texto web es una de las características que diferencian los textos web de los textos impresos (véase 5.6) (Shepherd and Watters, 1998).

Al recopilar sitios web completos, obtendremos páginas web en diferentes extensiones como: .html, .htm, .xml, .dhtml, xhtml, .jsp (*Java server pages*), .aspx (*active server pages*) etc. Además, se descargarán y formarán parte del corpus todas las imágenes, tablas, tipografía, etc. puesto que al igual que la tipografía, son parte integrante del texto (Göpferich, 1995).

Debido a las limitaciones de almacenamiento y descarga, así como las características de ciertos archivos, se excluyeron los archivos en formato .pdf que representan además un reflejo de los géneros *extant* (véase 4.2.2.6.1), y cuyo objetivo se indica sólo para la impresión de documentación que se distribuya por Internet (Nielsen, 2003). Se excluyeron además los archivos multimedia como los vídeos, los archivos de sonido y los archivos comprimidos<sup>117</sup>.

#### **8.4. Criterios para asegurar que se trata de un texto original**

En la marabunta de la web, es esencial definir claramente qué sitios web son producciones naturales en la *locale* es-ES, incluyendo los sitios web que se produjeron en catalán o cualquier otra lengua peninsular. Los siguientes criterios son de esencial importancia y se desarrollaron para determinar la autoría de sitio web como creación original en la *locale* es-ES, pudiendo usarse con otros fines en los estudios de traductología.

En la actualidad se usa extensivamente la web como corpus de documentación en cualquier campo, y el traductor necesita conocer los métodos de evaluación claros de la información que aparezca en la web como en cualquier búsqueda documental durante el proceso de traducción. Por ejemplo, los criterios generales que indica Palomares (1999: 179) son: la autoría, el contenido, el acceso y el diseño. En nuestro caso, la autoría incluye además la necesidad de verificar el texto web como creación natural en la *locale* es-ES.

---

<sup>117</sup> Archivos de vídeo y sonido excluidos: .wma, .wmv, .mpeg, .rm, .avi, .asf, .mov, .wav, .vob. Archivos de compresión: .rar, .tar, .zip, .tgz

Los criterios que se listan a continuación se elaboraron por el propio autor en la investigación previa (Jiménez, 2005), respondiendo a la necesidad de que los textos sean originales o naturales<sup>118</sup> (Halverson, 1998) en la recopilación de corpus monolingües. Debemos destacar que su concepción se basa en la teoría de prototipos de Eleanor Rosch (1977, 1978), por lo que los criterios de inclusión o no inclusión en el corpus no son rígidos. Según la teoría de prototipos, existen miembros más cercanos al centro de la categoría y otros más periféricos y no existen fronteras claras definidas de inclusión o no en el corpus.

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>119</sup>.
- La no inclusión de caracteres exclusivos del español tales como ¡, ¿, ñ.
- La presencia de comentarios<sup>120</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 saber si las páginas web que se suponen traducidas han sido originalmente desarrolladas en España. Por ejemplo, la página española de la multinacional Adecco<sup>121</sup>, incluye en sus *javascript* variables en español, con lo que asumimos que la página se desarrolló 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;')
```

---

<sup>118</sup> Halverson (1998) indica que deben ser «*naturally-occurring language*».

<sup>119</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>120</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 -->

<sup>121</sup> <http://www.adecco.es/>

- La presencia de subtextos 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, e inglés la página original tiene el dominio .com y la traducción al español .es
- Y como último, tomamos como creación original los sitios de multinacionales que pueden tener un sistema de GCM descentralizado (véase 3.4 ) y por lo tanto 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 siendo muy parecida a la página principal en inglés [www.adecco.com](http://www.adecco.com), contiene una tipología textual que la nos indica que es una creación original en español. No obstante, debido a los diferentes niveles de localización, es posible que un sitio globalizado en el nivel de adaptación (Brooks, 2000) pueda parecer una creación original en español. En ese caso es necesario observar el código fuente y encontrar los comentarios de los desarrolladores <!-->, o observar la estructura de los URL para observar si el sitio web es el producto de la localización.

## **8.5. Limitaciones**

La recopilación del corpus de manera sincrónica y las particularidades del texto web conllevan unas limitaciones que se pueden agrupar en dos grupos: las limitaciones debidas a las funcionalidades del software empleado combinadas con las particularidades del texto web, y aquellas que se surgen del volumen de textos recopilados.

El volumen de textos recopilados supuso la imposibilidad de descargar ciertos componentes del texto, como los archivos .pdf, los de video, sonido, o los archivos comprimidos. Sinclair (1991) apunta la necesidad de contar con un corpus de gran tamaño para las investigaciones siguiendo esta metodología, y en este caso, la descarga sincrónica de nuestro corpus con los medios disponibles supuso la necesaria eliminación de estos elementos.

En segundo lugar, el software empleado para la descarga del corpus en un solo día fue Htrack, que se basa en el análisis sintáctico del código de las páginas

web, incluyendo los *scripts*, para detectar los enlaces internos<sup>122</sup> y descargarlos a su vez. Los problemas se presentaron por: los *scripts* que presentan páginas web dinámicas, y por lo tanto distintas dependiendo de la sesión<sup>123</sup>, la existencia de Macromedia Flash en un gran número de sitios web, y al uso de marcos o *frames* en las páginas. En estos tres casos, el análisis sintáctico de las páginas web no detecta los enlaces y por lo tanto, el programa no puede descargar el sitio web completo. De esta manera, se tuvieron que eliminar 74 sitios web del total de 248 sitios inicialmente seleccionados, el 29,8% de los mismos. De igual modo, es imposible acceder y descargar las intranet o zonas reservadas.

### 8.5.1. Los marcos y la arquitectura servidor-usuario

En la actualidad, gran cantidad de sitios web se generan dinámicamente a través de un servidor activo, produciendo una página web virtual (véase 4.2.2.6.1), es decir distintas dependiendo del momento o del usuario. Estas páginas web se basan en programas de varios lenguajes de programación que necesitan de la conexión a Internet para funcionar correctamente. En estos casos, 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>124</sup>, por lo que la extensión de la página web que se obtiene es *.do*, indicando la existencia de un programa Java que funciona en un servidor activo. En estos casos, se hace imposible copiar un sitio para su estudio. Se podría no obstante estudiar las páginas web *virtuales* una por una, pero como mencionamos anteriormente, no consideramos que constituyan un texto completo. Las grandes empresas con grandes recursos están usando cada vez más estos lenguajes de programación con servidores Web, con lo que se dificultará más si cabe en el futuro el estudio de la textualidad web de sitios corporativos completos. En nuestro corpus, identificamos 20 sitios web que se excluyeron al no ser posible descargarlos debido a estas circunstancias, el 8,6% del total inicialmente seleccionado.

Los marcos o *frames* se basan en la posibilidad de incluir varias páginas web dentro de una 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 (2000: 85-87) indica la necesidad de prescindir de los marcos debido a los problemas que

---

<sup>122</sup> Los enlaces en hipertextos pueden ser internos, hacia el mismo sitio web, o externos, hacia otros sitios web (ver Janoschka, 2003: 178).

<sup>123</sup> En este caso suelen ser *Active Server Webpages*.

<sup>124</sup> Normalmente IBM WebSphere

producen en la categorización, navegación o búsqueda <sup>125</sup>. El programa Htrack presentó problemas en algunos sitios web puesto que el URL de la página web no es el mismo de las páginas que constituyen los marcos y por lo tanto, el programa considera estas páginas como pertenecientes a otros sitios web. Los marcos presentaron problemas en 11 de los sitios excluidos, algunos de ellos combinados con Flash, el 4.43%.

### **8.5.2. Macromedia Flash**

El uso de Macromedia Flash está aumentando de manera exponencial en las páginas web por proporcionar una experiencia multimedia, animada y muy interactiva. El uso de Flash en su formato FLA (*Flash Movie Authoring File*), convertido en formato SWF (*Small Web Format/ShockWave Flash*) para distribuirse por Internet, puede presentar varios problemas: el usuario puede no tener este componente instalado y su inclusión no debe molestar, distraer o confundir al usuario durante la sesión (Brinck et al. 2002). Una conexión lenta puede llevar al usuario a abandonar el sitio web<sup>126</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 las presentaciones y animaciones le distraen de este propósito básico. Por cuanto al proceso de traducción, Macromedia Flash se aleja de la simplicidad del código 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 a continuación (Macromedia, 2004). A menudo es necesario recurrir a paquetes de software especializados en la traducción de estos tipos de archivos, como Avral Tramigo<sup>127</sup>.

El uso de Flash en los sitios españoles 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. En la recopilación, los dos últimos casos presentaron problemas puesto que el analizador del programa Htrack no puede detectar los enlaces internos en las animaciones Flash, impidiendo descargar la totalidad de sitios web como [www.lechepascual.es](http://www.lechepascual.es), que usa Flash en su menú de navegación. Para intentar solucionar el problema se quitaron las presentaciones

---

<sup>125</sup> Véase Brinck et al. (2002:361-364) para un tratamiento de las ventajas y desventajas del uso de marcos en páginas web.

<sup>126</sup> Los manuales de desarrollo recomiendan incluir la opción de «saltar la introducción» e incluir una barra que indique el estado de la descarga de la presentación.

<sup>127</sup> Para información sobre el paquete de software : <http://www.avral.com/tramigo/>

Flash de inicio<sup>128</sup>, incluyendo las que sirven como selectores del idioma de preferencia. En 18 de los casos, el 10,46%, el sitio web se pudo descargar con normalidad.

El uso de Flash presenta además problemas por cuanto a la accesibilidad para personas con discapacidades, al no incluir etiquetas básicas como `<alt>` o no poder el usuario cambiar el tamaño de letra desde su navegador en caso de que sea necesario.

En caso de usar Flash, se recomienda presentar el sitio web en dos versiones html y Flash: el único sitio web en nuestro corpus que presenta ambas versiones con la posibilidad de elegir es [www.culture02.com](http://www.culture02.com), un 0,58% de los sitios web españoles.

### 8.5.3. Sitios web eliminados por distorsionar las estadísticas

En la recopilación de corpus web es difícil conocer a priori la extensión del hipertexto a estudiar. La concepción cíclica en el diseño y compilación de corpora (Biber 1993: 256) nos llevó a separar 2 sitios web durante el proceso de análisis del mismos. Estos 2 sitios web poseen unas características que distorsionarían las estadísticas, [www.fotolab.es](http://www.fotolab.es) y [www.masson.es](http://www.masson.es). El primero cuenta con 22.103 páginas web totales, y el segundo 4.082. A pesar de ser sitios web corporativos, el objetivo final de estos sitios web se centra más en la venta *online*, lo que implica la inclusión total del catálogo de venta. Se indicó anteriormente los sitios web de compra *online* constituyen un género diferenciado, aunque en ciertos casos la línea de separación entre géneros web en constante evolución no está clara y por lo tanto no se observó en un primer momento. Como ejemplo, la media de páginas por sitio web en nuestro corpus es de 161,99. En caso de que se incluyeran estos dos sitios web la media sería de 310,5.

---

<sup>128</sup> Aún cuando según la concepción del texto web defendida, forman parte del mismo.

## **9. RESULTADOS DEL ANÁLISIS**

### **9.1. Composición final del corpus**

El corpus final se descargó el 4 de mayo del 2006. En un principio se seleccionaron 248 sitios web con una representación de todas las áreas económicas. En la descarga definitiva no fue posible descargar 74 de los sitios web, el 29,8%, debido a las limitaciones indicadas anteriormente. Se separaron además dos sitios web por presentar características tipológicas y discursivas que distorsionarían las estadísticas.

El corpus final se compone de un total de 27.852 páginas web, con una media de páginas por sitio web de 161,93. A su vez, se compone de 67.868 ficheros, 394,58 por sitio web y 10.298 carpetas de archivo. El total de bytes es 3.001.247.410 bytes o 3,01 Gigabytes. La media sería 17.4449.112 bytes por sitio web, alrededor de 17,5 Megas.

La media de páginas web por sitio español en el corpus es de 161,93 páginas, con una media de 356,50 palabras por página. No obstante, es imposible contabilizar con programas estadísticos las palabras codificadas en imágenes o archivos Flash. En la siguiente tabla se puede observar la composición global del corpus con especial interés en los elementos textuales de interés en el proceso de traducción.

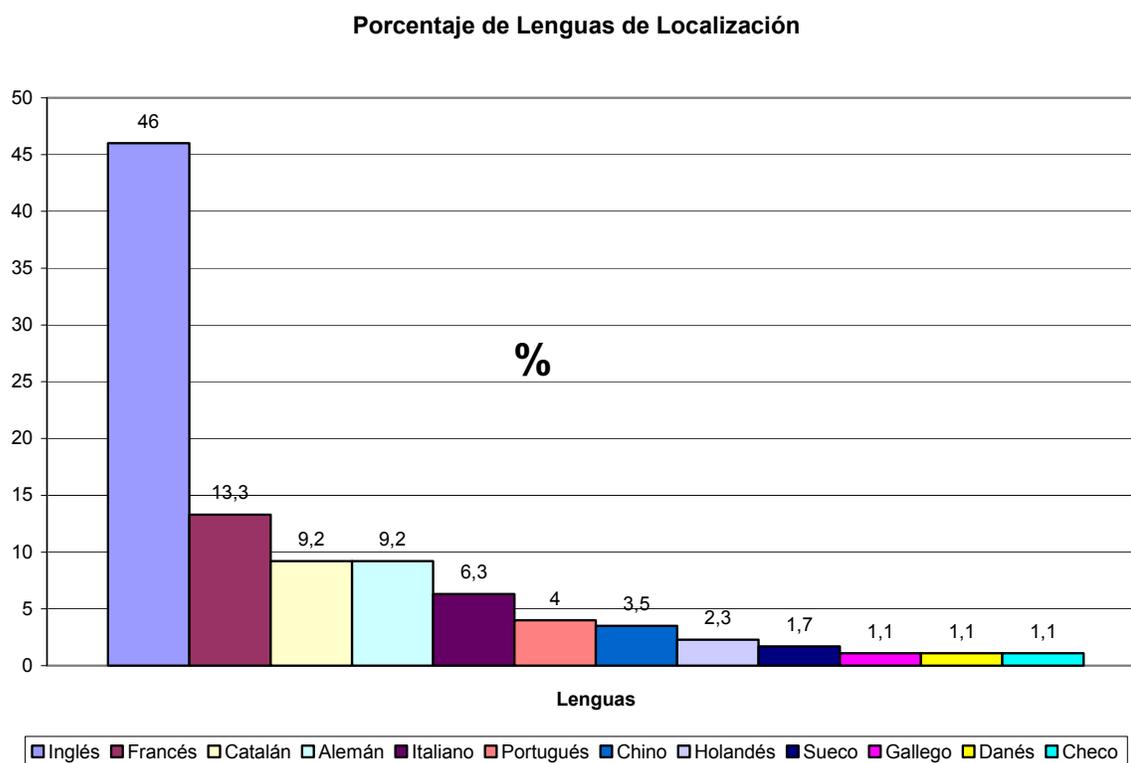
<b>Corpus total: 172 sitios web corporativos</b>		
<b>Tipo</b>	<b>Total</b>	<b>Media</b>
<b>Páginas web</b>	<b>27.852</b>	161,93 páginas por sitio web
<b><u>Palabras en Cuerpo de la página</u></b>	5.737.289	205,99 palabras por página web.
<b>Palabras total</b>	9.929.302	356,50 palabras por página
<b>Palabras en &lt;Meta&gt;</b>	934.361	33,54 palabras
<b>Palabras en &lt;alt&gt;</b>	317.605	11,40 palabras
<b>Palabras en &lt;input&gt; y &lt;select&gt;</b>	1.359.017	48,79 palabras
<b>Palabras en Scripts</b>	242.720	9,71 palabras por página
<b>Palabras en &lt;OnMouse&gt;</b>	1.338.310	48,05 por página
<b>Caracteres sin incluir los espacios</b>	73.786.696	2649,24 caracteres
<b>Número de enlaces</b>	936.975	33,64 enlaces por página
<b>Número de Imágenes</b>	852.938	30,62 imágenes por página

Tabla 3. Estadísticas del corpus realizada con el programa Webbudget.

### 9.1.1. Sitios web localizados del español a otras lenguas en el corpus

En la descarga del corpus fue imposible separar o incluir la totalidad de los sitios web españoles localizados a otras lenguas debido a los diferentes modelos de localización. Esto se debe a la falta en gran medida de un modelo claro de globalización para los sitios web españoles. En algunos casos los sitios web localizados se encontraron inmersos en el mismo directorio que el sitio español, en otros casos se encontraban en directorios separados y en varios casos se presentaban en dominios diferentes. Debido a esta particularidad se decidió que no era posible separar en una descarga sincrónica la totalidad las páginas web localizadas a otros idiomas y se permitió que el programa Htrack descargara los

sitios web con sus localizaciones <sup>129</sup>. La lengua con mayor número de localizaciones fue el inglés, con un 46% de los sitios web, con lo que se observa el estatus de *lingua franca* en la web (LISA 2003), seguida del francés con un 13,3% y del catalán y el alemán ambos con un 9,2 %. En la siguiente gráfica se puede observar la distribución de versiones localizadas en los sitios web españoles.



**Tabla 4. Distribución de versiones localizadas de sitios web españoles.**

## 9.2. Caracterización del sitio web corporativo desde una perspectiva traductológica

Son varios los estudios descriptivos que se han centrado en los géneros digitales web. Estos estudios se centran en ciertos géneros de sitios web como los de empresas informáticas (Bolaños, 2002), las empresas azulejeras (Rennau, 2004), las empresas bancarias (Fernández, 2003), de las universidades (Ruiz, 2003), de

<sup>129</sup> Sólo cuando la versión localizada se encuentre incluida en el mismo directorio del sitio web español.

publicidad web (Alamán, 2003; Janoschka, 2003), las páginas institucionales (Rehm, 2002) o las páginas de inicio (Askehave y Nielsen, 2005). El objetivo del presente estudio se basa en la descripción desde una perspectiva del proceso de traducción. En este caso, comenzaremos el análisis de la descripción de los textos web con la distinción entre: los distintos elementos textuales en los que se componen las páginas y en los distintos tipos de páginas de los que se compone un sitio web. A continuación, se analizará la estructura prototípica de este género web en español, para pasar a caracterizar los bloques comunicativos desde cuatro de los factores de caracterización del género: los elementos de la situación comunicativa, la función textual, la funcionalidad y los rasgos convencionales, incluyendo elementos extra e intratextuales.

### **9.2.1. Diferentes clases de páginas**

Los textos web que suponen los sitios web corporativos se componen de diferentes tipos de páginas web con ciertas características diferenciadas: las páginas web de *Splash*<sup>130</sup>, páginas principales, páginas interiores y páginas producidas por la arquitectura servidor-usuario.

#### **9.2.1.1. Las páginas web de Splash**

Estas páginas se presentan en el momento en el que el usuario accede al sitio web. Suelen presentar animaciones y un concepto gráfico que pretenden atraer al usuario y mostrarle la alta calidad del sitio web y evidentemente, la empresa. En general se utiliza el Flash para realizar estas animaciones. Las guías de usabilidad no recomiendan su uso ya que ralentizan el acceso del usuario al contenido que busca y en conexiones lentas pueden incluso hacer que el usuario abandone el sitio web (Brinck et al. 2002). Se recomienda que se incluya la opción de saltar o pasar la introducción, que en español se suele indicar con «Saltar introducción» o «Saltar intro».

---

<sup>130</sup> El término «página de presentación» se puede usar como sinónimo de página de inicio o *homepage*



**Figura 17. Pantalla animada Splash en formato Macromedia Flash. Usada además como selector de idioma.**

En nuestro corpus, debido a los problemas que supuso el uso de Macromedia Flash en la descarga del mismo, se eliminaron 18 páginas de Splash para que se pudiera descargar el sitio global completo<sup>131</sup>.

Uno de los usos más extendidos de éste tipo de páginas en España es el uso como selector de idiomas del mismo, como en el ejemplo incluido de [www.rianxeira.es](http://www.rianxeira.es).

### 9.2.1.2. Página principal

Ésta es la página de inicio y en ella se presenta la información más importante para el usuario y un resumen de los contenidos que puede encontrar en la misma. La página principal cumple dos funciones, una expositiva o informativa en la que se comunica los contenidos del sitio web y la información básica de la empresa y otra exhortativa o instructiva con ciertos elementos textuales y visuales que nos animan a quedarnos en el sitio web y visitar ciertas partes (Askehave y Nielsen, 2005: 123; Bolaños, 2002). Esta página sirve de «puerta de entrada» al sitio web aunque los usuarios no tienen por que empezar su visita en la misma, puesto que se puede acceder directamente a cualquier página sin pasar por la misma, un fenómeno que se conoce como *deep linking* (Janoschka, 2003; Nielsen, 2000: 179). Son muchas las metáforas con las que se identifica esta página, pero no todas se corresponden con las funcionalidades de este nuevo

---

<sup>131</sup> Aun cuando en nuestra definición del texto web forman parte del *hipertexto* (véase 5.8) que se recopila.

género, como la portada de una revista<sup>132</sup>, la puerta de entrada, la tabla de contenidos, el recepcionista de la empresa, folleto informativo etc. (Nielsen y Tahir, 2002: 19). Como en cualquier otro tipo de páginas, se hace esencial la brevedad, concisión y claridad para que el usuario comprenda la misma.

El análisis del corpus nos muestra la convencionalización del término «Inicio» en la denominación de este tipo de páginas, con un 60% de aparición de este término en las páginas de inicio. Un 24% de sitios usan el término «Home» en este caso. Se observa por tanto una influencia clara del género anglosajón en los sitios españoles. En varios casos se detectó cómo el uso de paquetes de software de desarrollo web anglosajones influye en el diseño de los sitios web, con la inclusión de iconos producto de metáforas anglosajonas, como una casa para «Inicio».



Figura 18. Uso de iconos anglosajones en página web de inicio española. [www.caditesting.es](http://www.caditesting.es)

Otros términos menos convencionalizados que se detectaron en el corpus para la denominación de este tipo de página son: página principal (6%), principal (4%), portada (4%) o página de salida (2%).

### 9.2.1.3. Páginas interiores

Este tipo de página se diferencia de la página de inicio en que incluye contenidos específicos de mayor profundidad para el usuario o que facilitan la funcionalidad con páginas de contacto, formularios, opciones de búsqueda etc. La páginas interiores deben incluir un menú de navegación y la posibilidad clara de llegar a la página de inicio de la empresa, ya que debido al *deep linking*, o la posibilidad de acceder al sitio web sin pasar por la página de inicio, el usuario necesita conocer claramente donde se encuentra y las relaciones estructurales del sitio (Nielsen 2000:179). La navegación suele ser más condensada de la que aparece en las páginas de inicio, con la inclusión clara de este vínculo. En ciertos tipos de páginas se justifica la desaparición del menú de navegación, como en

---

<sup>132</sup> La diferencia esencial que indican Nielsen y Tahir (2002) en este caso es que el usuario accede al sitio web con deliberadamente y en cambio, una portada de revista necesita competir con el resto de publicaciones.

los que se necesita que el usuario complete una tarea, por ejemplo en los formularios, encuestas, formalización de compra *online* etc. (Brinck et al, 2002). La inclusión del menú de navegación puede distraer al usuario de esta tarea básica.

Este tipo de páginas incluyen contenidos más amplios y detallados, ya que el usuario que llega a ellas ha elegido voluntariamente activarlas. Esto se refleja en un menor grado de escaneado de la página y una lectura más atenta cuanto más se profundice en la estructura del hipertexto (Janoschka, 2003: 109).

#### **9.2.1.4. Las páginas basadas en la interacción por la arquitectura servidor-usuario**

Las realizaciones de las páginas internas pueden estar basadas en la interacción, como los formularios, o las páginas que se producen a través de programas que se encuentran en el servidor web como las páginas con extensión .asp, .do, .cgi, o .aspx. 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. Estas páginas se construyen según las especificaciones del usuario y del servidor o sólo las preferencias del usuario y pueden producir páginas web diferenciadas dependiendo de la sesión o del usuario. Representan una realización del género web virtual (Shepherd y Watters, 1998).

#### **9.2.2. Elementos textuales a traducir en las páginas web.**

La textualidad de las páginas web se caracteriza por la repetición de ciertos elementos textuales a traducir. Estos elementos textuales son: los títulos, los elementos de navegación, el cuerpo de la página web, los pies de páginas, las etiquetas meta, los segmentos textuales en javascript u otro tipo de script, los segmentos textuales en el código Html como los incluidos en *OnMouse*, o *Input*, las columnas laterales (*sidebars*), los anuncios, las imágenes con texto, las animaciones flash y los demás elementos multimedia como vídeos, clips de sonido, anuncios etc. (Brinck et al, 2002). Cada elemento textual posee unas características y lenguaje específico.

### 9.2.2.1. Los títulos

Los títulos de las páginas aparecen en la barra superior del navegador y sirven para dar a conocer al usuario donde se encuentra. La etiqueta html usada en este caso es `<title> </title>`. Este componente es esencial para que los buscadores encuentren y clasifiquen el sitio web, y sirve como guía para el usuario, aunque éste raramente lo lee (Brinck et al. 2002: 273). Las guías de estilo web indican incluir el nombre de la empresa en el título y no comenzar el segmento con determinantes u otras palabras gramaticales (Nielsen, 2002).



Figura 19. 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.

En los sitios web del corpus se observan cuatro posibilidades para el título: incluir el URL del sitio web, incluir el nombre de la empresa, incluir una descripción de la empresa, o incluir el nombre de la empresa junto a una descripción de la misma. Es esencial además cambiar el título en cada página web para que el usuario sepa además en qué parte del sitio web se encuentra. La totalidad de los sitios web de nuestro corpus incluyeron este elemento textual.

### 9.2.2.2. Los elementos de navegación

Los menús de navegación son esenciales en todo sitio web. Su inclusión es un elemento esencial para distinguir entre el género digital *novel*, de aparición exclusiva en Internet, y un género *replicated*, como los folletos de empresas que simplemente se cuelgan en Internet sin funcionalidad alguna.

Estos menús pueden aparecer en la parte superior de la página web, en la columna de la derecha del mismo, o en ambos. Los pies de página presentan un tipo diferenciado de navegación puesto que están indicados para la navegación básica (Brinck et al, 2002). La brevedad, concisión y claridad son esenciales ya que el usuario no tiene tiempo de interpretar los mismos en caso de ambigüedad.



Figura 20. Menú de navegación superior. [www.mahou-sanmiguel.com/inicio.html](http://www.mahou-sanmiguel.com/inicio.html)

En el corpus se encontraron varios vínculos en los menús de navegación que no son intuitivos, y por lo tanto pueden no ser funcionales, como el vínculo «Genseat» en el sitio web [www.seat.es](http://www.seat.es). Este vínculo indica un bloque comunicativo dirigido a informar sobre las novedades en el diseño innovador y desarrollo de nuevos productos, y se observa la falta de intuitividad y relación semántica entre el término utilizado (trigger) y el contenido de la página. En los casos en los que es difícil de entender el término empleado en un enlace, el usuario dirigirá su atención hacia otro componente del texto (Janoschka, 2004).

### 9.2.2.3. Pies de página

Los pies de página constituyen otro elemento de navegación en la página web y se ha convencionalizado una tipografía con letra más pequeña. En esta área se suelen colocar los elementos que deben aparecer en toda página pero que el usuario raramente lee, como los avisos legales o los *copyrights*.



Figura 21. Pie de página. [www.azud.com/es/index.html](http://www.azud.com/es/index.html)

La inclusión de este elemento textual no se incluye en la totalidad de los sitios web en el corpus.

### 9.2.2.4. Los segmentos textuales en eventos Html

#### 9.2.2.4.1. Los textos en etiquetas *Meta*

Las etiquetas meta son aquellas que usa el desarrollador para insertar información sobre la propia página web que no observa el usuario final. El contenido incluido en estas etiquetas es vario, como el nombre, el conjunto de caracteres utilizado, etc. De interés en la traducción es el contenido de la etiqueta `<meta NAME="DESCRIPTION" CONTENT=...>` o `<meta NAME="KEYWORDS" CONTENT= ...>`, puesto que en ellas se encuentra la descripción de la página web y la lista de palabras clave que la describen. Estas etiquetas son usadas por los motores de búsqueda web en la clasificación de las páginas, de ahí la importancia de su traducción. En las etiquetas meta *CONTENT* se pueden observar las relaciones ontológicas con hiperónimos, hipónimos o sinónimos para facilitar que se encuentren y clasifiquen correctamente los sitios web.

En nuestro corpus encontramos una media de 33,54 palabras en etiquetas meta *CONTENT* por página web, aunque es muy común la repetición del mismo segmento textual en la totalidad de las páginas de un mismo sitio web, con lo que el total de segmentos *Meta* diferentes en el corpus es menor.

#### **9.2.2.4.2. Los segmentos textuales *Alt***

Los segmentos textuales `<Alt>` son aquellos que describen las imágenes en una página web. Son de utilidad para las personas con discapacidades o para aquellas personas con conexiones lentas que por defecto no descarguen las imágenes en las páginas o en caso de error de descarga. En el trabajo de Nielsen y Tahir (2002) encontraron que solamente el 42% de los sitios web incluían este elemento. En nuestro corpus la media de palabras en eventos *Alt* por página web es de 11,40. En los sitios web españoles es común no incluir este elemento en la totalidad de las páginas web, y en un número reducido de sitios no se incluye totalmente, como en [www.diode.es](http://www.diode.es).

El encargo de la localización de un sitio web puede incluir la creación de estos elementos textuales.

#### **9.2.2.4.3. Los segmentos textuales en *OnMouse***

Los segmentos en *OnMouse* se basan en la 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 opcional y la media en nuestro corpus es de 48,05 palabras por página web.



Figura 22. Segmento textual *OnMouse* al pasar el cursor por el área del gráfico.

#### 9.2.2.4.4. Segmentos textuales en *scripts* y `<Select>`

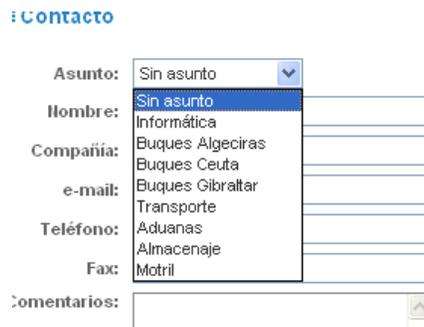
Los segmentos textuales que aparecen en scripts o eventos Html se basan en la funcionalidad que presenta el nuevo medio. Estos segmentos textuales aparecen en situaciones determinadas basadas en la interactividad del usuario con el texto web, sobre todo como respuesta o clarificación a una acción que 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 23. Página protegida que indica un área reservada de [www.diode.es](http://www.diode.es).

La etiqueta `<Select>` constituye un evento html que se usa para la inclusión de menús de navegación o de opciones que se despliegan una vez que se activa el área en la que se encuentran. En general, se desaconseja su uso como método de navegación puesto que el usuario no puede ver la totalidad de las opciones sin

activar el mismo, y por lo tanto, puede no llegar a encontrar la información que necesite (Nielsen, 2000). La brevedad en este tipo de segmento textual es esencial.



The image shows a web form titled "Contacto" with several input fields. The "Asunto:" field is a dropdown menu currently displaying "Sin asunto". The dropdown list is open, showing the following options: "Sin asunto", "Informática", "Buques Algeciras", "Buques Ceuta", "Buques Gibraltar", "Transporte", "Aduanas", "Almacenaje", and "Motril". Other fields include "Nombre:", "Compañía:", "e-mail:", "Teléfono:", "Fax:", and "Comentarios:". Each of these fields has a corresponding empty text input box to its right.

Figura 24. Segmentos textuales en evento «Select».

### 9.2.2.5. Texto en el cuerpo de la página

El cuerpo de la página web constituye el contenido diferenciado de cada página dentro del texto web global. La media de palabras en el cuerpo de página en nuestro corpus es de 205,99, reflejo claro de las peculiaridades de este nuevo tipo textual, marcado por la concisión, brevedad y la necesidad de dividir los textos para presentarlos en una pantalla (McGovern et al, 2002; Vilamor, 2001). Los manuales de estilo indican la recomendación de reducir un texto hasta un 50% en comparación con los textos que aparecen en otros medios, como el impreso (Nielsen, 2000).

Los estudios empíricos en lectura web concluyen que la mejor organización textual de las páginas web se compone de tres columnas, el menú de navegación a la izquierda, el cuerpo de la página en el centro, y otra columna con contenidos resumidos del sitio web a la derecha.



Figura 25. Distribución textual en una página web. Elaboración propia.

[www.microrent.es/index.html](http://www.microrent.es/index.html)

### 9.2.2.6. Elementos multimedia

Los elementos multimedia como sonido, imágenes con texto, animaciones flash, vídeos, clips de sonido etc. son elementos esenciales en la textualidad web. Todos estos elementos integran el texto global y forman parte del mismo (Göpferich, 1995). En el corpus se encontraron una media de 30,72 imágenes por página web. Este aparente alto número de imágenes se explica porque a menudo las imágenes se usan para incluir elementos de navegación, menús o anuncios web. La siguiente figura incluye un menú de navegación compuesto por 22 imágenes en formato .gif; cada opción cuenta con otra imagen que se superpone al pasar el cursor con la etiqueta *Onmouseover*.



**Figura 26. Menú de navegación con texto codificado en imágenes. De [www.pandasoftware.es/home.html](http://www.pandasoftware.es/home.html)**

Al igual que los segmentos textuales incluidos en Flash, no se han podido incluir los elementos textuales en el cómputo total de las estadísticas del corpus por ser imposible realizarlo con las herramientas utilizadas.

Por limitaciones de almacenamiento y descarga, no se recopilaron los videos y clips de sonido en el corpus.

### **9.3. Análisis de la superestructura: bloques y secciones en los sitios web corporativos.**

Cada género está compuesto de una serie de fragmentos textuales, los bloques comunicativos, que en los géneros tradicionales se encuentran organizados jerárquicamente de manera lineal. Estos bloques comunicativos son típicos de cada género (Swales, 1990) y en los textos web se pueden identificar con los diferentes vínculos que producen una superestructura multilineal (Askehave y Nielsen, 2005). Estos bloques constituyen las diferentes partes del texto global, el hipertexto, que tienen una función específica dentro de la función general del texto (Swales, 1990); por ejemplo, la función textual de las páginas de información sobre la «Empresa» es en una función expositiva al describir la historia y organización de la misma, con una función secundaria exhortativa con alternativa al intentar que el usuario confíe en ella y establezca una relación comercial con la misma. La función de este bloque es esencial puesto que consolida la veracidad y reputación del emisor del texto.

A su vez, los bloques comunicativos se dividen en secciones. Cada una de ellas tiene una función dentro del bloque en el que se ha incluido (Gamero, 2001: 107);

un ejemplo sería la sección «localización» dentro del bloque «empresa», por la que el usuario puede conocer donde se encuentra la presencia física de la misma.

El procedimiento que se ha seguido para identificar los bloques y secciones se basa en la identificación manual de los enlaces hacia los diferentes bloques en los menús de navegación. La creación de una página web independiente es indicador de la existencia de un bloque comunicativo independiente y sus respectivas secciones (Askehave y Nielsen, 2005). El análisis se basa en una de las peculiaridades básicas de la textualidad web: la necesidad de dividir los contenidos en páginas web debido a la presentación en pantalla (Brinck et al., 2002:284), con la necesidad de que el contenido de cada página represente un bloque comunicativo diferenciado y sea unitario. Argumentamos que esta particularidad da validez a nuestro análisis, y basamos por lo tanto la convencionalización de bloques y secciones en el porcentaje de existencia de páginas web diferenciadas. No obstante, la inexistencia de una página independiente no indica que tal bloque no exista, ya que un bloque comunicativo puede incluirse dentro de otra página web. Como ejemplo, la información de contacto se puede incluir en la página web de inicio; en las páginas web españolas esto sucede en un 13,96 % del total. Los porcentajes por tanto se basan en la existencia de páginas web independientes para tal bloque comunicativo, que relacionamos con la mayor convencionalización del mismo.



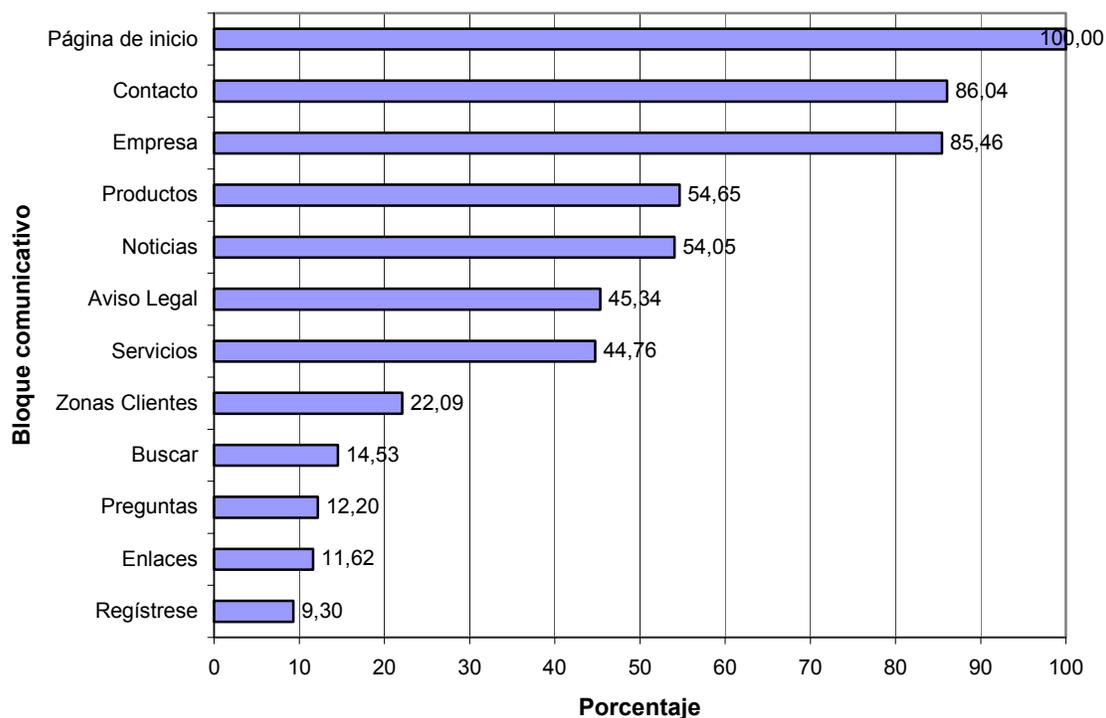
Figura 27. Bloque «contacto» incluido en la página de inicio.

A través del análisis empírico del corpus se han identificado ocho bloques característicos. A su vez, cinco de ellos pueden incluir distintas secciones con distinto grado de convencionalización en este género web. Los bloques comunicativos característicos identificados son:

1. Las páginas de inicio.
2. Páginas de contacto.
3. Páginas con información sobre la empresa.
4. Páginas con información legal.
5. Páginas de noticias.
6. Páginas de productos o servicios, incluyendo catálogos.
7. Zonas exclusivas de clientes.
8. Páginas basadas en la interactividad con el sitio web.

Como comparativa, en el estudio sobre sitios web corporativos anglosajones se identifican siete bloques comunicativos (Nielsen y Tahir, 2002:44-48): 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*).

En la siguiente gráfica se observa el grado de prototipicidad de cada bloque comunicativo en el corpus.



**Tabla 5. Grado de prototipicidad de bloques comunicativos en el género web.**

En la superestructura típica del género se observa la aparición de ciertos bloques con mayor grado de prototipicidad y que por lo tanto son obligatorios, tales como «contacto», «empresa» o «productos/servicios», y otros con menor grado de prototipicidad en la superestructura del género, como son las «noticias» o las diferentes «zonas exclusivas de clientes».

De igual manera, cada bloque comunicativo dentro del género se puede subdividir en secciones con mayor o menor grado de prototipicidad. En la siguiente tabla se analizan los diferentes bloques y secciones del género que se han identificado en el análisis del corpus recopilado. En los estudios de género se identifica la superestructura de un género como base comparativa con la superestructura prototípica en otros idiomas o *locales* diferentes (Gamero, 2001: 109). La siguiente tabla muestra por lo tanto la superestructura típica identificada en el análisis del corpus del género web de estudio y se presenta con la intención de investigar el grado de prototipicidad de los bloques y secciones en la *locale* es-ES. Como se mencionó anteriormente, la metodología empleada para identificar los diferentes bloques se basa en la existencia de páginas web independientes

para cada bloque y sección (Askehave y Nielsen, 2005). El porcentaje que se presenta en la tabla indica en consecuencia la aparición del bloque en el corpus como página independiente:

## **Bloques y Secciones identificados**

1. <b>Página de inicio</b>	<b>100%</b>
2. <b>Contacto [interacción]</b>	<b>86,04%</b>
i. Formulario	31,42%
3. <b>Información general sobre la empresa.</b>	<b>75 %</b>
i. Localización	28,48%
ii. Experiencia de la empresa	17,44%
iii. Misión	10,46
iv. Sistema de calidad	9,30%
v. Historia	8,13%
vi. Instalaciones	7,55%
vii. Logística	6,49%
viii. Proyectos	4,65 %
ix. Galerías de Imágenes	4,65%
x. I+D	4,06%
xi. Divisiones	1,16%
xii. Exportación	0,58%
4. <b>Noticias/Actualidad</b>	<b>54,06%</b>
5. <b>Productos -Servicios-Soluciones</b>	
i. Productos	53,48
ii. Servicios	44,76
iii. Ofertas-promociones	9,88%
6. <b>Legal</b>	<b>45,34%</b>
i. Aviso-Nota Legal	27,90%
ii. Declaración de privacidad	13,37 %
iii. Condiciones generales	4,65%
7. <b>Zonas de Clientes</b>	<b>22,09%</b>
i. Trabajo	19,76%
ii. Consejos	9,30%
iii. Formación	5,81%
iv. Tarifas	5,81%
v. Pedidos	5,23%
vi. Publicaciones	4,65%
vii. Profesionales	2,90%
viii. Presupuestos	2,90%
ix. Inversores	2,90%
x. Franquicia	2,32%
xi. Regalos	0,58%
xii. Financiación	0,58%
xiii. Su opinión	0,58%
8. <b>Interactividad con sitio web</b>	
i. Buscar	14,53%
ii. Preguntas o ayuda [faq]	12,20%
iii. Enlaces	11,62%
iv. Regístrese	9,30%
v. Glosario	1,74%

**Tabla 6. Bloques y secciones identificados en el corpus.**

A continuación se presenta un análisis de los diferentes bloques y secciones, centrandó el análisis en los elementos de la situación comunicativa, la función textual, la funcionalidad y los rasgos convencionales incluyendo la convencionalización de la terminología empleada para la denominación del bloque o sección.

### **9.3.1. Descripción del bloque «Contacto»**

#### **9.3.1.1. Elementos de la situación comunicativa**

El subgénero de contacto se basa en la interacción entre el usuario del sitio web y la empresa que le ofrece productos o servicios, emisora del texto. La comunicación se establece por tanto entre el emisor del texto y el receptor, el usuario. El medio permite que se establezca una comunicación con mayor rapidez (Janoschka, 2003) por medio de correo electrónico, formularios o incluso chats de contacto. De igual manera se incluyen los teléfonos y números de fax de contacto, con lo que se multiplican los medios de comunicación que puede seleccionar el receptor. Con la evolución del género han aparecido los formularios de contacto que agilizan la respuesta, con presencia en el 32,2 % de los sitios web del corpus.

#### **9.3.1.2. Función textual**

La función textual de estas páginas sería expositiva puesto que se presenta principalmente la información básica de contacto de la empresa. Debido a la funcionalidad que presenta el medio la función o foco secundario es exhortativo, ya que los enlaces con el correo de la empresa necesitan invitar al usuario a contactar con la misma (McGovern et al.2002). Los formularios de contacto presentarían una función textual diferenciada marcada por la interactividad del bloque comunicativo. Argumentamos que el foco contextual es exhortativo con alternativa, ya que la función básica es condicionar la «acción» del usuario para que rellene el formulario. El usuario que active el vínculo hacia esta página web llega con este propósito, con lo que se condiciona la aparición de un texto con formas mayoritariamente sustantivas, con los campos a rellenar y los botones «enviar» que representan el componente exhortativo del texto.

☒ Contacto

Ingrese los datos requeridos en el siguiente formulario.

Los campos con ( \* ) son obligatorios

- Nombre y Apellido: \*
- Teléfono:
- Fax:
- E-mail: \*
- Consulta:

Enviar

**Figura 28. Formulario de contacto con una función textual exhortativa. De [www.adea.es](http://www.adea.es)**

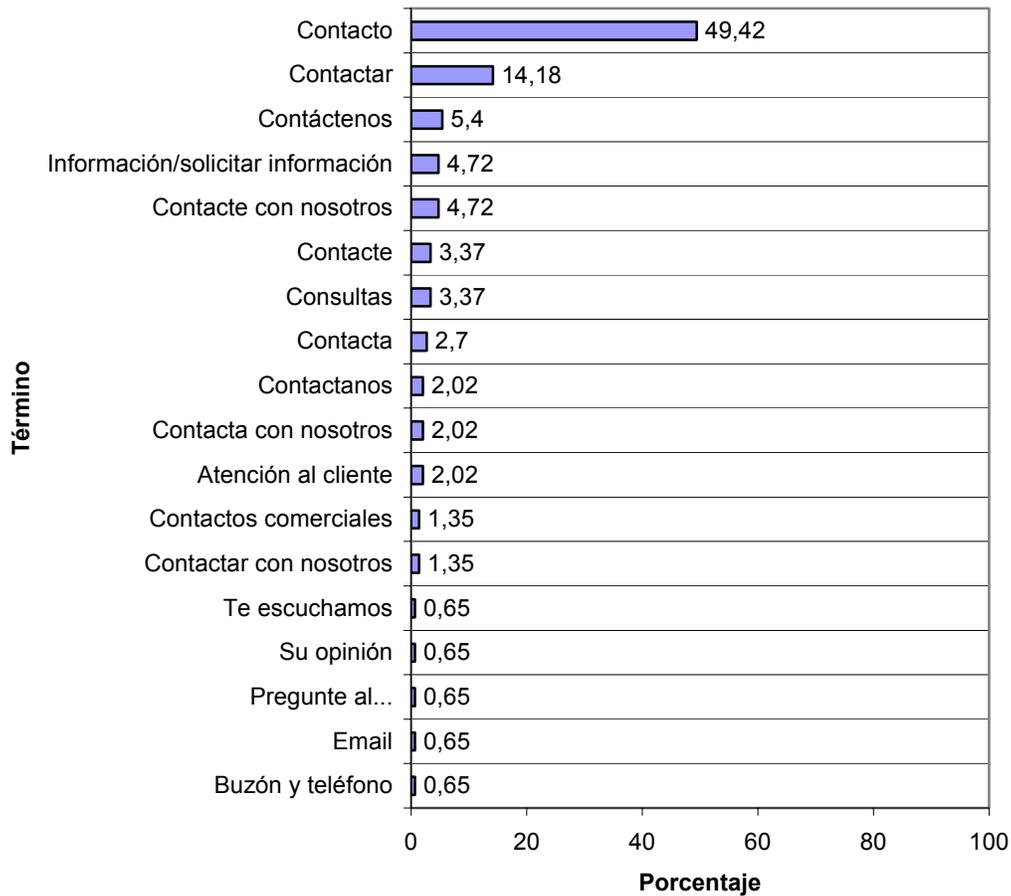
### 9.3.1.3. Rasgos convencionales

En los sitios web españoles se ha convencionalizado la aparición de un bloque de contacto con la existencia de una página separada para tal propósito en un 86,04% de los sitios web. Podemos indicar que este bloque es obligatorio puesto que aunque no se incluya una página independiente para este bloque, en el resto de sitios web se incluye directamente en la página de inicio.

En este bloque se presenta la información de contacto de la empresa. En nuestro caso se presenta el nombre, la dirección de la empresa, teléfonos, fax y el correo electrónico de la misma.

En los sitios web anglosajones se encuentra muy convencionalizado el término *Contact us*, con un 89% de los casos (Nielsen y Tahir, 2002). En nuestro caso se ha convencionalizado su denominación como «Contacto», con un 49,42% relativo de uso del término, un porcentaje mucho menor al anglosajón. En segundo lugar aparece el infinitivo del verbo «contactar», con el 14,18%. Ambas formas personales suponen una forma distinta de realizar el acto ilocutivo.

En la siguiente tabla se observa el grado de convencionalización de los diferentes términos empleados para denominar este bloque comunicativo.



**Tabla 7. Porcentaje de uso de términos para el bloque «Contacto».**

Si tomamos este caso, el bloque comunicativo de mayor aparición, podemos observar que las formas léxicas básicas empleadas en el menú de navegación para «contacto» es mayormente nominal.

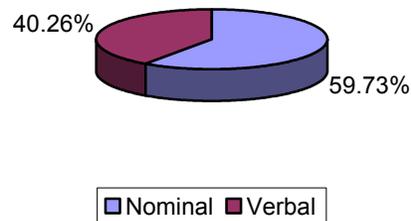


Figura 29. Presencia de formas sustantivadas frente a formas verbales en «contacto».

Son varios los elementos intratextuales que se pueden convencionalizar en cada género, entre ellos el tenor o el grado de formalidad. El análisis de la terminología empleada en este bloque puede revelar el grado de convencionalización de este elemento. Se observa la primacía de un alto grado de formalidad con el uso de sustantivos, 41,66%, o formas verbales con el pronombre usted, 40%, preferentemente.

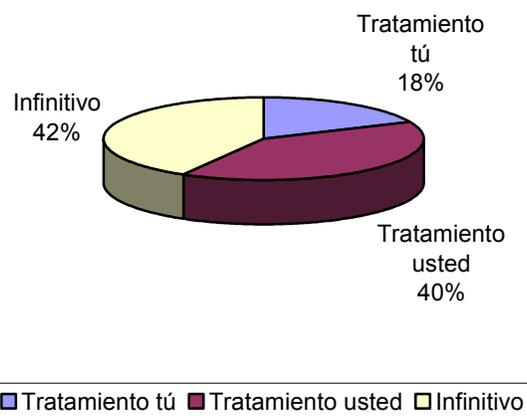


Figura 30. El tenor en los menús de navegación. Extracto de bloque «Contacto».

Solamente en el 18% de los casos se incluye una mención expresa al usuario con el tratamiento de «tú». Este aspecto, la mención expresa del “tú” en los textos web es una recomendación que se encuentra en la totalidad de las guías de estilo web anglosajonas (Kilian, 1999: 14; McGovern et al. 2002). Se observa que en español es más común usar sustantivos o formas verbales en infinitivo para expresar esta función comunicativa, con la preferencia del «usted» en el trato con el usuario o «cliente».

#### **9.3.1.4. Funcionalidad**

La funcionalidad en este caso 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, de chat de contacto y ayuda. Se incluye además la interactividad con el medio en caso de que la dirección de correo electrónico incluya una etiqueta de <mailto: > que produce una ventana de correo emergente y la posibilidad de interactuar con los formularios y páginas.

### **9.4. Descripción del bloque « La empresa»**

#### **9.4.1. Elementos de la situación comunicativa**

La web supone la presentación en un mismo plano a la totalidad de las empresas (Nielsen y Tahir, 2002: 46). En este contexto la empresa como emisora del texto necesita establecer sus credenciales y credibilidad, y a su vez supone un medio para distinguirse frente a la competencia con un sitio web de calidad. Este bloque se centra en proporcionar al usuario una descripción de la organización, presencia física, visión empresarial e historia de la misma. La importancia de este bloque comunicativo es mayor a la de otros textos emitidos por la empresa, como folletos, ya que los usuarios pueden no poseer elementos claros de valoración para entablar o mantener una relación comercial. A pesar de que estas páginas no suelen visitarse, la necesidad de establecer el estatus de la empresa conlleva el imperativo de que este bloque aparezca.

#### **9.4.2. Función textual**

El foco contextual sería expositivo-descriptivo con un foco secundario exhortativo. El foco contextual se basa en la descripción de los diferentes aspectos de importancia de la empresa, mientras que ciertas secciones posibles como los de experiencia o calidad pretenden potenciar la imagen comercial de la empresa y que el usuario confíe tanto en el sitio web como en la empresa en sí. Ciertas secciones como «Experiencia» poseen un foco contextual exhortativo con un foco secundario expositivo ya que la función básica se centra en la publicidad de la empresa, con una información más escasa y una apelación directa a la reacción positiva del receptor con un lenguaje claramente publicitario.

### 9.4.3. Rasgos convencionales

La aparición del bloque comunicativo «Información de la Empresa» se ha convencionalizado con un 85,6% de páginas independientes en los sitios web. En la siguiente gráfica se observa el porcentaje de aparición de los términos empleados para denominar este bloque comunicativo.

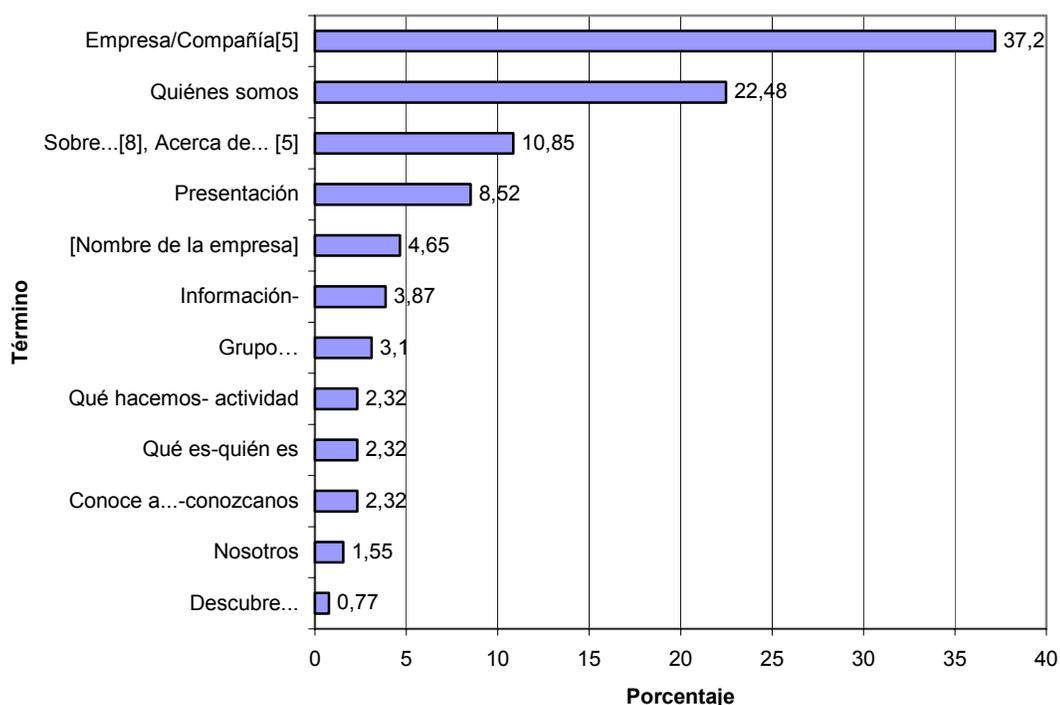
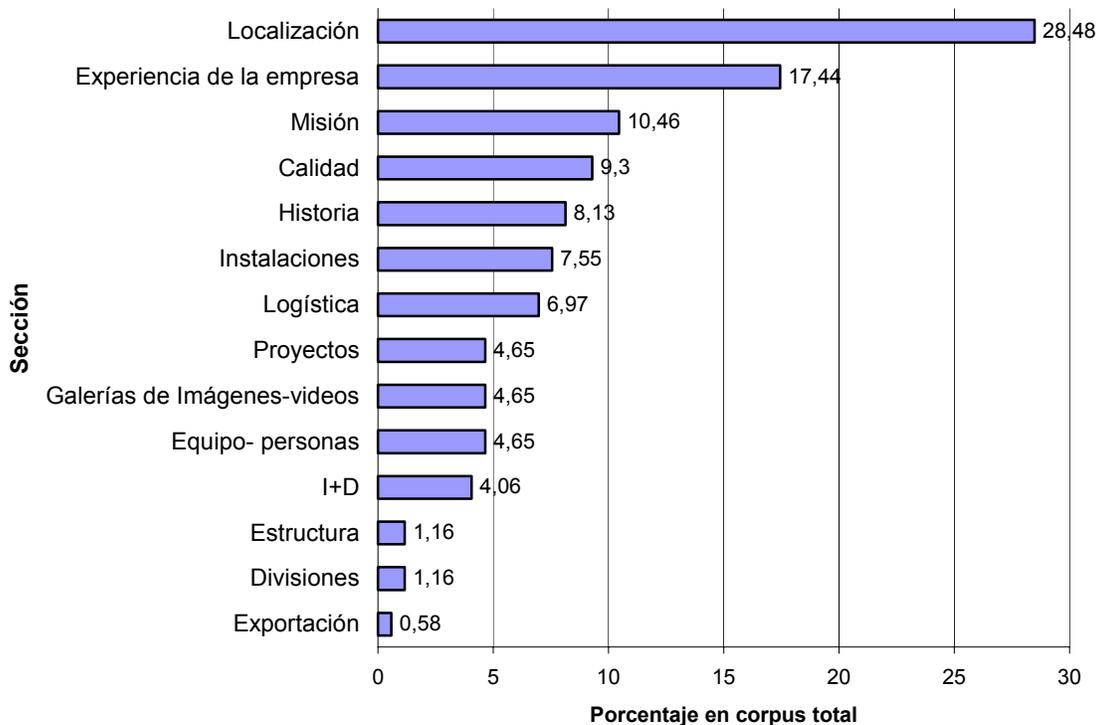


Tabla 8. Términos empleados en la denominación del bloque «Empresa».

El análisis de los términos más convencionalizados para denominar este bloque muestra una primacía de «Empresa», con un 37,2 % de los casos. Este término aparece con ciertas variantes como «La empresa» o «Nuestra empresa» y «Compañía», que incluimos como variante en el anterior pero con sólo el 3,8% del total. En segundo lugar aparece «Quiénes Somos», con un 22,48%. En inglés se encuentra altamente convencionalizado el término *About.../About us* con el 71% (Nielsen y Tahir, 2002). En nuestro corpus los términos sinónimos «Sobre.../Acerca...» aparecen solamente en el 10,85 % de los casos.

En el análisis de los vínculos independientes en el corpus se identifican las diferentes secciones que se pueden encontrar en el bloque comunicativo «La empresa», y se dividen entre la información general de la empresa y los distintos bloques comunicativos posibles:



**Tabla 9. Secciones en bloque comunicativo “Empresa”.**

Se observa que las secciones más convencionalizadas en este bloque son la localización, la experiencia, la calidad, la historia y las instalaciones. En ciertos casos la terminología para esta sección está muy convencionalizada, como en «historia», que aparece como único término posible. El análisis de la terminología empleada para denominar el resto de secciones muestra el grado de convencionalización de cada una.

#### **9.4.3.1. Localización**

Esta sección se basa en mostrar la presencia física de la empresa al usuario. El término «Localización» aparece en el 20,4%, al igual que «Oficinas» y «¿Dónde

estamos/está?» con el 20,4%. Aparece igualmente el sinónimo de localización «Situación» en un 4,08%.

En un gran número de casos se incluye un gráfico con el mapa de situación, y encontramos un 10,09 % de enlaces que incluyen la indicación expresa del término mapa- plano: «Ver mapa de situación»[ 1], «Ver plano de situación[1] / mapa de situación [1]», «Localízanos en el plano» [1]. Igualmente, aparecen los buscadores de oficinas o tiendas con un 10,05% de los casos,e indican la posibilidad de localizar las instalaciones más cercanas: «Busca tu tienda»[1], «Buscador de oficinas»[1], «Localizador de tiendas»[1], «Red de distribuidores» [1], «Concesionarios [SEAT]» [1], «Dónde comprar» [1].

#### 9.4.3.2. Experiencia

Mostrar la experiencia de la empresa es esencial para aquellas que no se encuentran altamente reconocidas, y esta sección sólo aparece en el 17,44% de los sitios web. El término más usado en esta sección es «clientes» con un 36,66% del total. En general se trata de mostrar los clientes de prestigio con los que se relacionan o mantiene una relación comercial para proporcionarles productos o servicios. En segundo lugar de aparición se encuentran «trabajos/servicios realizados», para mostrar la experiencia de la empresa, con un 30% del total. Los términos que aparecen son «trabajos realizados» [3], «servicios realizados»[2], «instalaciones» [3], «dossier» [1]. En esta sección se observa 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 el término «Experiencia», que en este bloque puede ser el elemento diferencial de la empresa. Así, en dos casos aparecen los términos «La diferencia» o «elemento diferencial» para la denominación de esta sección.

#### 9.4.3.3. Misión

Esta sección se basa en presentar los valores subjetivos que representan o a los que aspira la empresa. De nuevo, la función textual sería expositiva-descripción con un foco contextual secundario exhortativo, puesto se pretende condicionar al

usuario para que respete y valore a la misma. Esta sección presenta un índice medio de convencionalización al aparecer en el 10,46% de los sitios web.

La terminología empleada en la denominación de presenta una gran variabilidad con «Misión / nuestra misión» - 27,7% y «Filosofía / nuestra filosofía» con el 16,6%. Asimismo, son varios los términos empleados para la denominación de la sección, 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 los términos empleados, el mayor índice de inclusión del emisor en el término en la totalidad del corpus.

#### **9.4.3.4. Calidad**

La calidad de la empresa es esencial como elemento diferencial y de prestigio de la misma. Esta sección aparece en el 9,30% de los sitios web. El término «calidad» aparece directamente en el 57,25% de los casos, con la posibilidad de incluir «compromiso de...» o «certificados de...». La calidad se puede certificar por medio de normas internacionales: en el 47,75 % de los términos aparece el componente «certificación» o «acreditación». La calidad se puede certificar a través de normas internacionales ISO, y en un caso en el vínculo se denomina directamente con la denominación de las normas, la ISO 9001:2000, y la ISO 14001:2004<sup>133</sup>.

#### **9.4.3.5. Instalaciones**

Las instalaciones son una sección que aparece en el 7,55% del total de sitios web. El término más utilizado es directamente «instalaciones» con el 61,53% de uso. 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 sección «oficinas» incluidas en «localización», puesto que en ciertos casos, aparte de indicar la localización de la empresa, se describen las oficinas en sí mismas, con lo que estas dos secciones pueden aparecer combinadas. Se observa por lo

---

<sup>133</sup> en [www.diode.es](http://www.diode.es)

tanto que en ciertos casos la asignación de una página web a una u otra sección puede no ser totalmente clara, ya que pueden combinarse varias en una misma página. En este caso se procede a asignarla según la mayor presencia de contenido comunicativo de una sección o bloque determinado.

#### **9.4.4. Funcionalidad**

La funcionalidad de este bloque comunicativo se centra en la interactividad entre el sitio web y el usuario. Encontramos mapas de localización interactivos, galerías de imágenes interactivas o ciertas opciones de búsqueda de la localización de oficinas.

#### **9.4.5. Descripción del bloque «Noticias»**

##### **9.4.5.1. Elementos de la situación comunicativa**

Una de las razones por las que los usuarios se dirigen al sitio web es para obtener noticias que se desarrollen en el seno de las empresas, tanto en su estructura como en los productos y servicios que ofrecen. Esta constante actualización de contenidos permite que los usuarios puedan volver con cierta regularidad al sitio web, y por lo tanto, es uno de los mecanismos por los que la empresa puede ampliar el uso y distribución de los textos web que produce.

##### **9.4.5.2. Función textual**

La función básica en este bloque es expositiva, tanto en la descripción como la narración. El foco contextual secundario es exhortativo puesto que pretenden consolidar la reputación de la empresa. En ciertos casos se han encontrado en esta sección folletos informativos publicitarios o artículos comerciales que posee un foco secundario exhortativo (Gamero, 2001; 81).

##### **9.4.5.3. Rasgos convencionales**

El bloque comunicativo noticias se encuentra menos convencionalizado en los sitios web españoles puesto que aparece en el 54,03 % de los sitios.

El término más convencionalizado para este bloque es «Noticias» con un 29,34%. En segundo lugar aparece el término «Novedades», con el 15,21 %, que se suelen referir a noticias sobre los productos o servicios de la empresa. En tercer lugar, con el 15,21%, se observa la metáfora en la que se compara este bloque con una «sala de prensa», con cierta variabilidad en las dos unidades léxicas del término como *sala*, *gabinete*, *centro* o *nota* para el primer elemento y *prensa*, *comunicación* o *información* para el segundo.

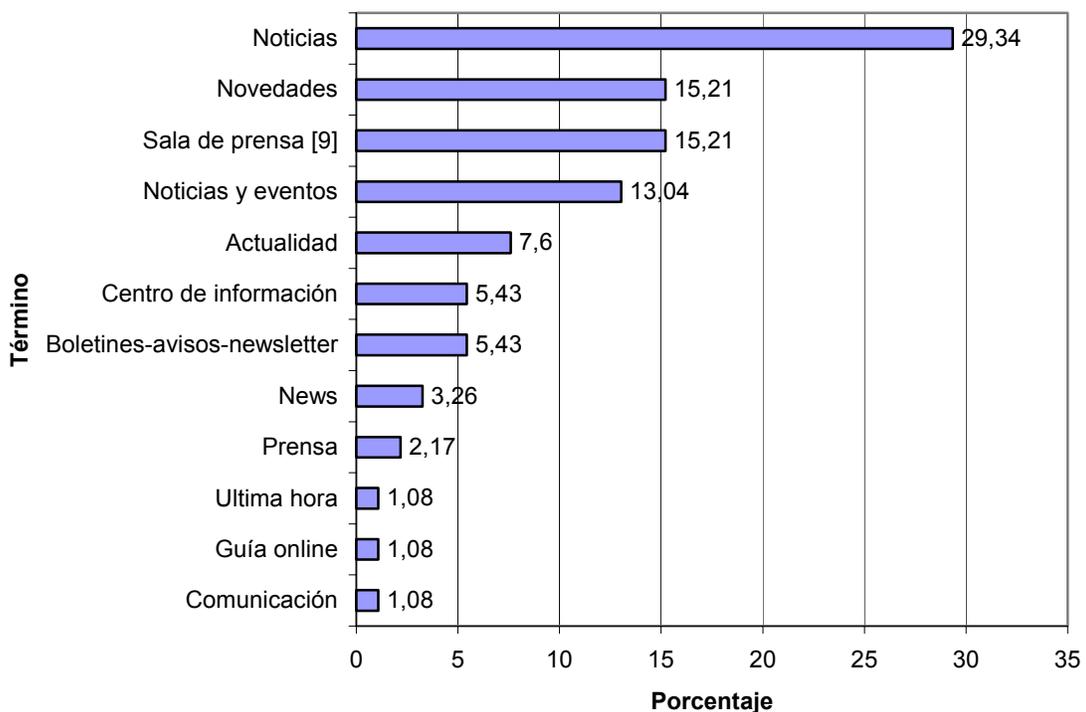


Tabla 10. Términos para la denominación del bloque “Noticias”.

En este bloque aparecen dos términos ingleses, reflejo de la influencia de la cultura anglosajona en el entorno de los negocios, *news* y *newsletter*.

#### 9.4.5.4. Funcionalidad

La funcionalidad específica de este bloque comunicativo se centra en la interactividad del usuario con el hipertexto, tanto en la activación de vínculos

para obtener las noticias como en la lectura selectiva de las mismas o la búsqueda de noticias por medio de fechas o por palabras clave.

#### **9.4.6. Descripción del bloque «Productos-Servicios»**

##### **9.4.6.1. Elementos de la situación comunicativa**

Las empresas basan su actividad económica en la venta de productos o servicios, y por lo tanto, 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 post-venta. La empresa como emisora del texto utiliza la web como plataforma adicional<sup>134</sup> de presentación y publicidad de sus productos o servicios, a la que acceden los usuarios conscientes de las funcionalidades que proporciona el medio.

##### **9.4.6.2. Función textual**

El foco contextual de este bloque es expositivo en descripción con un foco contextual exhortativo. La función básica de este bloque se centra en la descripción de los productos y servicios que ofrece la empresa que a través de una presentación atractiva que pretende la reacción positiva del usuario. En los casos en los que se presenten descripciones muy breves con una gran carga publicitaria el foco contextual es exhortativo con un foco contextual secundario expositivo, ya que supondría una descripción publicitaria informativa. Trosborg (1997:14), indica la necesidad de incluir esta multifuncionalidad en los textos, y propone este caso como ejemplo, ya que la publicidad o descripción de los productos puede ser ante todo persuasiva, y por lo tanto con una función exhortativa, o puede ser esencialmente descriptiva, sobre todo en las descripciones más técnicas o especializadas de los productos.

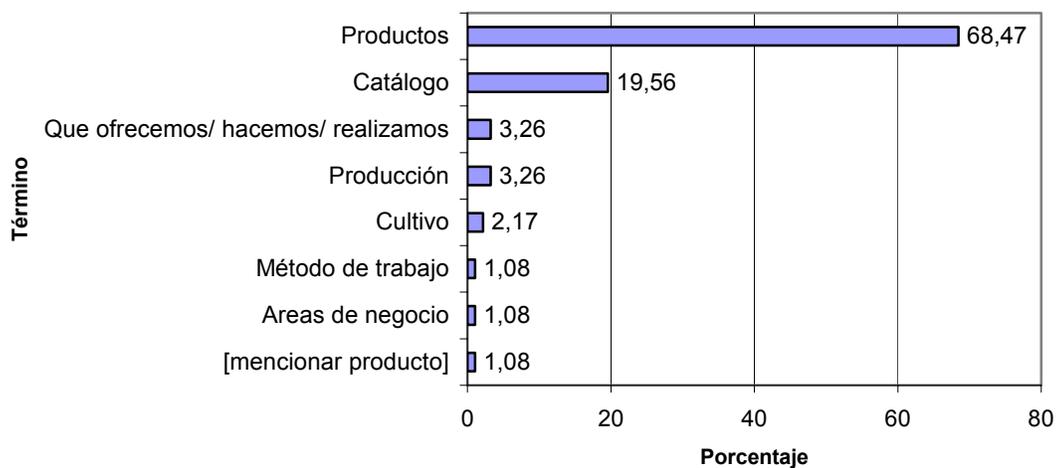
##### **9.4.6.3. Rasgos convencionales**

---

<sup>134</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)

La aparición de este bloque comunicativo es esencial en este género y se ha convencionalizado, con un 54,48% de sitios web con enlaces a «productos», un 44,76% con «servicios», y un 9,88% de sitios con la aparición de la sección «Ofertas-Promociones». La combinación de las tres posibles secciones sobrepasa el 100%, ya que en varios casos se presentan combinados en un mismo sitio web, productos y servicios, o productos y ofertas dependiendo de la actividad económica de la empresa. Las ofertas pueden constituir tanto productos como servicios; en el caso de [www.adeco.es](http://www.adeco.es) se mencionan las ofertas de trabajo, ya que constituye una empresa de empleo temporal, el servicio que ofrece.

En el caso de los productos, se encuentra convencionalizado el término «productos» con el 68,47% del total. Este término presenta dos variantes, «nuestros productos» con 4 apariciones en el corpus y «producto» en singular con una sola aparición. Esta sección constituye el catálogo de productos *online* de la empresa con las funcionalidades adicionales del medio, por lo que encontramos el término «catálogo» en un 19,56 % de los casos. En varios de estos casos la empresa a coloca en formato .html o en .pdf su propio catálogo de productos, que constituiría un género *extant* al simplemente volcar en Internet un texto ya existente en otro medio.



**Tabla 11. Frecuencia términos en la denominación de bloque «Productos».**

En el caso de los servicios, este término se encuentra altamente convencionalizado, con el 88,31% del total. El siguiente término es «soluciones», con presencia en el 6,49% del total que se asocia a ciertos campos como el informático.

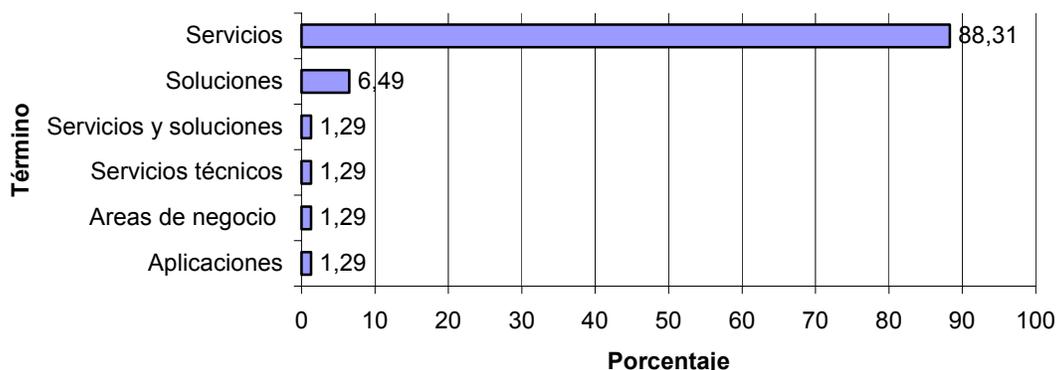


Tabla 12. Frecuencia términos en la denominación de bloque “Servicios”.

La única sección diferenciada que aparece en este bloque comunicativo sería «ofertas»: este mismo término se ha convencionalizado ( 70,58%).

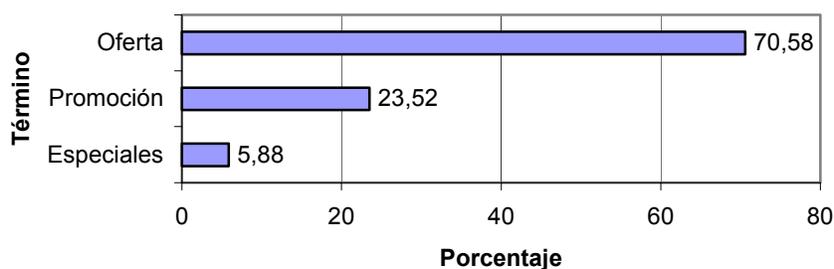


Tabla 13. Frecuencia de términos en la denominación de la sección “Ofertas”.

#### 9.4.6.4. Funcionalidad

La funcionalidad que presenta este bloque comunicativo se centra en la interacción que proporciona el medio, búsquedas de productos en el catálogo, la inclusión de gráficos interactivos o en diferentes tamaños, videos promocionales de los productos etc. , así como la posibilidad de configuración de ciertos productos por parte del usuario con la presentación final del modelo elegido y/o el precio total.

Ciertos sitios web presentan la posibilidad de compra o contratación en el mismo sitio web por medio del comercio *online*. No obstante, los web corporativos centrados exclusivamente en el comercio *online* poseen unas características

discursivas y estructurales distintas por lo que constituyen un género diferenciado.

### **9.4.7. Descripción del bloque «Legal»**

#### **9.4.7.1. Elementos de la situación comunicativa**

La actividad económica de la empresa que se constituye en emisora del texto se encuentra sometida a la legislación vigente en España e inmersa en ésta. Este aspecto de la situación comunicativa y sociocultural incide en la aparición de este bloque comunicativo que recoge el estatus y las responsabilidades jurídicas de la empresa emisora y de los usuarios del sitio web.

En el análisis de los avisos legales en el corpus se han identificado tres posibles relaciones legales que condicionan este bloque.

- 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 legal entre el usuario y la empresa por cuanto a la confidencialidad 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](#)), existiendo además la Agencia Española de Protección de Datos<sup>135</sup>. La totalidad de los avisos legales incluyen una cláusula indicando el cumplimiento de las disposiciones establecidas en esta ley.
- La relación jurídica vinculante entre la emisora del texto y el receptor del mismo por cuanto al texto web, expresado por en las «condiciones de uso».
- En el caso de localizaciones desde un país fuera de la Unión Europea, la legislación vigente de protección de datos para empresas que tenga presencia en la misma sería la Directiva Europea de Protección de Datos

---

<sup>135</sup> <https://www.agpd.es/index.php>

95/46/CE<sup>136</sup> del Parlamento Europeo y del Consejo del 24 de octubre de 1995. Esta directiva regula la protección y uso de los datos personales recogidos tanto dentro como fuera de la Unión Europea.

#### 9.4.7.2. Función textual

Como texto jurídico, la función textual de los avisos legales es exhortativa sin alternativa ya que es vinculante jurídicamente. El propósito retórico sirve simultáneamente dos funciones: informar al usuario de la relación entre la empresa y la normativa legal española, e igualmente el tratamiento de la información personal que proporcione el usuario.

#### 9.4.7.3. Rasgos convencionales

La forma más convencionalizada en el corpus es «aviso legal», que es el bloque que engloba la totalidad de información legal que proporciona la empresa en el sitio web, incluyendo la protección de datos y las condiciones de uso, y aparece en un 27,90% de los sitios web. La privacidad es una sección dentro del bloque de menor frecuencia, el 13,37%, al igual que las condiciones de uso, sólo un 4,63% de total de sitios web. Se observa que «aviso legal» correspondería con un bloque comunicativo en el que se pueden incluir estas otras dos secciones. En la siguiente gráfica se observa la distribución de las secciones en este bloque comunicativo:

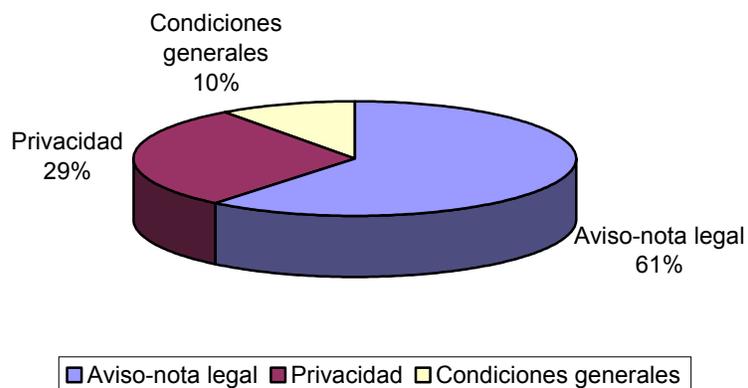


Tabla 14. Porcentaje de distribución del bloque «Legal».

---

<sup>136</sup> [http://ec.europa.eu/justice\\_home/fsj/privacy/](http://ec.europa.eu/justice_home/fsj/privacy/)

En la siguiente tabla se observa la distribución de los bloques y secciones que se encontraron en el corpus:

Bloque.-Legal	Número de aparición global en el corpus	Porcentaje
<b>Aviso-nota legal</b>	<b>48</b>	<b>27,90</b>
Aviso legal	33	19,18
Nota legal	6	3,48
Información legal	2	1,16
Advertencia legal	1	0,58
<b>Privacidad</b>	<b>23</b>	<b>13,37</b>
Política de privacidad	13	7,55
Privacidad	5	2,90
Protección de datos-política de protección de datos	4	2,32
Confidencialidad	1	0,58
<b>Condiciones generales</b>	<b>8</b>	<b>4,62</b>
Condiciones de uso	3	1,74
Términos y condiciones	1	0,58
Términos de uso	1	0,58
Condiciones	1	0,58
Normas de uso	1	0,58
Condiciones generales	1	0,58

Tabla 15. Porcentaje de aparición en el bloque «Aviso Legal» y terminología empleada.

En el análisis porcentual de la terminología empleada en este bloque se observa la convencionalización del término «aviso legal» con un 41,77% relativo de uso. El término «nota legal» aparece en el 7,59%. La sección de «política de privacidad» que aparece en un 16,45% sería un reflejo de la cultura anglosajona puesto que la normativa legal española establece claramente el tratamiento de los datos personales, con lo que la empresa no establece su propia política en este caso. En el corpus se observa la aparición de «Protección de datos» con un 5,06%, incorporando la terminología empleada en la ley vigente en España (Ley de Protección de Datos). Las condiciones de uso del sitio web se encuentran muy convencionalizadas en los sitios web anglosajones pero sólo se encontraron en el 4,62% en el corpus. Se puede por lo tanto concluir que esta sección no se encuentra convencionalizada en los sitios web españoles, y por lo tanto no aparece un término convencionalizado para su denominación, caracterizada por una gran variabilidad:

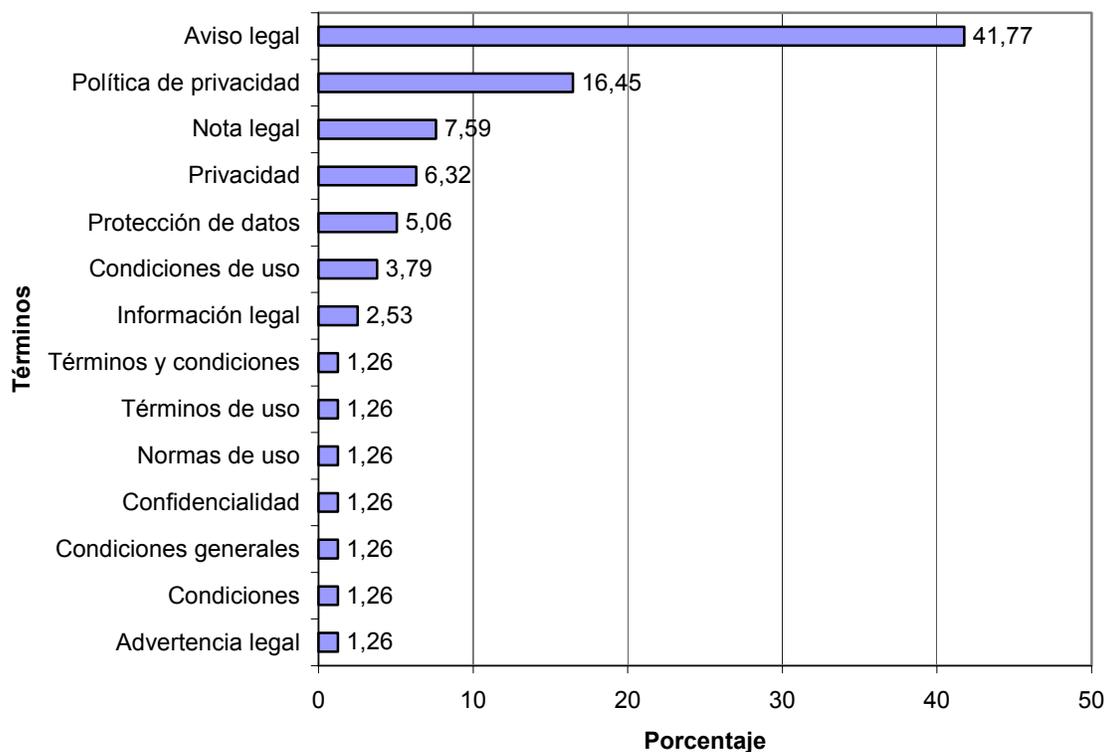


Tabla 16. Frecuencia de términos en la denominación del bloque «aviso legal».

#### 9.4.7.4. Funcionalidad

La funcionalidad en este bloque comunicativo se reduce a la interacción entre el hipertexto y el usuario, con la lectura selectiva de material y la activación de vínculos en el mismo. Se pueden encontrar vínculos hacia las fuentes de la legislación vigente que se menciona, así como un correo electrónico al que escribir en caso de modificar o eliminar los datos personales en posesión de la empresa.

#### 9.4.8. Descripción del bloque «Zona reservada-especial»

El bloque comunicativo «Zona especial» se basa en la inclusión de bloques que se encuentran especialmente destinados a ciertos usuarios, siendo necesario

registrarse en ciertos casos: el 22,09% del total de sitios web incluyen una zona reservada con contenidos especiales para clientes. El resto de zonas especiales, que se corresponden con secciones dentro de este bloque, cumplen una función específica destinada al directamente al usuario, como puede ser la búsqueda de trabajo, la solicitud de presupuestos o los consejos de distinta índole. La siguiente tabla muestra el grado de prototipicidad de las siguientes secciones, apareciendo las «zonas de clientes», 22,9% y la sección «trabajo», con el 19,76 % como las de mayor grado de convencionalización.

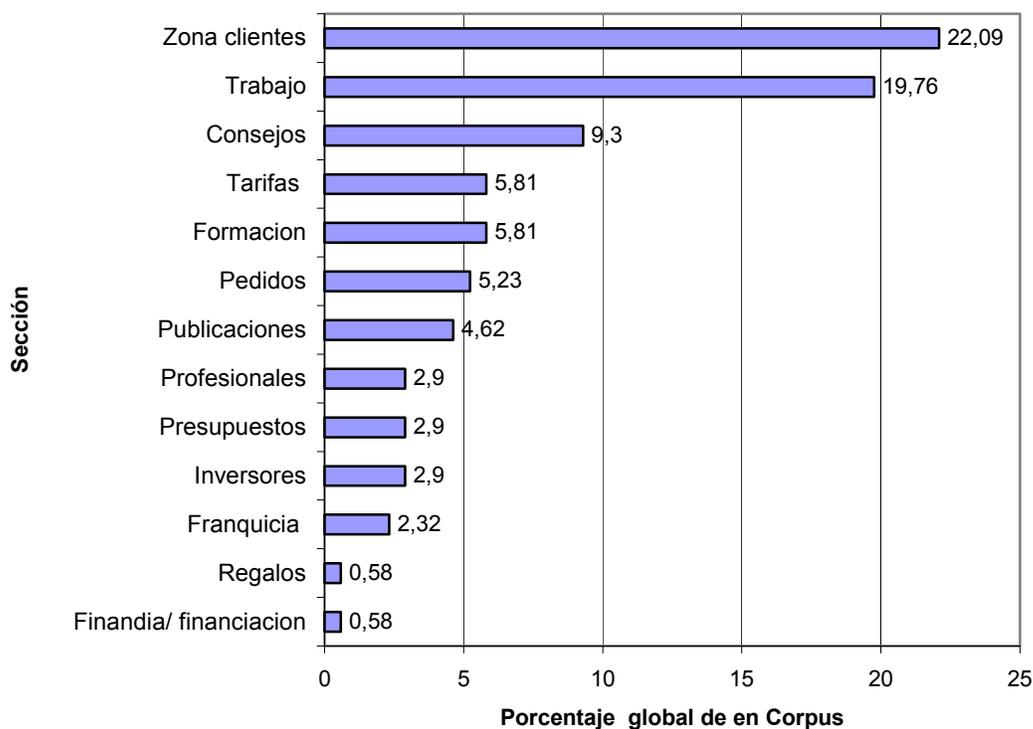


Tabla 17. Grado de prototipicidad de las secciones en bloque «Zona reservada».

El grado de convencionalización de la terminología empleada en las distintas secciones varía en gran medida, con ciertas secciones que emplean un término único o con escasa variación como «franquicia», «inversores» o «presupuestos», y otras secciones en las que presentan una gran variabilidad terminológica como puede ser la sección «trabajo», «consejos», en las que no se encuentra ningún término con un porcentaje de uso superior al 9%.

En la sección «Zona reservada» el término al que se denomina al usuario depende de la relación entre el mismo y la empresa. El mismo puede ser el

«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 acceden de igual manera se denomina de diferentes maneras: «Acceso...» - 29,72%, «Área»- 21,72%, «Zona...»-16,37%, o «Club...» con un 5,40%. Globalmente los términos más empleados son «Acceso clientes» 24,32%, «Clientes» - 10,81%, «Área de distribuidores» - 10,81%, «Área de clientes» -8,01% y «Zona de clientes» 8,01%.

La sección «Trabajo» es la zona que presenta una mayor variación denominativa del conjunto de las secciones del corpus, con tan sólo un 8,04% de uso en término más usado: «Bolsa de empleo». Los siguientes términos empleados con un 5,88% son: «Trabaja con nosotros», «Trabaje con nosotros"», «Trabajo», «Envíanos tu C.V.», «Recursos Humanos», «Empleo» y «Ofertas de empleo». En esta sección se encuentra el mayor índice de uso del tratamiento «tú» en la totalidad de las secciones. Concretamente, del total de formas verbales usadas en la denominación de esta sección, el 70% de las mismas usa el tratamiento «tú», el 20% el tratamiento de «usted», y el 10% la forma infinitiva.

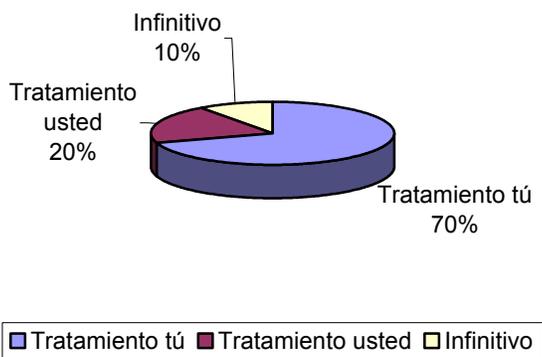


Tabla 18. El tenor digital. Tratamiento en la sección «Trabajo».

Se observa un contraste claro con otros bloques como «Contacto» (véase 9.3.1.3) en los que prima el tratamiento de «usted», 40%, o el infinitivo, 42%, con sólo el 18% de tratamiento de tú. En este caso, la necesidad de persuadir al usuario para que trabaje con la compañía resulta en un tratamiento más informal, con lo que se puede concluir que el grado de formalidad puede depender del bloque en el que se encuentre.

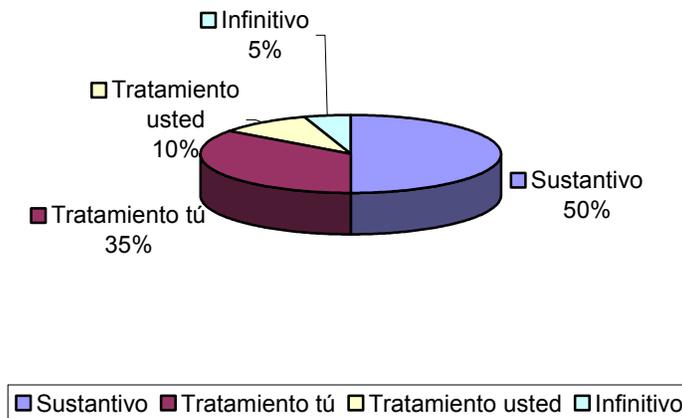


Tabla 19 . Tenor digital en la sección «Trabajo». Sustantivo o forma verbal.

La sección «Formación» aparece en un 5,81% de los sitios web del corpus. En la sección aparece tanto este término como el término «cursos», ambos con un 40% del total de términos en esta sección.

La sección «Tarifas», con un 5,81% de aparición global presenta igualmente tres términos: «Tarifas», «Precios», «Lista de precios», sin ningún término convencionalizado.

## 9.4.9. Descripción del bloque «Interactividad»

### 9.4.9.1. Elementos de la situación comunicativa

En las secciones de este bloque la comunicación se establece entre el sitio web y el usuario o receptor, por lo que se establece la interactividad que proporciona el medio (Janoschka, 2003). Este intercambio comunicativo está limitado a las posibilidades que incluya el desarrollador del sitio web, tales como búsquedas en el sitio de distinta índole, ayuda con el uso del sitio web o la posibilidad de registrarse como usuario para acceder a contenidos o funcionalidades reservadas.

### 9.4.9.2. Función textual

La función textual en casi todas las secciones es exhortativa con alternativa. El fin de estas secciones es que el usuario «actue» y sea parte participante activa en el intercambio comunicativo. En dos bloques identificados, las preguntas y el glosario, la función textual primaria sería expositiva con una función secundaria exhortativa sin alternativa, ya que se centran en informar al usuario.

### 9.4.9.3. Rasgos convencionales

Este bloque presenta unas secciones de menor grado de convencionalización en la superestructura del sitio web corporativo en España. La sección «Buscar» sería la de mayor convencionalización con un 14,35% de media en los sitios web españoles, seguida de las «preguntas o ayuda» sobre el sitio web con un 12,2% y los «enlaces» con un 11,72%.

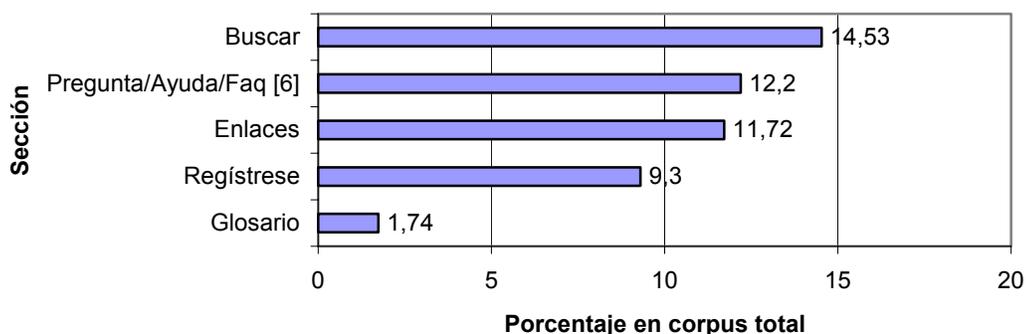


Tabla 20. Grado de prototipicidad de las secciones en bloque «Interactividad».

La sección «Búsqueda» se puede centrar tanto en la búsqueda de elementos dentro del propio sitio web, de productos o servicios que ofrece la empresa, o de oficinas o sedes de la misma en el país. El término convencionalizado para el primer caso es «Buscar», con un 80% del total. El estudio de (Nielsen y Tahir, 2002) presentan un 42% de aparición del término *Search* y un 40% de *Go*, recomendando finalmente *Search* por motivos de usabilidad. De interés es la aparición del término «Encontrar», con un 4% que supondría incluir el resultado del acto ilocutivo en la función realizada.

La sección «Ayuda» aparece en el 11.72% de los sitios web españoles. Los términos «ayuda» y «FAQ» se encuentran ambos ligeramente convencionalizados con un 24% de porcentaje de uso. Las «FAQ» se han presentado como un género web distintivo convencionalizado en el contexto anglosajón (Santini, 2005a). La traducción de este término, «Preguntas

Frecuentes» se encuentra en el 12% de en esta sección. Al igual que en el caso de «Política de Privacidad» se observa una influencia de la cultura anglosajona en la superestructura del género español debido en parte al uso de guías de desarrollo y estilo web anglosajonas traducidas al español (McGovern et al, 2002; Nielsen, 2000; Kilian, 1999).

El 11,72% de los sitios web incluyen una sección de «Enlaces», denominada con este mismo término en el 60% de los casos. Aparece igualmente el término «Links» en el 35% de los casos, un alto índice que muestra de nuevo la influencia de la cultura anglosajona en los textos web españoles. De igual manera se encontró un término que cláramente indica el objetivo de la sección: «Web recomendadas».

La sección de «Registro», que aparece en el 9,3% de los sitios web corporativos, presenta como término más convencionalizado «Regístrese», con un 25% de uso. «Hágase cliente», «Registro» con el 12,5 cada una seguirían al anterior. Se observa de nuevo una primacía del tratamiento de usted en el género de estudio. Esta sección presenta una gran variabilidad en los términos, como «Entrar en cuenta / crear una cuenta nueva», «Actívate», «Mi cuenta», «Registrarse», «Regístrate» o «Hágase socio».

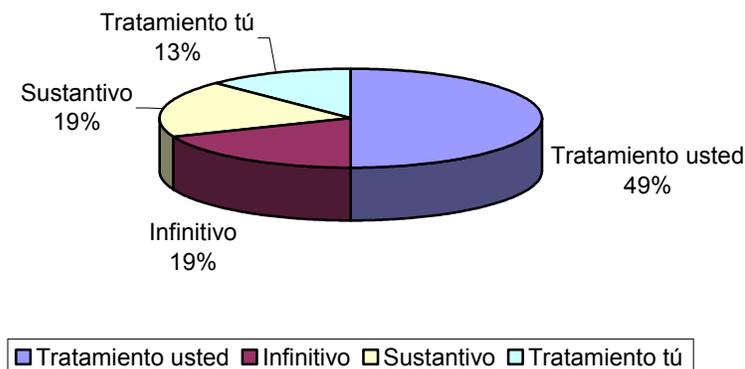


Tabla 21. El tenor digital en la sección «Registro».

Este caso, puesto que supone una apelación directa a que el usuario interactúe con el sitio web, es otro caso en el que se puede observar la convencionalización del tenor digital en el género, con una clara primacía del tratamiento de usted, con el 49%, o el infinitivo, con el 19%. Estos resultados se asemejan a los obtenidos anteriormente en el bloque «contacto».

#### **9.4.9.4. Funcionalidad**

Este bloque comunicativo se basa en la funcionalidad, en este caso la interactividad entre el usuario y el sitio web. Como se indicó anteriormente, la interactividad se centra en la posibilidad de realizar búsquedas en el sitio, registrarse para obtener contenidos exclusivos o realizar transacciones con la misma, o proporcionar ayuda al usuario con ciertas funcionalidades o aspectos del sitio web.

## 10. CONCLUSIÓN

El propósito de la presente investigación ha sido dilucidar los diferentes elementos que afectan al proceso de traducción en localización web y delimitar un marco de estudio claro de la textualidad en hipertextos. En especial, se aplica y adapta el concepto de *género* al estudio de esta textualidad, centrándonos en la descripción del género *sitio web corporativos* o «*corporate homepage*» en una *locale* determinada: español de España o es-ES.

Los *géneros* representan un concepto menos explorado en los estudios de traductología en comparación la *tipología textual*, concepto que se ha usado durante años como cajón de sastre. El género posee una larga tradición en los estudios literarios y surge de la capacidad de los hablantes de una lengua de reconocer las características y funciones de unos textos prototípicos que se repiten en una ocasión social determinada y suelen convencionalizarse para mejorar la efectividad comunicativa de los mismos. Los géneros combinan tres aspectos: los formales, los socioculturales y los cognitivos. Por cuanto a los aspectos formales, los textos representan formas convencionalizadas que se expresan a través de movimientos o etapas (Swales, 1990) que reconocen los hablantes de una comunidad de discurso y pueden variar entre diferentes lenguas y culturas. Debido a las peculiaridades de la textualidad web, estos movimientos se pueden asociar 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 se ha denominado Modelo de *Comunicación de Masas Interactivo* (Janoschka, 2003). Los géneros se pueden además convencionalizar de manera diferente en cada cultura, de ahí la importancia de los mismos en la labor del traductor. En la presente investigación, a pesar de la influencia de la cultura anglosajona por medio de manuales de estilo (Kilian, 2000; McGovern 2002) o de desarrollo y usabilidad (Nielsen, 2000; Nielsen y Tahir, 2002; Brinck et al. 2002; Krug, 2005), se observan diferencias básicas entre las convenciones textuales, de estilo y formato entre los sitios anglosajones y españoles, siendo ésta una de las posibles vías de profundización en la investigación. Por último, los aspectos cognitivos se basan en los propósitos de los hablantes en un intercambio comunicativo. En los sitios web, esta vertiente cognitiva se centra en las expectativas de los usuarios que debe tener en cuenta el emisor y traductor de un texto, y en este caso hay que añadir las expectativas por cuanto a la *funcionalidad*,

o cómo funcionan los textos web (Nielsen y Tahir, 2002). Esta funcionalidad se ha definido por medio de la *interacción* y la *interactividad* (Janoschka, 2003), es decir, entre la comunicación máquina-usuario o emisor-receptor.

La distinción clara entre los conceptos de género y *tipología textual* es esencial debido a la confusión terminológica que se puede observar en traductología. En el capítulo IV señalamos que las diferencias entre género y tipología textual se basan en que el primero se define por factores externos y socioculturales, mientras que los tipos textuales se definen por sus características lingüísticas internas (Trosborg, 1997). Ambos se encuentran asimismo convencionalizados en distintas lenguas: el género por cuanto a su estructura externa y su proyección sociocultural y la tipología por cuanto a su estructura lingüística. En el presente estudio se parte del concepto de género para el establecimiento de un marco de estudio contrastivo sobre la tipología textual en los géneros web. Esta metodología que combina ambos conceptos puede ser de gran utilidad al traductor (Trosborg, 1997: viii).

El estudio y delimitación en el capítulo IV del *cibergénero* sitio web corporativo o *corporate homepage* ha mostrado que representa un género *novel* (Kennedy y Shepherd, 1998), ya que no existe en ningún otro medio, y al contrario de lo concluido en otras investigaciones, va más allá de la simple evolución del folleto publicitario (Fernández, 2003) o catálogo empresarial (Rennau, 2004:43) al establecerse como un género propio. El análisis en el capítulo IX de este género muestra unos bloques comunicativos que se encuentran altamente convencionalizados (Swales 1990; Gamero, 2001: 107), como «Página de Inicio», «Contacto», «Empresa» o «Productos-Servicios», otros bloques con un grado de convencionalización media y por lo tanto opcionales, como «Noticias», «Avisos Legales», y finalmente otros de menor grado de prototipicidad como las «Zonas de clientes», o los bloques que se basan en la «Interactividad» con el sitio web. A su vez, cada bloque comunicativo se subdivide en ciertas secciones con distinto grado de prototipicidad, tales como la sección «Localización» en el bloque «Empresa», con un 28,48% de aparición, o la sección «Trabajo» en el bloque «Zonas de clientes», con un 19,76%.

El estudio de la terminología empleada en la denominación de cada bloque o sección muestra ciertos bloques cuya terminología está altamente convencionalizada, tales como «contacto», 49,42%, o «avisos legales», 61%, y otros bloques cuya terminología se caracteriza por la variabilidad en la denominación, como en la sección «trabajo», en la que el término más empleado no sobrepasa el 8,04% de uso. Esto contrasta con el grado de convencionalización

en sitios web corporativos anglosajones (Nielsen y Tahir, 2002), en los que *contact us* se emplea en un 89% de los casos o *jobs* en un 41%.

En el capítulo II se ha investigado el proceso de *localización* y sus efectos en la tarea de la traducción. En general, se observa cómo la industria de la localización se ha desarrollado instaurando sus propios modelos dejando a un lado el cuerpo de conocimiento de la traductología (O'Hagan y Ashworth, 2003). Además, la mayoría de los estudios desde una perspectiva traductológica se centran en la faceta técnica del proceso (Reineke, 2005; Mata, 2005), que se supone de mayor complejidad, sin que aparezcan estudios en profundidad centrados en las peculiaridades textuales, lingüísticas y contrastivas de este tipo de textos (Quirion, 2003). En la presente investigación se reclama la importancia de la base traductológica del proceso, en el que es imposible separar un componente lingüístico de una adaptación cultural o técnica, al no suponer más que una proyección hacia el futuro de la profesión del traductor (Quirion, 2003). En nuestra opinión, el complejo proceso de localización no se caracteriza exclusivamente por la faceta técnica que la industria reclama como distintiva (Pym, 2004: 25), sino que la localización se caracteriza igualmente por un nuevo modelo de comunicación que se define por el medio, un lenguaje o ciberlenguaje de características específicas (Crystal, 2001), un proceso de traducción caracterizado por el grado de localización que se encargue (Brooks, 2000) y un texto cuyos segmentos se encuentran organizados o no según criterios informáticos y no textuales.

Se ha establecido que las investigaciones en localización desde un punto de vista traductológico necesitan descomponer el proceso y realizar estudios que se centren en aspectos específicos. Sólo a partir de la división del objeto de estudio y de una perspectiva claramente traductológica se podrá reclamar la importancia de esta faceta en la localización. En parte, la dificultad de realizar investigaciones desde esta perspectiva se basa en la inexistencia de una delimitación clara de qué constituye un *texto en localización*. En el capítulo V se delimita y define que constituye un texto web, ya que el centro de la metodología es la recopilación y análisis de un corpus de tal tipo de textos, unidad de análisis de difícil delimitación debido a las características de los hipertextos. Los textos en localización se han definido como aquellos que se distribuye como una unidad y que se desarrollan o presentan de modo unitario, existiendo claras diferencias entre los textos en software y los *textos web*. En la necesaria delimitación del texto web se distingue entre los *E-textos*, los *hipertextos* y la *hiperweb* (Storrer, 2002). Tanto los *E-textos* como la *hiperweb* son de menor interés desde un punto de vista traductológico, siendo los *hipertextos* los que se establecen como una unidad clara de análisis y traducción. Estos hipertextos se delimitan por medio de una

función textual reconocible y una consistencia temática, como puede ser un sitio web ([www.jazztel.es](http://www.jazztel.es)), aunque son abiertos ya que los emisores pueden añadir más nodos. Las características básicas de los hipertextos en general se centran en la multilinealidad, la aparente falta de límites y sus implicaciones en el proceso de producción y recepción textual. De entre los siete parámetros que definen un texto (Beaudegrande y Dressler, 1981), los más importantes en los hipertextos son la coherencia y la cohesión, siendo la primera dirigida por el receptor, o *process-oriented* (Janoschka, 2003).

La investigación empírica, una vez definidos y delimitados los conceptos clave, se basa en la recopilación y de un *corpus* de hipertextos que se corresponden con el género web de mayor importancia por cuanto al volumen de traducción, el *sitio web corporativo*, y su posterior análisis con técnicas de lingüísticas de corpus. En el capítulo VIII se detallan la composición, criterios de recopilación y limitaciones del corpus. En la recopilación se siguieron los criterios de representatividad, estandarización y tipología textual. La representatividad, definida como la posibilidad de que todos los textos que forman parte de la población de estudio tengan las mismas posibilidades de formar parte del mismo, se estableció por medio del directorio de *Negocios* de *Google.com* con lo que se consiguió una representación de todas las áreas de negocio. No obstante, son varias las limitaciones en la recopilación y el análisis de hipertextos completos:

- La imposibilidad de recopilar *hipertextos* (sitios web) completos debido a la existencia de intranets o arquitecturas servidor-usuario que impiden la descarga completa de los mismos, y en ciertos casos la inestabilidad de ciertos tipos de textos web como los *virtuales*, distintos en cada sesión o acceso al sitio web.
- La imposibilidad de recopilar ciertos textos web sincrónicamente por las peculiaridades tanto del software empleado, Htrack, como de los textos web por el uso de Macromedia Flash, marcos y *scripts*.
- La existencia de elementos textuales codificados en imágenes o animaciones Flash que impiden su análisis con los paquetes de software de lingüística de corpus, tales como *Wordsmith Tools* o *Monoconc*. La extracción de elementos textuales manualmente de la animaciones Flash es posible con ciertos programas de software y podría ser una vía de futuros estudios.

- Las limitaciones que representa la recopilación sincrónica de ciertos tipos de textos web de grandes dimensiones que suele ser imposible conocer a priori, y la repetición de elementos textuales en cada página web. Ambas situaciones pueden distorsionar las estadísticas y por lo tanto invalidar los resultados del análisis.

El análisis del corpus recopilado, en capítulo IX, parte de perspectiva traductológica y se centra en la superestructura prototípica del género, en el grado de convencionalización de ciertos elementos intratextuales y la terminología empleada en la denominación de los bloques o secciones y por último el tenor digital. Éste último se analiza preliminarmente, y aun cuando en general se ha caracterizado por una reducción en el grado de formalidad (Janoschka, 2003; Giménez, 2000), en el género de estudio el trato suele ser formal, con lo que se puede concluir que el tenor digital se encuentra convencionalizado de manera diferente en cada cibergénero. En este caso se caracteriza por una preferencia del tratamiento de *usted* o la forma impersonal del *infinitivo* para dirigirse directamente al usuario. Por ejemplo, en el caso del bloque de mayor prototipicidad global, «Contacto», el infinitivo se usa en el 42% de los casos, el tratamiento de usted en el 40%, y el tratamiento tú en un 18%. En el caso de la sección «Regístrate», el tenor digital muestra unos resultados similares con la primacía del tratamiento de usted o el infinitivo. No obstante, en el presente trabajo se han observado claras diferencias dentro del mismo género, ya que en el caso de una sección de menor grado de prototipicidad como «Trabajo», el tratamiento preferido que se observa es el de tú, con un 70% de los casos, con un 20% de uso de usted y un 10% de uso del infinitivo.

En resumen, el presente trabajo ha establecido un marco de estudio de la textualidad web, con la descomposición del objeto de estudio, tanto por los bloques comunicativos que presenta un género determinado como por los distintos elementos textuales en los que se puede descomponer un sitio y página web. Este trabajo preliminar pretende asentar las bases para estudios posteriores que tengan un impacto claro en la localización de sitios web, puesto que se parte de una perspectiva textual y no exclusivamente técnica, enfoque este mayoritario en los estudios sobre localización. A su vez, la investigación parte de la cultura o *locale* meta, reclamando la necesidad de que los estudios partan de cada cultura y no desde el hegemónico entorno de desarrollo anglosajón hacia una imaginado destinatario global que acepta contenidos supuestamente globalizados y culturalmente neutros (Schäler, 2002).

Como conclusión, y de nuevo reclamando la importancia del texto en localización, concluimos con unas palabras del más reconocido investigador en usabilidad web:

**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).

## 11. FUTURO DE LA INVESTIGACIÓN

Son varias las posibles vías futuras de investigación una vez desarrolla una base de contraste y análisis de un género web determinado en español.

La primera opción se centra en la recopilación de un corpus comparable (Baker, 1995: 243) de sitios web localizados al español desde diferentes *locales* para investigar como afecta el proceso de traducción al producto, el texto localizado. Se estudiarían las diferencias en la convencionalización de elementos tanto extratextuales, el género, como de elementos intratextuales, la tipología textual. A partir de los diferentes bloques comunicativos identificados, se separarían subcorpus como base comparativa. El estudio no se centraría exclusivamente en la naturaleza del texto localizado (Laviosa, 1998; Baker, 1995), sino más bien en la importancia de la tipología textual y el género en el proceso de traducción en localización. El resultado del análisis descriptivo combinando ambos conceptos puede resultar en una investigación de gran utilidad al traductor (Trosborg, 1997).

De igual manera, y una vez establecidos ciertos parámetros básicos y recopilado un corpus de control, se puede realizar un estudio experimental en el que se estudiaran las peculiaridades en el proceso psicolingüístico en la traducción de hipertextos con una funcionalidad específica y la comparación con el proceso de traducción de textos lineales, estudiando como varían los mecanismos de establecimiento de las necesarias coherencia y cohesión entre ambos casos. El corpus recopilado supondría un corpus de «control» en la evaluación y preparación de los instrumentos. La proyección hacia el futuro de la profesión del traductor estará marcada por el aumento exponencial del uso de hipertextos como forma textual básica y por lo tanto, el proceso de traducción de los mismos

será de esencial importancia en la traductología, tanto para los profesionales como para la formación de futuros traductores.

Por último, y debido a este aumento exponencial en el uso y localización de hipertextos, la enseñanza de la localización de hipertextos como estructura textual básica se requerirá en la formación de traductores (Días Fouces, 2004; Gouadec, 2003). La última de las vías de investigación propuestas se centraría en el desarrollo de una serie de unidades didácticas a partir del corpus recopilado y la aplicación de un modelo de enseñanza de la localización de hipertextos en un grupo controlado de estudiantes de último curso de traducción que hayan estudiado un curso en localización. Este modelo partiría de las peculiaridades textuales observadas en los hipertextos españoles y supondría la evaluación por medio de instrumentos antes y después del desarrollo del módulo de las competencias adquiridas, para así investigar de la importancia del enfoque textual en la localización (Quirion, 2003) que parta de cada *locale* o cultura meta (Schäler, 2002). El grupo de control se centraría exclusivamente en la enseñanza de las peculiaridades técnicas del proceso de localización web, que en la actualidad suele ser la base de la formación en localización.

## 12. BIBLIOGRAFÍA

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. *Internet in linguistics, translation and literary studies*. Castellón de la Plana: Publicaciones de la Universidad Jaume I. 375-390.

Apple Computer Inc. 1992. *Macintosh Human Interface Guidelines*. Boston: Addison-Wesley Professionals.

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, S. Clear, J. y N. Olster. 1992. "Corpus Design Criteria". *Literary and Linguistic Computing*, 7 (1): 1-16.

Baker, M. 1996. "Corpus-based Translation Studies: The Challenges that Lie Ahead". En Harold 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. 1992. *In Other Words: A coursebook on translation*. Londres/Nueva York: Routledge.

Bassnett, S. y Lefevere, A. 1991. *Translation, history and culture*. London and New York: Pinter Publishers.

Barthes, R. 1977. *Image - Music - Text*. Hilla and Wang: Nueva York.

Beaugrande, R de y W. Dressler. 1981. *Introduction to Text Linguistics*. Londres: Longman.

Berners-Lee, T. 1999. *Weaving the Web*. San Francisco: Harper Collins.

Bhatia, V.K. 1993. *Analysing genre. Language use in professional settings*. Londres: Longman.

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., Conrad, S. 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 <http://jan.ucc.nau.edu/~biber/Web%20text%20types.ppt>).

Bloomfield, L. 1935. *Language*. London: George Allen and Unwin.

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. 2005. "Organización del conocimiento para la traducción jurídica a través de sistemas expertos basados en el concepto de género textual". En Isabel García Izquierdo (ed), *El género textual y la traducción. Reflexiones teóricas y aplicaciones pedagógicas*. Berna: Peter Lang.

Borja Albi, A. 2000. *El texto jurídico inglés y su traducción al español*. Barcelona: Ariel.

- Bowker, L. 1998. "Using Specialised Monolingual Native-Language Corpora as a Translation Resource: A Pilot Study". *Meta* 43 (4)
- Brinck, T., Gergle, D. y Scott D. Wood. 2002. *Usability for the Web*. San Francisco: Morgan Kaufman
- Brooks, D. 2000. "What Price Globalization? Managing Costs at Microsoft". En Robert C. Sprung (ed), *Translating into Success. Cutting-edge strategies for going multilingual in a global age*. Amsterdam-Filadelfia: John Benjamins. 42-59.
- Cheng, S. 2000. Globalizing an e-Commerce Website. En Robert C. Sprung (ed), *Translating into Success. Cutting-edge strategies for going multilingual in a global age*. Amsterdam-Filadelfia: John Benjamins. 29-41.
- Chomsky, N. 1965. *Aspects of the Theory of Syntax*. Cambridge: MIT Press.
- Corpas Pastor, G. 2001. "Compilación de un corpus ad hoc para la enseñanza de la traducción inversa especializada". *Trans*, 5: 155-184.
- 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.
- Duff, D. 2000. *Modern Genre Theory*. Harlow: Pearson.
- 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. (Disponível em <http://www.confluencias.net/n1/fouces/html>.)

Esselink, B. 2001. *A Practical Guide to Localization*. Amsterdam - Filadelfia: John Benjamins.

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., López, C. I. y Tercedor, M. 2001. "Utilización de técnicas de corpus en la representación del conocimiento médico". *Terminology*, 7 (2): 167-197

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.

Fortanet, I., Posteguillo, S., Plamer, J.C y Coll, J. F (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 Rouet, J. F., Levonen, J. J., Dillon, A. y Rand J. Spiro. *Hypertext and Cognition*. Nueva Jersey: Lawrence Erlbaum. 109-136.

Fritz, G. 1998, "Coherence in hypertext". En Bublitz, W., Lenk, U. y Ventola, E. (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. (Disponível em [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 .

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. 2002. "El género: plataforma de confluencia de nociones fundamentales en didáctica de la traducción". *Discursos. Série Estudos de Tradução*, 2. Universidade Aberta: Lisboa

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*. Tubinga: Narr.

Gommlich, K. 1993. "Text Typology and Translation-Oriented Text Analysis". En Wright, S. E. y L. D. Wright (eds), *Scientific and Technical Translation*. Amsterdam/ Filadelfia: John Benjamins.

Göpferich, S. 1995. "A Pragmatic Classification of LSP Texts In Science and Technology". *Target*, 7: 305-326.

Gouadec, D. 2003. "Le bagage spécifique du locuteur/localisateur. La vrai «nouvel profil» requis". *Meta*, 28 (4): 526-545.

Hall, E. 1976. *Beyond Culture*. Garden City, NY: Anchor Press.

- Halliday, M.A.K. y R. Hasan. 1985. *Language, Context and Text. Aspects of Language in a Social-semiotic Perspective*. Victoria: Deakin University Press.
- Halverson, S. 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)
- 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.
- Hofstede, G. 1991. *Cultures and Organizations: Software of the Mind*. Londres: McGraw-Hill.
- Hunt, Kevin, 1995. "The Design of World Wide Web Home Pages: Using Visuals to Establish Organizational Ethos". Actas de la *Society for Technical Communication (STC)*. (Disponible en <http://www.stc.org/confproceed/1995/PDFs/PG364367.PDF>)
- Hurtado Albir, A. 1996. "La traducción: classification et éléments d'analyse". *Meta* 41, (3) (Disponible en <http://www.erudit.org/revue/meta/1996/v41/n3>).
- Hurtado Albir, A. 2001. *Traducción y Traductología. Introducción a la traductología*. Madrid: Cátedra.
- Hymes, D. 1974. *Foundations in sociolinguistics: an ethnographic approach*. Filadelfia: University of Pennsylvania Press.
- Janoschka, A. 2003. *Web Advertising*. Amsterdam-Filadelfia: John Benjamins.
- Jiménez, M. A. 2005. "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.

- Johansson, S. y S. Oksefiell, eds. 1998. *Corpora and Cross-linguistic Research: Theory, Method and Case Studies*. Rodopi: Amsterdam, Atlanta.
- Jucker, A. H. 2003 "Mass Media Communication at the Beginning of the Twenty Century". *Journal of Historical Pragmatics*, 4 (1): 129-148.
- Jung, Linus, 2000. *La escuela traductológica de Leipzig*. Granada: Comares.
- 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. 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
- Kersten, G.E., Kersten, M.A. y W. Rakowski, 2001. "Application Software and Culture: Beyond the Surface of Software Interface". *InterNeg.org*. Disponible en (<http://interneg.concordia.ca/interneg/research/papers/2001/01.pdf>)
- Kilian, C. 1999. *Writing for the Web*. Self-Counsel Press: North Vancouver, Canada.
- Kress, G. 1985. *Linguistic Processes of Sociocultural Practice*. Victoria: Deakin University Press.
- Krug, S. 2005. *Don't make me think*. Indianápolis: New Riders.
- Kussmal, P. 1997. "Text-Type conventions and Translating: Some Methodological Issues". En A. Trosborg (ed.) *Text Typology and Translation*. John Benjamins: Amsterdam. 67-82.
- Lan, L. 2000. "Email: a challenge to Standard English?". *English Today*, 16 (4): 23-29.

Landow, George. 1992. *Hypertext: The convergence of contemporary Critical Theory and Technology*. Baltimore: The John Hopkins University Press.

Laviosa, S. 2002. *Corpus-based Translation Studies*. Amsterdam: Rodopi.

Laviosa, S. 1998. "The English Comparable Corpus: a Resource and a Methodology. En Bowker, L., Cronin, M., Kenny, D. 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.

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.

Leech, G. 1992. "Corpora and Theories of Linguistic Performance". En 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.

LISA. 2006. *LISA Best Practice Guide: Managing Global Content. Globan 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*. Lommet, A. (ed). Ginebra: The Localization Industry Standards Association (LISA).

LISA. 2003. *Localization Industry Primer*. Fry, D. (ed). Ginebra: The Localization Industry Standards Association (LISA). (Disponible en <http://www.lisa.org/products/primer.html>)

Lockwood, R. 2000. "Machine Translation and Controlled Authoring at Carterpillar". En Robert C. Sprung (ed) *Translating into Success. Cutting-edge strategies for going multilingual in a global age*. Amsterdam-Filadelfia: John Benjamings. 187-202.

Lockwood, R. 1999. "You snooze, you lose". *Language International* 11(4):12-16.

Lockwood, R. y K. Scott. 1999. *A Writer's Guide to the Internet*. London: Allison and Busby.

López Rodríguez, Clara Ines. 2001. *Tipologías textuales y conexión en la traducción biomédica inglés*. Tesis doctoral. Universidad de Granada. (Disponible en [http://www.ugr.es/~dpto\\_ti/prof/lrci-tesis.zip](http://www.ugr.es/~dpto_ti/prof/lrci-tesis.zip)).

López Quero, S. 2003. *El lenguaje de los «chats». Aspectos gramaticales*. Granada: Port-Royal Ediciones.

Luna, D., Peracchio, L. A. y M. D. de Juan. 2002. "Cross-Cultural and Cognitive Aspects of Web Site Navegation". *Journal of the Academy of Marketing and Science*, 30 (4): 397-410

Macizo, P. y M. T. Bajo. 2005. "Memoria de trabajo y traducción". *Cognitiva*, 17 (1): 29-54.

Macromedia, 2004. *Using Macromedia Flash MX as a multilingual authoring tool*. (Disponible en <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. *The Meaning of Meaning*. Londres: Kegan Paul.

Mata Pastor, M. 2005. "Localización y traducción de contenido web". En Reineke, D. (ed), *Traducción y Localización*. La Palmas de Gran Canaria: Anroart Ediciones. 187-252.

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 Gonzalo García, C. Y V. García Yebra. (eds), *Manual de documentación y terminología para la traducción especializada*. Madrid: Arco Libros.

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

McEnery, A. y A. Wilson. 1996. *Corpus Linguistics*. Edimburgo: Edinburgh University Press.

McGovern, G., Norton, R. y Catherine O'Dowd. 2002. *The Web Content Style Guide*. Edimburgo: Pearson Education.

Nielsen, J. 2003. "[PDF: Unfit for Human](http://www.useit.com/alertbox/20030714.html)". (Disponible en <http://www.useit.com/alertbox/20030714.html>).

Nielsen, Jakob. 2000. *Designing Web Usability: the practice of simplicity*. Indianápolis: News Riders

Nielsen, J. y M. Tahir. 2002. *Homepage usability: 50 Websited deconstructed*. Indianápolis: News Riders

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 <http://www.razonypalabra.org.mx/anteriores/n38/mneumuller.html>).

Nelson, T. 1993. (Primera edición 1981) *Literary Machines*. Sausalito: Mindful Press

Nielsen Netratings. 2003. "Global Internet Population Grows an Average of Four Percent Year-Over-Year". Comunicado de prensa del 20 de febrero del 2003. (Disponible en [http://www.nielsen-netratings.com/pr/pr\\_030220\\_hk.pdf](http://www.nielsen-netratings.com/pr/pr_030220_hk.pdf))

Nord, C. 1997. " A functional typology of translations". En A. Trosborg (ed). *Text Typology and Translation*. Amsterdam- Filadelfia: John Benjamins. 43-66.

Nord, C. 1996. *Translating as a Purposeful Activity. Functionalist Approaches Explained*. Manchester: St. Jerome.

O'Brien, S. 1998. "Practical Experience of Computer-Aided Translation Tools in the Software Localization Industry". En Bowker, L., Cronin, M., Kenny, D. 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.

Orozco, Mariana, 2002. "Revisión de investigaciones empíricas en traducción escrita". En *Trans*, vol. 6 2002, Pp 63-85

Palomares, R. (1999): "Evaluación de recursos de comunicación para el traductor en Internet". En Pinto, M. y J. A. Cordon (eds), *Técnicas documentales aplicadas a la traducción*. Madrid: Síntesis. 179-193.

Pearson, 1999. *Terms in Context*. Amsterdam-Filidelfia: John Benjamins.

Perez Hernandez, M. Chantal. 2002. *Explotación de los corpóra textuales informatizados para la creación de bases de datos terminológicas basadas en el conocimiento*. Madrid: CSIC/Elides, volumen 18 (disponible en : <http://elies.rediris.es/elies18/>).

Posteguillo, S. 2003. *Netlinguistics: languages, discourse and ideology in Internet*. Castellón: Publicaciones de la Universidad Jaume I.

Puurttinen, T. 1995. *Linguistic Acceptability in Translated Children's Literature*. Tesis Doctoral, University of Joensuu Publications in the Humanities N. 15, University of Joensuu.

Quirion, M. 2003. "La formation en localisation à l'université : pour quoi faire ?". *Meta* 48 (4):546-558 (Disponible en <http://www.erudit.org/revue/meta/2003/v48/n4/008725ar.html>)

- 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.
- 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.
- Reiss, K. 1971. *Möglichkeiten und Grenzen der Übersetzungskritik*. München: Hueber.
- Reiss, K. 1976. *Texttyp und Übersetzungsmethode. Der operative Text*. Kronberg: Scriptor.
- Reiss, K. 1981. "Type, Kind and Individuality of Text: Decision Making in Translation". *Poetics Today*, 2 (4)
- Reiss, K. y J. Vermeer. 1984. *Grundlegung einer Allgemeinen Translationsstheorie*. Tubinga: Niemeyer. (Traducción al español de S. García Reina y C. Martín de León, *Fundamentos para una teoría funcional de la traducción*. Madrid: Akal, 1997).
- 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 [www.tdx.cesca.es/TESIS\\_UJI/AVAILABLE/TDX-0126105-124531//renau.pdf](http://www.tdx.cesca.es/TESIS_UJI/AVAILABLE/TDX-0126105-124531//renau.pdf))
- Rosch, E. 1978. "Principles of categorization". En Rosch, E., 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.

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.

Sánchez Trigo, E. 2002. "Tipologías textuales y traducción". *Trans*, 6: 121-133.

Santini, M. 2005a. "Clustering Web Pages to Identify Emerging Textual Patterns". Ponencia en la Conferencia [RÉCITAL 2005](#), Dourdan, Francia. (Disponible en <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". Ponencia en el Congreso [AAACL/ICAME](#), Ann Arbor, MI. (Disponible en <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". Ponencia en el Congreso *LISA Forum Cairo 2005: Perspectives from the Middle East and Africa*. (Disponible en <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 [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". Reineke, D. (ed), *Traducción y Localización*. La Palmas de Gran Canaria: Anroart Ediciones.

Scott, Mike, 1996. *WordSmith Tools (Program)*. Oxford: University Press Baker.

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.

Singh, N. y A. Pereira. 2005. *The culturally customized Web site: customizing web sites for the global marketplace*. Oxford: Elsevier.

Singh, N., Furrer, O. 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.

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 Verlang. 33-65.

Stubbs, M. 1996. *Text and Corpus Analysis: Computer-assisted studies of language and culture*. Londres: Blackwell Publishers.

Swales, J. M 1990. *Genre Analysis. English in Academic and Research Settings*. Cambridge: Cambridge University Press

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 <http://elies.rediris.es/elies6/>)

- Torruella, J. y J. Llisteri. 1999. "Diseño de corpus textuales y orales". En J. M. Blecua *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. 1980. *In Search of a Theory of Translation*. Tel Aviv: The Porter Institute for Poetics and Semiotics.
- Trosborg, A. 1997. "Text Typology: Register, Genre and Text Type". En A. Trosborg (ed), *Text Typology and Translation*. Amsterdam-Filadelfia: John Benjamins. 3-23.
- 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.
- Vilamor, J. R. 2001. *Cómo escribir en Internet*. Madrid: Editorial Universitat.
- Yates, J. y Orlikowski, W. J. 1992. "Genres of organizational communication: A structural approach to studying communications and media". *Academy of Management Review*, 17 (2): 299-326.
- Yates, J. y Orlikowski, W. J. 1998. "Genre systems: structuring interaction through communicative norms". Cambridge, MA: MIT. Disponible en <http://ccs.mit.edu/papers/CCSWP205/>
- Yunker, J. 2003. *Beyond Borders: Web Globalization Strategies*. Indianápolis, Indiana: New Riders.
- W3C Consortium. Disponible en [<http://www.w3c.es/>]
- Werlich, E. 1976. *A Text Grammar of English*. Heidelberg: Quelle and Meyer.
- Werlich, E. 1975. *Typologie der Texte*. Heidelberg: Quelle & Meyer

Zanettin, F. 1998. "Bilingual Comparable Corpora and the Training of Translators". *Meta* 43 (4)

Zanettin, F. 2000. "Parallel Corpora in Translation Studies: Issues in Corpus Design and Analysis". En Maeve Olohan (ed.) *Intercultural Faultlines. Research Models in Translation Studies I: Textual and Cognitive Aspects*. Manchester: St. Jerome, 105-118.

## 13. APÉNCIDE I. COMPOSICIÓN DE 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 en azul se seleccionaron previamente 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)
      - 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)
    - 2.1. [Aceites y vinagres](#) (35) [Laboratorios Ordesa](http://www.ordesa.es/) - http://www.ordesa.es/  
[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/) - http://www.risco-sa.com/
- 2.14. [Mayoristas y distribuidores](#) (13) [Grupo Dia](http://www.dia.es/) - http://www.dia.es/
- 2.15. [Pan y pasteles](#) (24) [Croexsa](http://www.croexsa.com) - http://www.croexsa.com
- 2.16. [Pescados y mariscos](#) (30) [Conservas Valcarcel](http://www.valcarcel-sa.es/) - http://www.valcarcel-sa.es/
- 2.17. [Productos lácteos](#) (16) [Leche Pascual](http://www.lechepascual.es/) - http://www.lechepascual.es/
  
- 3. [Artes gráficas](#) (179)
  - 3.1. [Edición](#) (67)
    - 3.1.1. [Editoriales](#) (57) [Mapas murales](http://www.mapasmurales.com/) - http://www.mapasmurales.com/
    - [Masson](http://www.masson.es/) - http://www.masson.es/
    - [Grupo Anaya](http://www.anaya.es/) - http://www.anaya.es/
  - 3.2. [Impresión](#) (45)
    - 3.2.1. [Imprentas](#) (4) [Sticard](http://www.sticard.com/) - http://www.sticard.com/
    - [Mawijo](http://www.mawijo.com/) - http://www.mawijo.com/
  - 3.3. [Maquinaria y equipo](#) (10) [Empress Servicios](http://www.franquicopy.com/) - http://www.franquicopy.com/
  - 3.4. [Materiales y suministros](#) (6) [Milan-Factis](http://www.areamilan.com/) - http://www.areamilan.com/
  
- 4. [Artes y entretenimiento](#) (265)
  - 4.1. [Fotografía](#) (29) [Aymerich Golf Management](http://www.aymerichgolf.com) - http://www.aymerichgolf.com
  - 4.2. [Música](#) (45) [Fotolab 30](http://www.fotolab30.com/) - http://www.fotolab30.com/
  - 4.2.1. [Producción discográfica](#) (38) [Ars Antiqua](http://www.ars-antiqua.com/) - http://www.ars-antiqua.com/
  - 4.3. [Producción audiovisual](#) (108) [Acuarela Discos](http://www.acuareladiscos.com) - http://www.acuareladiscos.com
  - 4.3.1. [Madrid](#) (13) [Especialistas de cine - Xtrem Solutions](http://www.especialistasdecine.com) - http://www.xtremolutions.com
  - 4.3.1.1. [Madrid](#) (13) [Yllana Producciones](http://www.yllana.com/) - http://www.yllana.com/
  
- 5. [Comercio internacional](#) (35)
  - [Consultorías](http://www.consultorias.com) (13)
  - [Ocre-bl](http://www.ocre-bl.com/) - http://www.ocre-bl.com/
  - 5.1. [Transporte y logística](#) (3) [Europamar](http://www.europamar.com/) - http://www.europamar.com/
  
- 6. [Construcción](#) (644)
  - 6.1. [Constructoras](#) (49) [Hansa Urbana](http://www.hansaurbana.com/) - http://www.hansaurbana.com/
  - [Grupo Masa](http://www.grupomasa.com) - http://www.grupomasa.com

- 6.2. [Consultores](#) (4) [ERS - Geotecnia](http://www.ersgeotecnia.com/) - <http://www.ersgeotecnia.com/>
- 6.3. [Contratistas](#) (56) [Enerpal](http://www.enerpal.com/) - <http://www.enerpal.com/>
- 6.4. [Diseño](#) (64) [Teledetección y Gis](http://www.geosys.es/) - <http://www.geosys.es/>
- 6.4.1. [Arquitectura](#) (19) [Campos y Pradel](http://www.campospradel.com) - <http://www.campospradel.com>
- 6.4.2. [Interiores](#) [Almagra y Siena](http://www.almagraysiena.com/) - <http://www.almagraysiena.com/>
- 6.5. [Ingeniería Ambiental](#) (4) [Aisnorsa](http://www.aisnorsa.com/) - <http://www.aisnorsa.com/>
- 6.6. [Mantenimiento](#) (5) [3i Multihogar](http://www.3i-multihogar.com/) - <http://www.3i-multihogar.com/>
- 6.7. [Maquinaria y equipo](#) (51) [Recamasa](http://www.recamasa.com/) - <http://www.recamasa.com/>
- 6.8. [Materiales y suministros](#) (319) [Alkor](http://www.alkor.es/) - <http://www.alkor.es/>
- 6.8.1. [Ascensores](#) (6) [Ascensores Imem](http://www.imem.com/) - <http://www.imem.com/>
- 6.8.2. [Climatización](#) (26) [Aplicaciones Naturales del Sol, Ansol](http://www.ansol.es/) - <http://www.ansol.es/>
- 6.8.2.1. [Aire acondicionado](#) (11) [Ventilación industrial](http://www.ventilacion.com/) - <http://www.ventilacion.com/>
- 6.8.3. [Electricidad](#) (3) [Jung Electro Ibérica](http://www.jungiberica.es/) - <http://www.jungiberica.es/>
- 6.8.4. [Fontanería](#) (4) [Saniber](http://www.saniber.com/) - <http://www.saniber.com/>
- 6.8.5. [Hormigón](#) (11) [Sika](http://www.sika.es/) - <http://www.sika.es/>
- 6.8.6. [Mayoristas y distribuidores](#) (5) [Cemento Natural](http://www.cementonatural.com/) - <http://www.cementonatural.com/>
- 6.8.7. [Metales](#) (11) [Europerfil](http://www.euoperfil.es/) - <http://www.euoperfil.es/>
- 6.8.8. [Mobiliario](#) (3) [Duscholux](http://www.duscholux.es/) - <http://www.duscholux.es/>
- 6.8.8.1. [Cocina y baño](#) (12) [Burgalia](http://www.buraglia.com/) - <http://www.buraglia.com/>
- 6.8.8.2. [Oficina](#) (15) [Mobiofi - Muebles de Oficina](http://www.mobiofi.com/) - <http://www.mobiofi.com/>
- 6.8.9. [Pinturas y recubrimientos](#) (8) [MK2](http://www.mk2.es/) - <http://www.mk2.es/>
- 6.8.10. [Puertas y ventanas](#) (19) [Decortina](http://www.decortina.com/) - <http://www.decortina.com/>
- 6.9. [Piscinas](#) (10) [Acuarama](http://www.aquarama.org/) - <http://www.aquarama.org/>
- 6.9.1. [Maquinaria y equipo](#) (8) [La Hita](http://www.lahitaalquiler.com/) - <http://www.lahitaalquiler.com/>
- 6.9.2. [Materiales y suministros](#) (7) [Textura Decoración S.L.](http://www.instaladoresonline.com/) - <http://www.instaladoresonline.com>
- 6.9.3. [Mercado Inmobiliario](#) (59) [Casas Roda Golf](http://www.rodagolf.com) - <http://www.rodagolf.com>
- 6.10. [Plantas industriales](#) (6) [Ingeniería y Construcción de Naves, Iconsa](http://www.iconsa.es/) - <http://www.iconsa.es/>
- 6.11. [Promotoras](#) (9) [Restaura](http://www.restaura.es/) - <http://www.restaura.es/>
- 6.12. [Urbanización](#) [Nuevo Mundo](http://www.promociones-inmobiliarias-promotoras.com/) - <http://www.promociones-inmobiliarias-promotoras.com/>
- 6.12.1. [Constructoras](#) (49) [Construcciones especiales y dragados](http://www.drace.com/) - <http://www.drace.com/>
- 6.12.2. [Consultores](#) (4) [Corman Sistemas](http://www.cormansis.com/) - <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/) - http://www.belnature.com/
- 12.2. [Automotriz](#) (53) [Seat España](http://www.seat.es) - http://www.seat.es
- 12.3. [Calzado y complementos](#) (50) [PabloSky](http://www.pablosky.com/) - http://www.pablosky.com/
- 12.4. [Eléctrica y electrónica](#) (136) [Instrumentos de Medida](http://www.idm-instrumentos.es/) - http://www.idm-instrumentos.es/
  - 12.4.1. [Automatismos y PLC](#) (7) [ASING](http://www.asing.es/) - http://www.asing.es/
  - 12.4.2. [Cuadros y electricidad](#) (3) [Unión Fenosa](http://www.unionfenosa.es/) - http://www.unionfenosa.es/
  - 12.4.3. [Energía e industrial](#) (4) [Endesa Oficina On-Line](http://www.endesaonline.com) - http://www.endesaonline.com
  - 12.4.4. [Equipos e instrumentación](#) (1) [Mp Electronics](http://www.mpelectronics.info/) - http://www.mpelectronics.info/
  - 12.4.5. [Iluminación](#) (15) [Salvi](http://www.salvi.es) - http://www.salvi.es
  - 12.4.6. [Ingeniería y diseño](#) (2) [Diselec](http://www.diselec.com/) - http://www.diselec.com/
  - 12.4.7. [Materiales y suministros](#) (5) [Cebec](http://www.cebek.com/) - http://www.cebek.com/
  - 12.4.8. [Mayoristas y distribuidores](#) (8) [HT Instruments](http://www.htinstruments.com/) - http://www.htinstruments.com/
  - 12.4.9. [Productos de consumo](#) (7) [Bosch España](http://www.bosch-ed.com/) - http://www.bosch-ed.com/
  - 12.4.10. [Sonido](#) (3) [Reflexion Arts](http://www.reflexion-arts.com) - http://www.reflexion-arts.com
- 12.5. [Empaques y embalajes](#) (60) [Saint Gobain La Granja](http://www.saintgobainlagranja.es/) - http://www.saintgobainlagranja.es/
- 12.6. [Energía](#) (29) [Iberdrola](http://www.iberdrola.es/) - http://www.iberdrola.es/
- 12.7. [Extractivas](#) (8) [Tino](http://www.tino.es) - http://www.tino.es
- 12.8. [Maquinaria y equipo](#) (177) [Korta S.A.](http://www.korta-sa.com/) - http://www.korta-sa.com/
  - 12.8.1. [Agrícola](#) (45) [Arteriego](http://www.arteriego.com/) - http://www.arteriego.com/
  - 12.8.2. [Comercial](#) (9) [Freecaad](http://www.freecaad.com/) - http://www.freecaad.com/
  - 12.8.3. [Laboratorios](#) (3) [Flores Valles](http://www.floresvalles.es/) - http://www.floresvalles.es/
    - 12.8.3.1. [Textil](#) (12) [Jose Fuster](http://www.fuster.com) - http://www.fuster.com
  - 12.8.4. [Materiales](#) (115) [Codeca](http://www.cauchocodeca.com/) - http://www.cauchocodeca.com/
    - 12.8.4.1. [Madera](#) (26) [Maderas Manuel Sanz](http://www.bemarnet.es/manuelsanz/) - http://www.bemarnet.es/manuelsanz/
    - 12.8.4.2. [Metales](#) (20) [Fundesa](http://www.fundesa.es/) - http://www.fundesa.es/
    - 12.8.4.3. [Papel y cartón](#) (46) [Fabregas](http://www.fabregas.com/) - http://www.fabregas.com/
      - 12.8.4.3.1. [Pasta, papel y cartón](#) (38)
        - 12.8.4.3.1.1. [Distribuidores](http://www.keapapel.com/empresa.htm) (4) [Kea](http://www.keapapel.com/empresa.htm) [Papel](#) - http://www.keapapel.com/empresa.htm
        - 12.8.4.3.1.2. [Fabricantes](#) (11) [ENCE: Empresa Nacional de Celulosas](http://www.ence.es/) - http://www.ence.es/
        - 12.8.4.3.1.3. [Maquinaria](#) (3) [Maquinaria Ceres](http://www.macesa.com/) - http://www.macesa.com/
        - 12.8.4.3.1.4. [Transformadores](#) (14) [Cartonajes Levante](http://www.cartonajeslevante.com) - http://www.cartonajeslevante.com
      - 12.8.4.4. [Plásticos](#) (42) [Grup Sigma](http://www.grupsigma.es/) - http://www.grupsigma.es/

- 12.8.4.5. [Vidrio](#) (16) [Vidrieras Victor](http://www.vidrieras-victor.com/) - <http://www.vidrieras-victor.com/>
- 12.8.5. [Muebles](#) (128) [Colchones Relax](http://www.relax.es) - <http://www.relax.es>
- 12.8.6. [Química](#) (186) [Cepsa](http://www.cepsa.es) - <http://www.cepsa.es>
- 12.8.6.1. [Farmacéutica](#) (21) [Almirall Prodesfarma](http://www.almirallprodesfarma.com/) - <http://www.almirallprodesfarma.com/>
- 12.8.7. [Siderurgia](#) (36) [Olarra](http://www.olarra.com) - <http://www.olarra.com>
- 12.8.8. [Suministros](#) (66) [Carbueros Metálicos](http://www.carbueros.com/) - <http://www.carbueros.com/>
- 12.8.9. [Óptica](#) (1) [Grupo Prats](http://www.grupoprats.com) - <http://www.grupoprats.com>
  
- 13. [Informática](#) (943)
  - 13.1.1. [Consultores](#) (49) [Indra Sistemas](http://www.indra.es/) - <http://www.indra.es/>
    - 13.1.1.1. [Internet](#) (25) [Davinci Consulting Tecnológico](http://www.dvc.es/) - <http://www.dvc.es/>
      - 13.1.1.1.1. [Internet](#) (25) [F3](http://www.f-3.net) - <http://www.f-3.net>
      - 13.1.1.1.2. [Usabilidad](#) (21) [Xperience Consulting](http://www.xperienceconsulting.com) - <http://www.xperienceconsulting.com>
    - 13.1.1.2. [Formación](#) (23) [Cursos Aulacrea](http://www.aulacrea.com/) - <http://www.aulacrea.com/>
    - 13.1.1.3. [Hardware](#) (67) [Microrent](http://www.microrent.es) - <http://www.microrent.es>
      - 13.1.1.3.1. [Fabricantes](#) (18) [Inves](http://www.inves.es/) - <http://www.inves.es/>
      - 13.1.1.3.2. [Periféricos](#) (14) [Laboratorios Electrónicos Reset](http://www.todoimpresoras.com/) - <http://www.todoimpresoras.com/>
  - 13.1.1.4. [Internet](#) (404) [Acceso a la Red Internet Española, ARI](http://www.ari.es/) - <http://www.ari.es/>
    - 13.1.1.4.1. [Comercio electrónico](#) (10) [YLos](http://www.ylos.com) - <http://www.ylos.com>
    - 13.1.1.4.2. [Hospedaje de sitios](#) (209) [Giralweb](http://www.giralweb.com) - <http://www.giralweb.com>
      - 13.1.1.4.2.1. [Promoción de sitios](#) (11) [Discovery Net](http://www.discoverynet.es/) - <http://www.discoverynet.es/>
      - 13.1.1.4.2.2. [Proveedores de acceso](#) (52) [TeleCable](http://www.telecable.es) - <http://www.telecable.es>
      - 13.1.1.4.2.3. [Registro de dominios](#) (15) [Namelia](http://www.namelia.net/) - <http://www.namelia.net/>
      - 13.1.1.4.2.4. [WAP](#) [Terra Mobile](http://www.terra-mobile.es/) - <http://www.terra-mobile.es/>
  - 13.1.1.5. [Mayoristas y distribuidores](#) (33) [Diode España](http://www.diode.es) - <http://www.diode.es>
  - 13.1.1.6. [Seguridad](#) (13) [Internet Security Auditors](http://www.isecauditors.com/) - <http://www.isecauditors.com/>
  - 13.1.1.7. [Software](#) (207) [Panda Software](http://www.pandasoftware.es/) - <http://www.pandasoftware.es/>
    - 13.1.1.7.1. [ERP](#) (34) [Aitana](http://www.aitanasbs.com) - <http://www.aitanasbs.com>
  - 13.1.1.8. [Tiendas de informática](#) [Tiendapc](http://www.tiendapc.com) - <http://www.tiendapc.com>
  
- 14. [Marketing](#) (205)
  - 14.1.1. [Consultores](#) (14) [Sinergia: Imagen, marketing y comunicación](http://www.sinergia.es/) - <http://www.sinergia.es/>
  - 14.1.2. [Exposiciones y ferias](#) (25) [Altamira Congresos](http://www.altamira-congresos.com) - <http://www.altamira-congresos.com>
    - 14.1.2.1. [Traductores](#) [Ricard Sierra](http://www.ricardsierra.com/) - <http://www.ricardsierra.com/>

- 14.1.3. [Investigación de mercados](#) (16) [Netquest](http://www.netquest.es/) - <http://www.netquest.es/>
- 14.1.4. [Marketing en Internet](#) (11) [Interactive Network](http://www.antevenio.com/) - [www.antevenio.com/](http://www.antevenio.com/)
- 14.1.5. [Marketing promocional](#) (34) [Publiestilo](http://www.publiestilo.com) - <http://www.publiestilo.com>
- 14.1.6. [Publicidad](#) (76) [Margen Design](http://www.margin.com) - <http://www.margin.com>
  - 14.1.6.1. [Agencias](#) (17) [Tasmanet](http://www.tasma.net/) - <http://www.tasma.net/>
  - 14.1.6.2. [Medios](#) (14) [Tecnodimensión](http://www.tecnodimension.com/) - <http://www.tecnodimension.com/>
- 14.1.7. [Relaciones públicas](#) (13) [Atento Telecomunicaciones España](http://www.atento.es) - <http://www.atento.es>
- 14.1.8. [Servicios de apoyo](#) (10) [Buzoneofacil](http://www.buzoneofacil.com) - <http://www.buzoneofacil.com>
- 14.1.9. [Servicios globales](#) (7) [BGS](http://www.bgssl.com) - <http://www.bgssl.com>
  
- 15. [Mayoristas y distribuidores](#) (25) [Ortus Fitness](http://www.ortus.com/) - <http://www.ortus.com/>
  - 15.1.1. [Suministros](#) [Santiago Vargas](http://www.santiagovargas.com) - <http://www.santiagovargas.com>
  
- 16. [Medio ambiente](#) (109) [Integración y Sistemas de Energías Renovables, ISER](http://www.iser.es/) - <http://www.iser.es/>
  - 16.1. [Consultores](#) (17) [Ingenieros Asesores](http://www.ingenierosasesores.com/) - <http://www.ingenierosasesores.com/>
  
- 17. [Mercado inmobiliario](#) (221) [Agrofinca](http://www.agrofinca.com/) - <http://www.agrofinca.com/>
  - 17.1. [Administradores de fincas](#) (9) [Administrador de fincas en Valencia](http://www.joseroig.com) - <http://www.joseroig.com>
  - 17.2. [Agencias inmobiliarias](#) (44) [Remax España](http://www.remax.es/) - <http://www.remax.es/>
  - 17.3. [Alquiler](#) (6) [CEVASA](http://www.cevasa.com/) - <http://www.cevasa.com/>
  - 17.4. [Consultorías](#) (12) [Aguirre Newman](http://www.aguirrenewman.es/) - <http://www.aguirrenewman.es/>
  - 17.5. [Tasadoras](#) (25) [Tecnitasa](http://www.tecnitasa.es/) - <http://www.tecnitasa.es/>
  
- 18. [Oportunidades](#) (21)
  - 18.1. [Cursos en línea](#) (5) [Nomaster](http://www.nomaster.net/) - <http://www.nomaster.net/>
  - 18.2. [Franquicias](#) (15) [Tormo](http://www.tormo.com/) - <http://www.tormo.com/>
  
- 19. [Productos de consumo](#) (269)
  - 19.1.1. [Deportivos](#) (5) [Conor Sports](http://www.conor.es/) - <http://www.conor.es/>
  - 19.1.2. [Oficina](#) (5) [Sie](http://www.siesl.com) - <http://www.siesl.com>
  - 19.1.3. [Ropa](#) (73) [Etxe-Ondo](http://www.etxeondo.com/) - <http://www.etxeondo.com/>
  
- 20. [Recursos humanos](#) (21) [Verja y Asociados](http://www.verdejayasociados.com/) - <http://www.verdejayasociados.com/>
  - 20.1. [Software](#) (42) [Seresco](http://www.seresco.es) - <http://www.seresco.es>

21. [Salud](#) (51) [Hiperion Biotech](#) - <http://www.hiperionbiotech.com/>  
21.1. [Discapitados](#) (47) [Intervención Social, IS](#) - <http://www.intervencionsocial.com/>
22. [Seguridad](#) (141) [Securitas](#) - <http://www.securitas.es/>  
22.1. [Contra incendios](#) (9) [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) [Tracasa](#) - <http://www.tracasa.es>  
23.1. [Cerrajería](#) (5) [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) [Organizacion de eventos y fiestas](#) - <http://www.jptres.com/>  
23.4.1. [Fiestas y celebraciones](#) (8) [Iluminaciones Truman](#) - <http://www.codsc.com/iluminacionestruman/>  
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/>  
23.7.2. [Maquinaria y equipo](#) (29) [Alberto Casañal](#) - <http://www.maquinarialimpieza.com>  
23.7.3. [Tintorerías y lavanderías](#) (5) [Lavapant](#) - <http://www.lavapant.com/spanish/index.asp>  
23.7.4. [Viaria](#) [Prodesa Medioambiente](#) - <http://www.prodesa.net/>  
23.8. [Mensajería](#) (12) [STU Express Service](#) - <http://www.stu.es/>  
23.9. [Mudanzas](#) (47) [Cajas para mudanzas](#) - <http://www.cajasdemudanzas.com/>

- 23.10. [Veterinarios](#) (48) [Analítica Veterinaria - http://www.analiticaveterinaria.com/](http://www.analiticaveterinaria.com/)
24. [Telecomunicaciones](#) (160) [Unitronics](#) [Comunicaciones](#) -  
<http://www.comunicaciones.unitronics.es/>
- 24.1. [Consultores](#) (24) [Satec](#) - <http://www.satec.es/>
- 24.1.1. [Telefonía](#) (3) [Ypertron](#) - <http://www.ypertron.com>
- 24.2. [Fabricantes](#) (7) [Tecatel](#) - <http://www.tecatel.com/>
- 24.3. [Mayoristas y distribuidores](#) (6) [Jovitel Telecomunicaciones](#) - <http://www.jovitel.com/>
- 24.4. [Operadores](#) (13) [Jazztel Telecomunicaciones](#) - <http://www.jazztel.com/>
- 24.4.1. [ADSL](#) (13) [ADSL Ya](#) - <http://acceso.ya.com/adsl/>
- 24.4.2. [Cable](#) (4) [Ono](#) - <http://www.ono.es/>
- 24.4.3. [Telefonía](#) (8) [Retevisión](#) - <http://www.retevision.es/>
- 24.4.3.1. [Móvil](#) (26) [Mobipay](#) - <http://www.mobipay.es>
- 24.4.4. [Fabricantes](#) (5) [Nokia](#) - <http://www.nokia.es/>
- 24.4.5. [Móvil](#) (20) [Movistar](#) - <http://www.movistar.es/>
- 24.4.6. [Operadores](#) (8) [Telefónica](#) - <http://www.telefonica.es/>
25. [Textiles](#) (172) [Alfombras](#) [y](#) [Moquetas](#) [Belka](#) -  
<http://www.alfombrasbelka.com/>
- 25.1. [Materiales y suministros](#) (9) [Ideatint](#) - <http://www.ideatint.com/>
- 25.2. [Mayoristas y distribuidores](#) (6) [Carmen Rey](#) - <http://www.carmenrey.com/>
26. [Transportes y logística](#) (292) [Alcoarte](#) - <http://www.alcoarte.com/>
- 26.1. [Alquiler de barcos](#) (16) [Cosas de Barcos](#) - <http://www.cosasdebarcos.com/>
- 26.2. [Aéreos](#) (18) [Iberia](#) - <http://www.iberia.es>
- 26.3. [Consultores](#) (2) [Geswinner](#) - <http://www.geswinner.com/>
- 26.4. [Ferrovianos](#) (7) [RENFE](#) - <http://www.renfe.es/>
- 26.5. [Marítimos y fluviales](#) (19) [Trasmediterránea](#) - <http://www.trasmediterranea.es/>
- 26.6. [Por carretera](#) (40) [Europcar](#) - <http://www.europcar.es/>
27. [Turismo](#) (86)
- 27.1. [Agencias de viajes](#) (62) [Viajar](#) - <http://www.viajar.com/>
- 27.2. [Alquiler de vehículos](#) (32) [Victoria Rent a Car](#) - <http://www.victoriacars.com/>

## 14. APÉNDICE II. FIGURAS Y TABLAS

### 14.1. FIGURAS

FIGURA 1. ESQUEMA DEL CICLO DE DESARROLLO DE UN PRODUCTO GLOBAL SEGÚN LISA (2004: 15) ...	10
FIGURA 2. DIFERENTES GRADOS DE LOCALIZACIÓN EN IKEA.COM.....	20
FIGURA 3. EVOLUCIÓN DEL USO DE INTERNET EN ESPAÑA. FUENTE: ASOCIACIÓN DE USUARIOS DE INTERNET.....	27
FIGURA 4. PREDICCIÓN EN EL VOLUMEN DE NEGOCIO EN 2007. FUENTE IDC RESEARCH (CF. SCHÄLER, 2005).....	31
FIGURA 5. DISPOSICIÓN LÍQUIDA EN PANTALLA CON POSIBILIDAD DE VARIAR EL TAMAÑO DE LA LETRA. ICONO DE WWW.ONO.ES. ....	36
FIGURA 6. PRODUCTO DE SOFTWARE COMPLETO. ....	36
FIGURA 7. PROPUESTA DE SITUACIACIÓN DEL GÉNERO OBJETO DE ESTUDIO. ....	45
FIGURA 8. MODELO DE COMUNICACIÓN DE MASAS INTERACTIVO. ADAPTADO DE JANOSCHKA (2003) .	53
FIGURA 9. RESUMEN DE CONTENIDO FORMA Y FUNCIONALIDAD DE LOS 6 CIBERGÉNEROS DESCRITOS POR SHEPHERD Y WATTERS (1999) .....	56
FIGURA 10. LOS CIBERGÉNEROS. ADAPTADO DE SHEPHERD Y WATTERS, (1998), Y UBICACIÓN DEL OBJETIVO DE ESTUDIO, COMBINADO CON KENNEDY Y SHEPHERD, (2005).....	57
FIGURA 11. CLASIFICACIÓN DE GÉNEROS SEGÚNA LA NIELSEN-NETRATINS.....	60
FIGURA 12. MODELO DE TEXTO LINEAL. ADAPTADO DE FRITZ (1997: 222).....	66
FIGURA 13. MODELO DE LA ESTRUCTURA DE UN HIPERTEXTO EN EL CONTEXTO DE LA WWW. ADAPTADO DE JANOSCHKA (2003: 173) .....	67
FIGURA 14. ESQUEMA DE GLOBAL CONTENT MANGAMENT SYSTEM. ....	71
FIGURA 15. ESQUEMA DE LAS TRES DIMENSIONES DEL CONTEXTO ( HATIM Y MASON 1990: 58).....	76
FIGURA 16. FUENTE, HATIM Y MASON (1990:75). PRAGMÁTICA Y SEMIÓTICA DEL CONTEXTO. ....	79
FIGURA 17. PANTALLA ANIMADA SPLASH EN FORMATO MACROMEDIA FLASH. USADA ADEMÁS COMO SELECTOR DE IDIOMA. ....	105
FIGURA 18. USO DE ICONOS ANGLOSAJONES EN PÁGINA WEB DE INICIO ESPAÑOLA. WWW.CADITECTESTING.ES .....	106

FIGURA 19. TÍTULO DE LA PÁGINA WEB DE WWW.INNOVA31.ES, CON INDICACIÓN DEL TÍTULO GENERAL DEL SITIO WEB Y DE LA PÁGINA ESPECÍFICA DEL MISMO.....	108
FIGURA 20. MENÚ DE NAVEGACIÓN SUPERIOR. WWW.WWW.MAHOU-SANMIGUEL.COM\INICIO.HTML	109
FIGURA 21. PIE DE PÁGINA.WWW.AZUD.COM\ES\INDEX.HTML .....	109
FIGURA 22. SEGMENTO TEXTUAL <i>ONMouse</i> AL PASAR EL RATÓN POR EL ÁREA DEL GRÁFICO. ....	111
FIGURA 23. PÁGINA PROTEGIDA QUE INDICA UN ÁREA RESERVADA DE WWW.DIODE.ES .....	111
FIGURA 24. SEGMENTOS TEXTUALES EN EVENTO « <i>SELECT</i> ».....	112
FIGURA 25. DISTRIBUCIÓN TEXTUAL EN UNA PÁGINA WEB. ELABORACIÓN PROPIA. ....	113
FIGURA 26. MENÚ DE NAVEGACIÓN CON TEXTO CODIFICADO EN IMÁGENES. DE WWW.PANDASOFTWARE.ES\HOME.HTML.....	114
FIGURA 27. BLOQUE «CONTACTO» INCLUIDO EN LA PÁGINA DE INICIO. ....	115
FIGURA 28. FORMULARIO DE CONTACTO CON UNA FUNCIÓN TEXTUAL EXHORTATIVA. DE WWW.ADEA.ES.....	121
FIGURA 29. PRESENCIA DE FORMAS SUSTANTIVADAS FRENTE A FORMAS VERBALES EN «CONTACTO»..	123
FIGURA 30. EL TENOR EN LOS MENÚS DE NAVEGACIÓN. EXTRACTO DE BLOQUE «CONTACTO».....	123

## 14.2. TABLAS

TABLA 1. FRECUENCIA DE USO EN CORPUS ORIGINAL Y LOCALIZADO. DE JIMÉNEZ (2005) .....	29
TABLA 2. CICLOS DE VIDA DE PRODUCTOS DE SOFTWARE <i>versus</i> CONTENIDO WEB (MATA, 2005: 194). 37	
TABLA 3. ESTADÍSTICAS DEL CORPUS REALIZADA CON EL PROGRAMA WEBBUDGET.....	102
TABLA 4. DISTRIBUCIÓN DE VERSIONES LOCALIZADAS DE SITIOS WEB ESPAÑOLES. ....	103
TABLA 5. GRADO DE PROTOTIPICIDAD DE BLOQUES COMUNICATIVOS EN EL GÉNERO WEB. ....	117
TABLA 6. BLOQUES Y SECCIONES IDENTIFICADOS EN EL CORPUS.....	119
TABLA 7. PORCENTAJE DE USO DE TÉRMINOS PARA EL BLOQUE «CONTACTO».....	122
TABLA 8. TÉRMINOS EMPLEADOS EN LA DENOMINACIÓN DEL BLOQUE «EMPRESA».....	125
TABLA 9. SECCIONES EN BLOQUE COMUNICATIVO “EMPRESA” .....	126
TABLA 10. TÉRMINOS PARA LA DENOMINACIÓN DEL BLOQUE “NOTICIAS” . ....	130
TABLA 11. FRECUENCIA TÉRMINOS EN LA DENOMINACIÓN DE BLOQUE «PRODUCTOS».....	132
TABLA 12. FRECUENCIA TÉRMINOS EN LA DENOMINACIÓN DE BLOQUE “SERVICIOS” . ....	133
TABLA 13. FRECUENCIA DE TÉRMINOS EN LA DENOMINACIÓN DE LA SECCIÓN “OFERTAS” .....	133
TABLA 14. PORCENTAJE DE DISTRIBUCIÓN DEL BLOQUE «LEGAL». ....	135

TABLA 15. PORCENTAJE DE APARICIÓN EN EL BLOQUE «AVISO LEGAL» Y TERMINOLOGÍA EMPLEADA. .....	136
TABLA 16. FRECUENCIA DE TÉRMINOS EN LA DENOMINACIÓN DEL BLOQUE «AVISO LEGAL».....	137
TABLA 17. GRADO DE PROTOTIPICIDAD DE LAS SECCIONES EN BLOQUE «ZONA RESERVADA».....	138
TABLA 18. EL TENOR DIGITAL. TRATAMIENTO EN LA SECCIÓN «TRABAJO».....	139
TABLA 19 . TENOR DIGITAL EN LA SECCIÓN «TRABAJO». SUSTATIVO O FORMA VERBAL.....	140
TABLA 20. GRADO DE PROTOTIPICIDAD DE LAS SECCIONES EN BLOQUE «INTERACTIVIDAD».....	141
TABLA 21. EL TENOR DIGITAL EN LA SECCIÓN «REGISTRO».....	142