

ASPECTOS TÉCNICOS Y METODOLÓGICOS DE LA LOCALIZACIÓN DE CONTENIDO WEB¹

María José Pérez Cabrera
Email: mariajoseperezcabrera@gmail.com

Resumen: Este artículo incide sobre aspectos teórico-prácticos que permiten realizar una categorización de los conocimientos necesarios para llevar a cabo un proyecto de localización de contenido web.

Se ha pretendido explicar detalladamente el proceso que debe seguir un traductor al recibir un encargo de este tipo. Se ha realizado un análisis de las herramientas disponibles para trabajos de localización basándonos en el estudio de Esselink (2000). Posteriormente, se han evaluado los recursos disponibles y se han extraído conclusiones sobre los que pueden resultar más idóneos para el traductor.

Una vez elegidas las herramientas, se detalla su uso mediante una rigurosa explicación e imágenes para que cualquier profesional pueda aprender a utilizarlas sin dificultad.

Palabras clave: localización, contenido web, herramientas de localización.

Título en inglés: Web content localization: technical and methodological aspects

Abstract: This article deals with the theoretical and practical features that lead to a categorization of aspects related to the web content localization.

This is an attempt to explain the process a translator must follow when receiving an assignment of this kind. We have developed an analysis of the existing tools for this task based on the studies carried out by Esselink (2000). Furthermore, these resources have been assessed in order to find the most suitable ones for the translator.

Once the implements are chosen, we have explained their use in detail with the support of images so that any professional can learn to use them easily.

Keywords: localization, web format, localization tools.

1. INTRODUCCIÓN

En una época en la que el contenido web es el método principal de difusión de información, el traductor se enfrenta a una nueva forma de trabajo en la que son muchas las ocasiones en las que se le requiere la traducción de textos enmarcados en un soporte relativamente actual en el mundo de la traducción profesional para el que los recursos, aunque cada vez son más, aún son más escasos de lo deseable y su uso sigue siendo dificultoso para muchos profesionales.

¹ **Date of reception:** 29 May 2012

Date of acceptance: 6 September 2012

Con este artículo pretendemos acercarnos a uno de los tipos de encargos de traducción que podemos decir más frecuentes en este ámbito: la traducción de un texto sobre cualquier temática, general o especializada, presentado en un soporte web.

En un caso así, es necesario cubrir todos los aspectos imprescindibles para obtener un resultado adecuado, es decir, hay que conocer en profundidad la temática de la que trata el texto y llevar a cabo un riguroso análisis textual, sobre todo en el caso de que este sea especializado para conocer las peculiaridades del género al que nos enfrentamos, y, por último y probablemente el paso más complejo, es imprescindible conocer el método de traducción de contenido web² y ponerlo en práctica.

1.1. Localización de contenido web

En primer lugar, y para ser rigurosos en la denominación del tema que nos ocupa, debemos apuntar que es común la referencia a este tipo de traducción como “traducción o localización de páginas web”, de hecho, son estos términos a los que se recurre con mayor frecuencia para expresar el concepto que aquí tratamos, sin embargo, parece más adecuado acuñar el término “localización de contenido web”.

En los primeros años de Internet, las páginas web no eran como lo son actualmente, se trataba de una página que carecía de enlaces, jerarquías, etc., mientras que ahora el traductor al que se le encarga una traducción de este tipo se encuentra ante un conjunto de páginas y documentos enlazados, por lo que actualmente hablar de “páginas web” no es del todo correcto, ya que cada una de ellas está compuesta por varias páginas que a su vez pueden enlazar con otras. Por lo tanto, como podemos observar, se trata de un término bastante desfasado, pero que sigue utilizándose con asiduidad para denominar un concepto al que no hace referencia en realidad. Aún así, por frecuencia de uso, muchos de los especialistas en esta materia lo utilizan para poder llegar al receptor con mayor facilidad.

Dicho esto, lo que acostumbramos a denominar como “localización de páginas web” queda mejor reflejado si nos referimos a ello como “localización de contenido web”, ya que prescindimos del término “página”, que puede llevar a confusión cuando nos referimos al conjunto del proceso traductor. Además, de esta manera queda perfectamente reflejada la labor real del traductor en este proceso, la traducción del texto y la adaptación del resto del contenido sin influir en ciertos aspectos informáticos, pero en esto profundizaremos más adelante.

Esta nueva expresión, en cuanto a que se trata de un contenido abierto, dinámico y cambiante, se utiliza en contraposición a “localización de productos informáticos”, que incluye la traducción de enciclopedias electrónicas, videojuegos, aplicaciones informáticas, etc., de contenido cerrado y estático. Por tanto, nuestro campo de estudio se reduce a la localización de contenido web, similar en muchos aspectos a la de productos electrónicos, pero con las suficientes diferencias como para tratarlas de manera independiente.

Al referirnos a este tipo de traducción, entendemos que se trata de un proceso que, además de las destrezas lingüísticas y del dominio del campo para el que se ha desarrollado la aplicación, exige también conocimientos básicos en lo referente al proceso de desarrollo

² Mata Pastor, 2005.

del mismo. Sólo así conseguiremos afrontar el trabajo adecuadamente y de una manera eficaz. Por lo tanto, queda claro que nos encontramos ante un tipo de traducción especializada con una serie de propiedades únicas y específicas que la hacen especial y requieren una formación muy concreta por parte del profesional que la llevará a cabo. Además, el auge de esta disciplina nos lleva a una necesidad mayor de especialización, por el cada vez mayor número de especialistas que se dedican a este campo.³

Sería útil aportar una definición de página web que nos sirva de ahora en adelante para delimitar dicho concepto. Se trata de un (sub)conjunto acotado de información en diversos formatos electrónicos, generalmente estructurados y organizados hipertextualmente, que se puede consultar en un ordenador, en una red de ordenadores o varios identificadores alfanuméricos en forma de URL⁴, y que en la actualidad se articula sobre la base técnica de uno o varios lenguajes de etiquetado y de programación, y de varios protocolos de telecomunicaciones, fundamentalmente “HTTP⁵” y “TCP/IP⁶”.

La creación de toda página web comienza con el análisis de los requisitos que deberá cumplir, y posteriormente se lleva a cabo el diseño de la misma y su codificación. Estos dos pasos son los que darán forma y contenido a la misma. Finalmente, aunque debe realizarse a lo largo de todo el proceso, debe probarse el resultado final para verificar la adecuación y buen funcionamiento del sitio web.

Puesto que nuestra labor consiste en localizar contenido web, los conocimientos sobre desarrollo web que debemos adquirir los aplicaremos durante dicha actividad; eso sí, no tenemos que profundizar en ellos de la misma forma que lo haría un profesional de la informática, ya que serán ellos los que creen el sitio web y traten de facilitar al traductor la labor de traducción y localización. Sin embargo, no siempre el creador de la página web la desarrolla teniendo en cuenta su futura localización, en dichos casos la traducción deberá llevarla a cabo o bien un informático con amplios conocimientos lingüísticos de ambas lenguas, o un traductor con conocimientos de localización web.

Para empezar debemos conocer datos imprescindibles como que la mayoría de las páginas web se programan en código HTML⁸, aunque también es muy frecuente el uso de XML⁹, por lo que debemos conocer de forma básica dichos lenguajes. El traductor nunca debe modificar el código de programación, sino que modificará, en caso de necesitarlo, el código de etiquetado, de manera que debe saber reconocerlo, leerlo y cómo trabajar con él.

Cuando se prevé que la página web va a ser traducida, el creador debe internacionalizarla, es decir, debe separar el texto traducible del resto del código fuente. El conjunto de partes que conforman un sitio web, que posteriormente es presentado ante los usuarios, debe funcionar en cualquier circunstancia, país, idioma y cultura¹⁰. Por este motivo, el programador debe crear el documento de manera que el código de programación sea compatible con

³ Hann, Michael (2004).

⁴ *Uniform Resource Locator*.

⁵ *Hypertext Transfer Protocol*.

⁶ *Transmission Control Protocol/Internet Protocol*.

⁷ Reineke, 2005.

⁸ *Hypertext Markup Language*.

⁹ *Extensible Markup Language*.

¹⁰ W3C. “Guía breve de Internacionalización”.

el uso de caracteres y contenidos de los idiomas en cuestión¹¹. Si no se lleva a cabo este procedimiento, la localización será mucho más complicada, tanto para el traductor que se enfrenta a la misma, como para el informático que posteriormente debe enlazar la versión traducida a la versión original.

Un producto totalmente internacionalizado, además de permitir al usuario explotar el software en su propia lengua ajustándose al contexto cultural¹², es aquel que se puede localizar completamente sin necesidad de realizar cambios en el código de programación, sino solamente traduciendo texto, sustituyendo imágenes e iconos y personalizando los parámetros para las opciones de visualización y diseño visual.

El traductor puede encontrarse con dificultades de diversa índole, una de las más comunes es el formato del texto que se le ofrece. El texto puede encontrarse en formato binario, para lo que se necesita utilizar una herramienta de localización; o puede encontrarse en formato fuente, es decir, separado del código fuente, para lo que le bastaría con cualquier editor de textos. Hace tan sólo unos años era prácticamente impensable realizar una localización web en formato binario, pero poco a poco se han ido desarrollando nuevas herramientas que facilitan la labor, muchas de las cuales contienen funciones traslativas. De esta manera, los textos de la interfaz de usuario se visualizan sin el código fuente colindante. Son referenciados mediante un identificador con el lugar que ocupan en el archivo del que han sido extraídos para que los textos traducidos puedan reinsertarse sin ningún problema.

En este trabajo no vamos a adentrarnos en el estudio del desarrollo de contenido web, ya que no es competencia del traductor llevarlo a cabo y hemos apuntado los conceptos básicos que debemos conocer. Sin embargo, hemos de saber que se realiza de manera cíclica, siguiendo unas fases determinadas que se repiten periódicamente para que el contenido no caiga en desuso o quede perdido en la amplitud de la Red de Redes. Esto es relevante para el localizador porque su labor puede llevarse a cabo en cualquiera de las fases y, aunque no es lo habitual, lo ideal sería afrontarla desde el principio para no plantear más problemas de los usuales.

La manera de plantear la localización dependerá en gran medida del tipo de web al que se enfrenta el traductor. Existen infinidad de clasificaciones del contenido web, unas lo catalogan según el público al que se dirige (genérico o específico), otras según los objetivos (informar, entretener, convencer, vender, etc.), pero desde el punto de vista de la localización nos interesa si la Web es dinámica o estática. Un sitio web será dinámico o estático en función de varios parámetros:

- La volatilidad del contenido y frecuencia de su actualización
- El dinamismo visual
- El grado de interactividad que ofrece al usuario
- La arquitectura cliente-servidor
- Etc.

A continuación vamos a proceder a desgranar la estructura de los sitios web, diferenciando para ello dos grandes niveles en los que enmarcar los aspectos que pueden

¹¹ Df. Dal Dosso (2003).

¹² Uren et al., 1993.

resultar relevantes para la localización: comenzaremos con la macroestructura de la Web, abordando los aspectos más generales de la misma, para luego pasar a la microestructura.

1.1.1. Macroestructura

Los sitios web albergan dos niveles de contenido: virtual/visual o real/subyacente.

- *Virtual*: se corresponde con el mapa de navegación que percibe el usuario de la Web, las pautas para organizar el contenido en secciones y subsecciones. Se trata de un contenido vital para la internacionalización y la localización porque los esquemas de navegación y de estructuración mental varían de una cultura a otra, por lo tanto, un mismo sitio web puede tener que modificar las pautas de organización según la lengua en la que se presente.
- *Real*: se trata del conjunto de código informático, contenido textual, gráfico y multimedia que se almacena en contenedores de dos tipos:
 - o Programas o archivos ejecutables: llevan a cabo actividades.
 - o Documentos: se limitan a almacenar contenido.

Este contenido real alberga una cantidad muy elevada de archivos y carpetas, por lo que su control y manipulación es uno de los retos de todo proyecto de localización web.

Es vital que el traductor conserve la estructura organizativa en las remisiones cruzadas entre programas y documentos, y entre instrucciones y contenidos de los resultados deseados, para que, al igual que en la versión original, la Web se visualice y sus elementos dinámicos funcionen. No obstante, si al crear la página web no se prevé que puede ser traducida en un futuro, habrá que llevar a cabo una reestructuración de las carpetas, ya que, además de la carpeta principal, que alberga el resto, existirá una nueva carpeta de similares dimensiones que contiene el resto de archivos y subcarpetas de la traducción. Por lo tanto, la carpeta principal y la nueva pasarán a ser subcarpetas de una mayor en la que se disponga un menú inicial con opciones de elección de idioma, o bien la nueva carpeta pasará a ser subcarpeta de la principal, pudiendo acceder a su contenido el usuario mediante un enlace que se añade a ella.

1.1.2. Microestructura

Una vez vista la estructura general del contenido web, pasamos a analizar su microestructura, que abarca contenido textual, gráfico, multimedia y código informático. El primero siempre tiene que ser modificado, mientras que el resto se modificará o no según el caso, las exigencias de localización y las peticiones del cliente.

En la Web coexisten, habitualmente, lenguaje de programación¹³, lenguaje de etiquetado¹⁴ y texto, por lo que es imprescindible saber diferenciarlos y delimitarlos. Si es

¹³ Lenguaje artificial que puede ser usado para controlar el comportamiento de una máquina, especialmente una computadora. Estos se componen de un conjunto de reglas sintácticas y semánticas que permiten expresar instrucciones que luego serán interpretadas.

¹⁴ Lenguaje artificial que permite dar formato a un texto.

necesario, se debe extraer uno para trabajar con él sin modificar el otro y luego volver a insertarlo una vez modificado.

En nuestro caso, como traductores, tan sólo nos hemos tenido que ocupar del lenguaje de etiquetado y del texto. Este último ya será ampliamente analizado al estudiar el corpus por lo que, en adelante, en este subapartado, nos centraremos en el lenguaje de etiquetado y en los aspectos relacionados en éste que el traductor de contenido web debe conocer, es decir, su gramática, sistematizada por el SGML¹⁵. Por ejemplo, HTML es un lenguaje de etiquetado, mientras que PHP es un lenguaje de programación.

Los lenguajes de etiquetado más empleado son el HTML¹⁶ y el XML, como ya hemos mencionado anteriormente, y albergan tanto contenido textual como informático. El texto localizable suele ir entre etiquetas, las cuales no se traducen, pero puede ser necesario modificarlas.

Las etiquetas que sirven para delimitar pueden ser sencillas, del tipo <P> (párrafo nuevo), o doble, del tipo texto (texto en negrita). La finalidad de las mismas es establecer la manera en la que el contenido se presenta en un navegador. Es algo comparable a una partitura de música, la cual puede ser leída por un intérprete y convertida en música. La partitura son las etiquetas y el intérprete es el navegador, que da forma al código. A continuación se recogen las etiquetas más comunes:

- Una de las normas básicas del lenguaje de etiquetado es que los archivos deben estar compuestos por una cabecera y un cuerpo, cuyos contenidos van delimitados respectivamente por las etiquetas dobles <HEAD><HEAD> y <BODY><BODY>.

El cuerpo de la Web es el grueso del contenido de la misma, incluyendo además otra serie de etiquetas que puedan corresponder. Es recomendable comenzar a traducir el cuerpo y dejar la cabecera para el final, al igual que otros tipos de textos dejamos el título o la introducción para cuando ya hemos terminado el resto.

- El comienzo y el final del archivo van indicados por la doble etiqueta <HTML><HTML>.
- Es frecuente también encontrar instrucciones de programación insertadas mediante etiquetas <SCRIPT><SCRIPT>.
- La etiqueta doble <TITLE><TITLE> es un elemento obligatorio que delimita el título de la página que posteriormente aparecerá en la barra de título de la ventana del navegador. Los buscadores lo consultan para indexar la Web.
- La etiqueta simple <META> no es obligatoria, pero sí muy frecuente. Contiene cuatro tipos de atributos:
 - o Name=
 - o Content=
 - o Scheme=
 - o Http-equiv=

¹⁵ *Standard Gneralised Markup Language.*

¹⁶ Álvarez García (1997).

De los cuatro, tan sólo se traducen los valores del atributo “content=”, en caso de que los atributos “name=” o “http-equiv=” no tengan el valor “Generator” ya que, en tal situación, el valor del atributo “content=” será el nombre completo y la versión de la aplicación con la que se ha creado el documento HTML.

- La presencia de la etiqueta doble <A><A> tampoco es obligatoria, pero suele aparecer con frecuencia. Su contenido aparece en la Web como hipervínculo, remitiendo a una de las partes de la Web, a un documento, a una dirección de correo, etc. Sólo se traduce el valor del atributo “href=”.
- La etiqueta simple tampoco es obligatoria en la Web. Su contenido aparece en el sitio web, pero como texto alternativo que aparece al situar el cursor sobre una imagen. Se traducen los valores del atributo “alt=”, mientras que los del atributo “src=” hay que ajustarlos según dónde se almacene la imagen en cuestión.

Al localizar un archivo es recomendable extremar la atención y no depositar una confianza total en la herramienta que se esté utilizando, ya que es bien sabido que, aunque suelen estar preparadas para afrontar este tipo de situaciones, las máquinas y herramientas automáticas pueden cometer errores. Para que el traductor elija la mejor opción para su labor de localización, a continuación explicamos el funcionamiento de las herramientas de localización.

1.1.3. Herramientas para la localización de lenguaje de etiquetado y texto del contenido web

Como ya hemos apuntado anteriormente, la localización de un sitio web implica la adaptación del contenido textual, gráfico y multimedia, y retocar el código informático para localizarlo cuando sea necesario. Las herramientas que deben utilizarse en estos casos deben combinar funciones de programación y funciones lingüísticas.

Las estrategias de localización más empleadas, ambas con sus seguidores y detractores, son la de traducir los textos en su formato original o bien extraerlos para localizarlos con alguna herramienta y volver a insertarlos. Sea cual sea la técnica que el traductor utilice, debe asegurarse de delimitar perfectamente las partes localizables de las no localizables, de lo contrario el buen funcionamiento de la Web se verá afectado.

A diferencia de lo que ocurre en la localización de contenido informático, en el que la localización sólo puede ser realizada por aquellos que dispongan de una serie de materiales muy exclusivos y unos conocimientos excesivamente amplios, en este tipo de localización el mercado es mucho más amplio y más abierto.

Esselink (2000) lleva a cabo un análisis de las ventajas e inconvenientes del uso de algunos de los tipos de herramientas de localización de archivos HTML y XML más empleados, agrupándolos en cinco grupos: editores de texto (básicos y avanzados), aplicaciones ofimáticas y de autoedición, aplicaciones de creación y diseño web, gestores de memorias de traducción, herramientas de traducción de páginas web.

Presentamos a continuación un esquema resumen de las mismas:

Tabla 1. Valoración de la utilidad de las herramientas de localización.

HERRAMIENTA → ANÁLISIS ↓	EDITORES DE TEXTO	APLICACIONES OFIMÁTICAS Y DE AUTOEDICIÓN	APLICACIONES DE CREACIÓN Y DISEÑO WEB	GESTORES DE MEMORIAS DE TRADUCCIÓN	HERRAMIENTAS DE TRADUCCIÓN DE PÁGINAS WEB
VENTAJAS	<ul style="list-style-type: none"> - Fácil acceso al código para adaptarlo o retocarlo - No presenta los problemas que el resto en cuanto a reinterpretación del código - Puede incorporar funciones de tratamiento de textos - No convierte caracteres especiales o acentuados 	<ul style="list-style-type: none"> - Entorno gráfico¹⁷ WYSIWYG⁴ - Código oculto, protegido de posibles cambios - Funciones de tratamiento de textos 	<ul style="list-style-type: none"> - Entorno gráfico WYSIWYG - Código oculto, protegido de posibles cambios - Prestaciones de validación de código 	<ul style="list-style-type: none"> - Todas las relacionadas con el empleo de memorias de traducción en cualquier tipo de texto - Diferenciación variable de etiquetado y texto 	<ul style="list-style-type: none"> - Válidas para proyectos pequeños - Funciones de tratamiento de textos - Diferenciación de código y texto - Entorno gráfico

¹⁷ *What you see is what you get*

<p>INCONVENIENTES</p>	<ul style="list-style-type: none"> - Entorno no gráfico - Riesgo de alteración accidental del código - Riesgo de no detección del texto - Carencia de memorias de traducción 	<ul style="list-style-type: none"> - Carencia de memorias de traducción - Riesgo de reintepretación de código - Sólo viable con herramienta y versión idénticas 	<ul style="list-style-type: none"> - Carencia de funciones de tratamiento de textos - Carencia de memorias de traducción - Riesgo de reintepretación de código 	<ul style="list-style-type: none"> - Carencias en los filtros de conversión de lenguajes de etiquetado - Problemas para la identificación de scripts¹⁸ - Problemas de incompatibilidad con algunas herramientas 	<ul style="list-style-type: none"> - No permite trabajar directamente sobre el lenguaje de etiquetado para realizar modificaciones puntuales
------------------------------	--	--	---	---	---

¹⁸ Código de programación adicional a la web en cuestión.

<p style="text-align: center;">COMENTARIOS</p>	<p>Existen muchos editores de texto avanzados que identifican con colores código y texto, de manera que quedan claramente delimitados. Además incluyen potentes funciones para la gestión de archivos, búsqueda y sustitución. Muchos también permiten la interconexión con un navegador y la visualización del resultado.</p> <p>Aún así, resulta inconcebible su uso para grandes cantidades de texto. Su empleo es ideal para modificaciones concretas de lenguaje de etiquetado, siempre que se cuente con los conocimientos básicos de la sintaxis de los mismos.</p>	<p>Su uso es totalmente desaconsejable para un proyecto de localización de contenido web.</p>	<p>Su empleo suele ser desaconsejable salvo si se trata de herramientas que garanticen la integridad del código.</p>	<p>Trabaja con filtros demasiado genéricos, por lo que el usuario que no domine con demasiada solvencia los lenguajes de programación y etiquetado puede encontrarse con bastantes dificultades.</p>	<p>Es la herramienta ideal para este tipo de labores, ya que agrupa numerosas ventajitas, y tan sólo presenta un inconveniente significativo, que puede ser compensado sin problemas por otra herramienta como un editor de textos.</p>
---	--	---	--	--	---

Como conclusión de este riguroso análisis podemos extraer que lo ideal es aunar aplicaciones que faciliten tanto la labor de traducción en sí del texto, como de modificación de otro tipo de elementos de índole informática, pero, como podemos observar, no existe ninguna herramienta ideal que aporte todos los beneficios deseados. Por este motivo, para llevar a cabo nuestra labor de localización de un contenido web determinado, tendremos que utilizar conjuntamente los dos tipos de herramientas que mejor realizan las funciones que necesitamos, por un lado, una herramienta de traducción de páginas web que cubra el aspecto textual y, por otro lado, un editor de texto que cubra el aspecto referente al código de etiquetado.

Una vez conocidos los aspectos teóricos básicos sobre localización de contenido web, tenemos cubiertos los conocimientos técnicos necesarios para abordar la traducción de una web. Ahora sólo nos queda profundizar en el tema que en él se aborda, y considerar las herramientas para el análisis del texto.

A continuación vamos a ver la aplicación práctica en un ejemplo que consideramos bastante significativo, ya que representa a la mayoría de webs a las que puede enfrentarse un localizador de contenido web. Se trata de un texto sobre una temática ajena a la informática, en este caso sobre psicología del lenguaje y análisis de la conducta, presentado en un soporte informático en Internet.

2. ETAPAS DEL PROCESO DE LOCALIZACIÓN DE CONTENIDO WEB

Como ya hemos apuntado, son dos los niveles temáticos que debemos tener en cuenta: los aspectos informáticos y aquellos correspondientes al tema en cuestión; y, por supuesto, cada uno de ellos en su relación con la traducción y localización de este tipo de textos.

1. El primer paso es, sin duda, el *acercamiento a la Web* en cuestión. Este acercamiento consiste en una *lectura minuciosa* de la Web atendiendo a su *estructura externa e interna*, y la *comprobación de los enlaces y funciones* de la misma.
2. A este primer contacto le sigue la *recopilación del material* necesario para la *documentación* que necesitaremos a la hora de llevar a cabo nuestra labor. En el caso concreto que presentamos, los materiales versan sobre localización web, psicopedagogía y atención temprana, y sobre análisis textual.
3. Una vez recopilado el material, proceso que evidentemente se alarga prácticamente tanto como la realización del encargo por el estallido de nuevas dudas y nuevas necesidades, procedemos a su *riguroso análisis y estudio*. Esta fase nos conduce a un estado de conocimiento suficiente para comenzar a realizar la práctica en cuestión.
4. En esta nueva etapa del proceso, dirigimos nuestra atención hacia la búsqueda de una herramienta adecuada, fiable y de calidad para la traducción automática del contenido web. Como hemos podido observar en los fundamentos teóricos relacionados con la localización, la elección se realiza mediante un balance de ventajas y desventajas. Tras un minucioso análisis de las posibilidades que se nos presentan, la opción más conveniente es el uso de una *herramienta de traducción de páginas web* para

cubrir los aspectos textuales, y de una *herramienta avanzada de edición de texto* para modificar aquellas partes de etiquetado que fueran necesarias.

Tras tomar esta decisión, es necesario valorar varias herramientas del mismo tipo antes de escoger las dos que finalmente utilizaremos.

En nuestro caso, y consideramos que para cualquier encargo de características similares, llegamos a la conclusión que las mejores opciones son CatsCradle 3.8 (herramienta de traducción de páginas web) y Notepad++ (editor de texto avanzado). La decisión de utilizar estas dos herramientas en lugar de otras de su misma categoría es debida a la claridad de su interfaz, que permite trabajar y localizar los contenidos de manera más sencilla que otras. Esto es muy importante para profesionales de la traducción no expertos en herramientas informáticas, ya que estas herramientas permiten unos resultados óptimos que, por la carencia de conocimientos, no serían posibles si no dispusiéramos de ellas.

Para escoger dichos programas ya es necesario analizar cuestiones sobre su modo de empleo y sus prestaciones, al igual que la de otros cuyo uso sopesamos, sin embargo, dicha tarea se hace más relevante al tomar la decisión. En ese momento es imprescindible leer muy atentamente los manuales de usuario para conocer a fondo ambos programas.

- **CatsCradle 3.8:** se trata de una herramienta de traducción de páginas web para profesionales especializados en este ámbito. Permite traducir documentos HTML separando el texto traducible del resto (código de programación y de etiquetado), de manera que el traductor no tiene que preocuparse del riesgo de modificar código y puede ceñirse al texto y a los problemas de traducción del mismo.

CatsCradle extrae automáticamente el texto y lo muestra en un editor incorporado donde se puede traducir. Además, se trabaja con pantalla partida, la mitad inferior se corresponde con el editor, y la mitad superior con la vista de un navegador web en el que, a medida que se va traduciendo el texto, aparece en el formato web del original, permitiendo así la comprobación del resultado sin necesidad de guardarlo.

Este programa, además, ofrece la posibilidad de realizar búsquedas en el texto, contar palabras, elegir la fuente, realizar correcciones ortográficas y utilizar glosarios y memorias de traducción.

- **Notepad++:** se trata de un editor de texto y de código fuente que soporta varios lenguajes de programación.

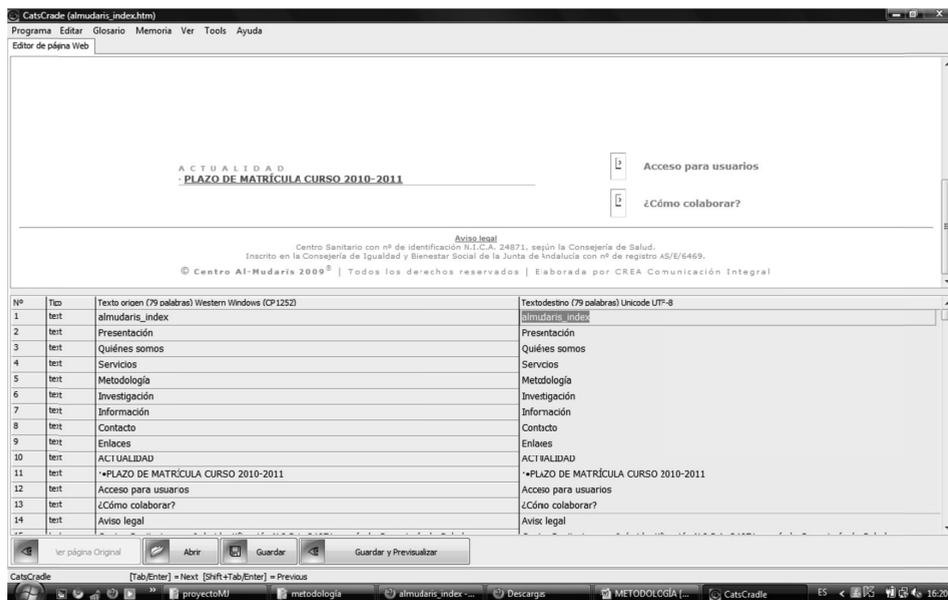
Notepad++ se caracteriza por sus grandes ventajas visuales, ya que colorea el código para diferenciarlo del texto traducible, de esta manera, su localización es mucho más rápida. Permite comprobar el resultado sin guardar, al igual que ocurría con CatsCradle, admite trabajar sobre varios textos al mismo tiempo, separándolos en distintas pestañas, y además, permite buscar y reemplazar palabras o expresiones.

Una vez analizadas las prestaciones de cada uno de los programas vamos a describir el proceso:

En primer lugar, acudimos a la Web que queremos traducir y guardamos todas las páginas de las que se compone en una carpeta creada para tal fin. En este paso debemos asegurarnos de que en dicha carpeta se nos presenta tanto el archivo .htm o .php como la carpeta que corresponde a sus “files”.

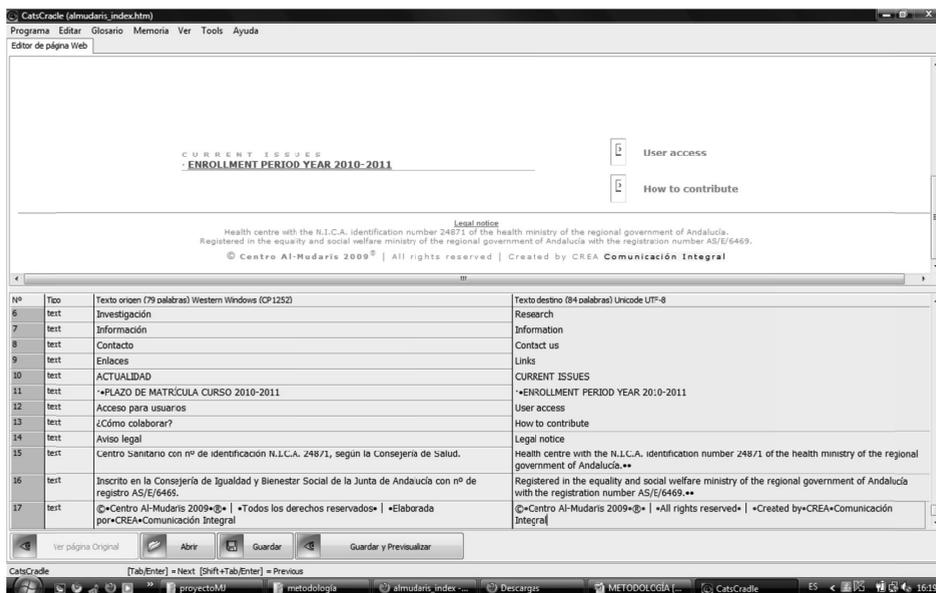
El siguiente paso es entrar en *CatsCradle* y “abrir” la página que vamos a traducir. Una vez abierta, en la mitad inferior de la pantalla aparece el texto contenido en la página, en dos columnas distintas para poder ser modificado, como bien podemos apreciar en la imagen que añadimos más abajo.

Además, observamos que en la mitad superior, como apuntábamos anteriormente, se nos presenta el texto según la forma y contenido que nosotros hayamos dispuesto en la columna de la derecha.



Captura de *CatsCradle*

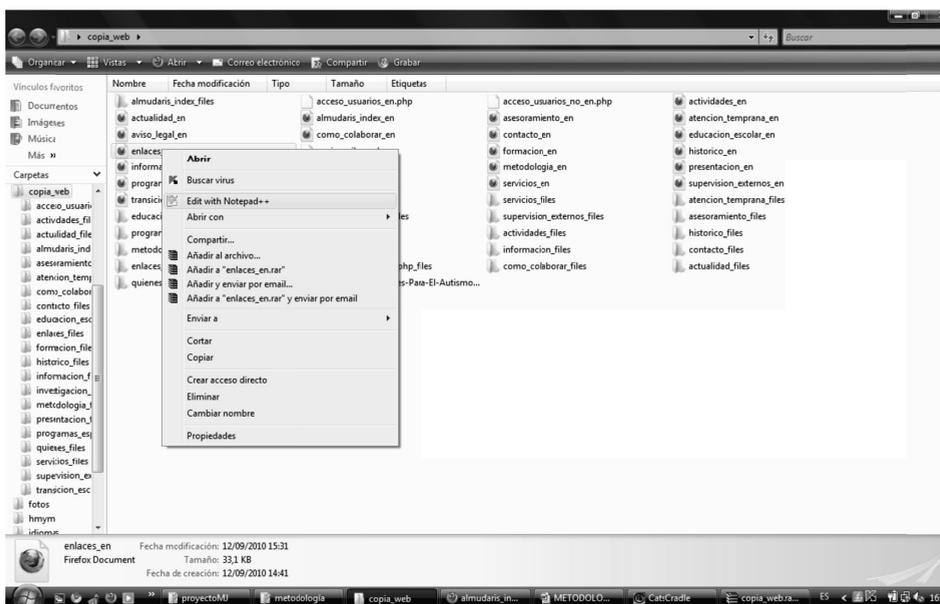
Una vez extraído el texto, podemos comenzar a traducirlo. No sin antes haber analizado el corpus textual de manera detallada, como se hace en el capítulo 4.



Captura de CatsCradle

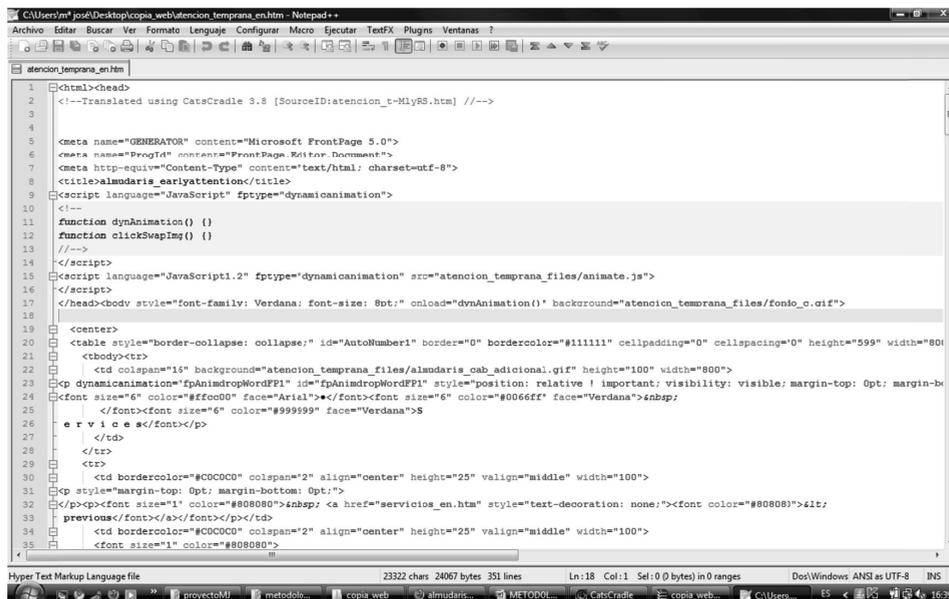
Al finalizar la traducción de cada una de las páginas de nuestra Web, debemos guardarlas mediante la opción “guardar” que aparece en la parte de abajo. Automáticamente se guarda en la carpeta desde la que partimos. Una vez traducidas todas las páginas, acudimos a la carpeta inicial y modificamos los nombres de cada uno de los archivos de la siguiente manera: si el archivo original se denominaba “contacto.htm”, debemos cambiar el nombre por el de “contacto_en.htm”, de manera que sigamos los estándares del lenguaje HTML, según los cuales el nombre del archivo traducido debe ir seguido de las dos primeras letras de la lengua de llegada.

Una vez hecho todo esto, tenemos nuestra página web traducida. El siguiente paso es modificar aquellas partes del código de etiquetado que deben modificarse por exigencias de la lengua de llegada, o bien para mejorar aspectos estructurales del texto. Para este paso utilizamos *Notepad++*. Desde la carpeta en la que guardamos nuestros archivos traducidos, accedemos a aquellos que queramos modificar seleccionándolo con el botón derecho y eligiendo la opción de “*EditwithNotepad++*”. De esta manera somos dirigidos de forma automática hacia el programa, el cual abre el archivo seleccionado.



Captura de la carpeta contenedora de los archivos de la Web

Una vez dentro de Notepad++, podemos ver el código fuente de la página web y el texto de la misma, diferenciados por colores. Asimismo, podemos ver que en la parte superior aparece una referencia al programa con el que se ha traducido la Web.



Captura de Notepad++

En este punto, el profesional puede encontrarse ante un código bien estructurado o ante un código con ordenación caótica, lo cual dificulta bastante la labor del profesional.

Como podemos apreciar, en este caso el código resulta bastante caótico a simple vista. Para quien no conozca este ámbito, vamos a mostrar una imagen del texto de otra página web bien estructurada. El primero se corresponde con la página web de la UCO y el segundo con la del Centro AI-Mudarís.

Observamos que en el segundo caso resulta mucho más difícil encontrar un término o etiqueta concretos, ya que las líneas son demasiado largas y las etiquetas muy imprecisas en muchos casos. Además, son muchas las ocasiones en las que operaciones que podrían realizarse en muy poco espacio tienden a alargarse y a ocupar líneas completamente innecesarias. Todo esto tiene sus consecuencias en el proceso de localización por parte del traductor, ya que debe invertir mucho más tiempo del que se emplearía en condiciones normales en una labor de este tipo, y los errores que éste puede cometer se deben, en gran medida, a la caótica estructuración del código de etiquetado.

```

Código fuente de http://www.uco.es/ - Mozilla Firefox
Archivo Editar Ver Ayuda
<tr>
<td width="180" height="61" rowspan="2" class="C2">
<table cellpadding="0px" cellspacing="0px" class="T2">
<tr>
<td height="6px" colspan="3"></td>
</tr>
<tr>
<td width="14px"></td><td width="152" height="18px"><a href="..">Organización institucional</a></td>
<td width="14px"></td></tr>
<tr>
<td colspan="2" height="18px"><a href="investigacion/">Investigación y empresa</a></td>
<td></td>
</tr>
<tr>
<td colspan="2" height="18px" class="C4"></td><td></td>
</tr>
<tr>
<td height="18px"><a href="/docencia/">Docencia en la Universidad </a></td>
<td colspan="2"></tr>
<tr>
<td colspan="2" height="18px" class="C4"></td><td></td>
</tr>
<tr>
<td colspan="2" height="18px"><a href="/estudios/">Estudios y oferta académica</a></td>
<td colspan="2"></tr>
<tr>
<td colspan="2" height="18px" class="C4"></td><td></td>
</tr>
<tr>
<td colspan="2" height="18px"><a href="/servicios/">Servicios para universitarios</a></td>
<td colspan="2"></tr>
<tr>
<td colspan="2" height="18px" class="C4"></td><td></td>
</tr>
<tr>
<td colspan="2" height="18px"><a href="/centros/">Centros, campus y departamentos</a></td>
<td colspan="2"></tr>
<tr>
<td colspan="2" height="18px" class="C4"></td><td></td>
</tr>
<tr>
<td colspan="2" height="18px"><a href="/gestion/">Gestión administrativa</a></td>
<td colspan="2"></tr>
<tr>
<td colspan="2" height="18px" class="C4"></td><td></td>
</tr>
<tr>
<td colspan="2" height="18px"><a href="/internacionalcooper/">Internacional y cooperación</a></td>
<td colspan="2"></tr>
<tr>
<td colspan="2" height="18px" class="C4"></td><td></td>
</tr>
<tr>
<td colspan="2" height="18px"><a href="/congresos/">Congresos</a></td>
<td colspan="2"></tr>
<tr>
<td colspan="2" height="18px"></td>
<td colspan="2"></tr>
</table>
</td>
</tr>
<td width="199" height="110" class="C1b">
<table cellpadding="0px" cellspacing="0px" class="T2">
<tr>
<td height="5px" colspan="2"></td>
</tr>
<tr>
<td width="7"></td>
<td height="19px"><a href="/docencia/">Docencia en la Universidad</a></td>

```

Captura Notepad++ (UCO)

```

196 <td bordercolor="#C0C0C0" style="border-style: none; border-width: medium; height="25" valign="top" width="50">&nbsp;
197 </td>
198 <td style="border-style: none; border-width: medium; colspan="3" bordercolor="#C0C0C0" align="left" height="25" valign="top" width="150">
199 <p style="margin-left: 5px; margin-top: 0pt; margin-bottom: 0pt; dynamicanimation="fpAnimFormatRollerFP1" fprolloverstyle="background-color
200 <font size="2"><font color="#3296fa">&nbsp;&nbsp;&nbsp;&#9679</font> </font>
201 <font size="1"><a href="informacion_en.htm" style="text-decoration: none;">
202 <font color="#000000">Información</font></a></font></p></td>
203 </td>
204 <td bordercolor="#C0C0C0" style="border-style: none; border-width: medium; height="25" valign="top" width="50">&nbsp;
205 </td>
206 <td colspan="3" style="border-style: none; border-width: medium; bordercolor="#C0C0C0" align="left" height="25" valign="top" width="150">
207 <p style="margin-left: 5px; margin-top: 0pt; margin-bottom: 0pt; dynamicanimation="fpAnimFormatRollerFP1" fprolloverstyle="background-color
208 <font size="2"><font color="#55a8eb">&nbsp;&nbsp;&nbsp;&#9679</font> </font>
209 <font size="1"><a href="contacto_en.htm" style="text-decoration: none;">
210 <font color="#000000">Contacto</font></a></font></p></td>
211 </td>
212 <td bordercolor="#C0C0C0" style="border-style: none; border-width: medium; height="25" valign="top" width="50">&nbsp;
213 </td>
214 <td colspan="3" style="border-style: none; border-width: medium; bordercolor="#C0C0C0" align="left" height="25" valign="top" width="150">
215 <p style="margin-left: 5px; margin-top: 0pt; margin-bottom: 0pt; dynamicanimation="fpAnimFormatRollerFP1" fprolloverstyle="background-color
216 <font size="2"><font color="#8e6fd">&nbsp;&nbsp;&nbsp;&#9679</font> </font>
217 <font size="1"><a href="enlaces_en.htm" style="text-decoration: none;">
218 <font color="#000000">Enlaces</font></a></font></p></td>
219 </td>
220 <td bordercolor="#C0C0C0" style="border-style: none; border-width: medium; height="25" valign="top" width="50">&nbsp;
221 </td>
222 <td colspan="3" style="border-style: none; border-width: medium; bordercolor="#C0C0C0" align="left" height="25" valign="top" width="150">
223 <p style="margin-left: 5px; margin-top: 0pt; margin-bottom: 0pt; dynamicanimation="fpAnimFormatRollerFP1" fprolloverstyle="background-color
224 <font size="2"><font color="#8e6fd">&nbsp;&nbsp;&nbsp;&#9679</font> </font>
225 <font size="1"><a href="usuarios_en.htm" style="text-decoration: none;">
226 <font color="#000000">Usuarios</font></a></font></p>

```

Captura Notepad++ (Centro AI-Mudarís)

Todo esto puede deberse, en cierto modo, a la herramienta elegida para crear la Web, que aparece al principio del código de cada una de las páginas: Microsoft FrontPage 5.0. Este recurso no es muy utilizado por los programadores web por la cantidad de imprecisiones y errores que acarrea.

Pues bien, una vez dentro del código fuente procedemos a realizar aquellas modificaciones que el traductor estima necesarias, tan sólo en el código delimitado por las etiquetas `<body></body>`:

- Cursiva en títulos de obras, artículos, etc., mediante la etiqueta ``, o `<i></i>`
- Dobles espacios o espacios que no aparecen, mediante la etiqueta ` `;
- Cambios de párrafo mediante la etiqueta `<P></P>` (puede usarse la versión simple “`<P>`” o la doble “`<P></P>`”).
- Cambio de la etiqueta “`lang="ES-TRAD"`” por “`lang="EN-TRAD"`”, por el cambio de idioma
- Modificación de algunas palabras en negrita mediante la etiqueta ``
- Cambio de línea mediante la etiqueta `
`
- Cambio de los enlaces entre las distintas páginas, ya que todas las páginas traducidas deben ir enlazadas de la misma manera que lo están las originales en la página web original. Los enlaces se encuentran delimitados por las etiquetas `<a>`, en el código nos aparecen los enlaces de la página web, sin embargo, debemos modificarlos

por el enlace a la ubicación de la carpeta en la que se encuentra la traducción de dicha página.

Así, por ejemplo, el enlace “<http://www.centroalmudaris.com/presentacion.htm>” pasará a ser “file:///C:/Users/m%C2%AA%20jos%C3%A9/Desktop/copia_web/presentacion_en.htm”. Conseguimos el mismo efecto si el enlace es “presentacion_en.htm”.

- Lo mismo ocurre con las imágenes, ya que algunas es necesario modificar o comprobar su enlace. Aparecen delimitadas por la etiqueta ``

Una vez realizados los cambios que se han considerado oportunos y comprobado su funcionamiento y adecuación, tanto durante el proceso como después de finalizarlo, la localización de la Web está completada. El resultado final es una carpeta que contiene todos los archivos, textuales y no textuales, que conforman la página web. Para visualizarla basta con abrirla en cualquier navegador. Por último, sólo queda entregarla para que sea un informático quien la una a la versión original, ya sea mediante un menú inicial en el que se pueda elegir cualquiera de las dos versiones, o mediante un enlace desde la página de inicio de la versión en español, que nos conduzca a la versión en inglés.

3. CONCLUSIONES

Con el presente artículo hemos pretendido dar a conocer a los profesionales de la traducción un aspecto actual y relevante en este campo como es la localización de contenido web, dados los múltiples cambios, sobre todo de índole tecnológica, que se están produciendo y que afectan tanto al proceso traductológico como a las herramientas que deben utilizarse. La inmediatez cada vez más requerida en los encargos obliga a los traductores a adaptarse a los programas de traducción emergentes para cumplir plazos y expectativas.

Basándonos en el estudio de Esselink (2000), hemos realizado un detallado análisis de las herramientas disponibles para este tipo de trabajos, es decir, *editores de texto, aplicaciones ofimáticas y de autoedición, aplicaciones de creación y diseño web y gestores de memorias de traducción*. Posteriormente, se han evaluado dichas herramientas y se han extraído conclusiones, indicando la mayor o menor idoneidad de su uso por parte del traductor.

Tras la selección de las herramientas, se ha detallado su uso mediante una minuciosa explicación acompañada de imágenes, con la intención de facilitar el aprendizaje de su manejo.

Por todo esto, consideramos que este trabajo supone una aportación de utilidad para todo profesional que reciba un encargo relacionado con el proceso de localización de contenido web o que esté interesado en conocer más acerca de este nuevo campo de nuestra disciplina.

REFERENCIAS BIBLIOGRÁFICAS

- ÁLVAREZ GARCÍA, A. 1997. *HTML. Creación de páginas Web*. Madrid: Anaya.
- AREVALILLO DOVAL, J. J. 2002. “Y tú, ¿traduces o proyectas?”. *La linterna del traductor* 1: 22-27.

- DAL DOSSO, H. “La internacionalización de sitios web”. <http://www.traductores.org.ar/wwwisis/BASE-1/PDF/01828.pdf> .
- ESSELINK, B. 2000.“Practical Guide to Localization”.BibliotecaUniversitaria de Córdoba. http://medina.uco.es/search~S6*spi?/Ylocalization&searchscope=6&SORT=R/Ylocalization&searchscope=6&SORT=R&SORT=RZ&extended=0&SUBKEY=localization/1,88,88,B/1856~b1281322&FF=Ylocalization&searchscope=6&SORT=R&4,4,,1,0 .
- ESSELINK, B. ED. 2000.*A Practical Guide to Localization*.Ámsterdam: John Benjamins.
- HANN, M. 2004. “Basis for Scientific and Engineering Translation: English-German-English”. Biblioteca Universitaria de Córdoba. http://medina.uco.es/search~S6*spi?/Yscientific+translation&searchscope=6&SORT=R/Yscientific+translation&searchscope=6&SORT=R&SORT=RZ&extended=0&SUBKEY=scientific%20translation/1,15,15,B/1856~b1309517&FF=Yscientific+translation&searchscope=6&SORT=R&1,1,,1,0 .
- LOUREIRO PERNAS, M. 2007. “Paseo por la localización de un videojuego”. *Tradumática* 5.
- MATA PASTOR, M. 2005. “Localización y traducción de contenido web”. En: DetlefReineke: *Traducción y localización. Mercado, gestión y tecnologías*. Las Palmas de Gran Canaria: Anroart Ediciones.
- MUÑOZ SÁNCHEZ, P. 2006. “Electronic Tools for Translators in the 21st Century”. *Translation Journal* 10. <http://www.bokorlang.com/journal/38tools.htm>
- PÉREZ FERNANDEZ, L. 2010. “La localización de videojuegos inglés-español. Aspectos técnicos, metodológicos y profesionales” (tesis doctoral). Universidad de Málaga. Dirección: Emilio Ortega Arjonilla y Miguel Ángel Candel Mora. Universidad de Málaga. TRANS 2010.
- REINEKE, D. 2005. “Herramientas de localización”. En: DetlefReineke: *Traducción y localización. Mercado, gestión y tecnologías*. Las Palmas de Gran Canaria: Anroart Ediciones. 187-252.
- SOMERS, H. ED. 2003.*Computers and Translation: A translator's guide*. Ámsterdam: John Benjamins.
- UREN, E. ET AL. 1993. *Software Internationalization and Localization: an Introduction*. New York: Van NostrandReinhold.
- W3C. “Guía breve de Internacionalización”. <http://www.w3c.es/Divulgacion/GuiasBreves/internacionalizacion> .