

# GESTORES DE MEMORIAS DE TRADUCCIÓN DE SOFTWARE LIBRE

José Manuel Muñoz Muñoz y Mercedes Vella Ramírez  
Universidad de Córdoba

## Resumen

Los programas de gestión de memorias de traducción constituyen una de las principales herramientas en el ámbito de la Traducción Asistida por Ordenador (TAO), especialmente cuando se trabaja con textos muy repetitivos. En el presente artículo se analizan las memorias de traducción de software libre y se proponen criterios que es recomendable tener en cuenta a la hora de optar por una u otra herramienta. Tras una breve exposición de los principios que inspiran el movimiento a favor del software libre y su extensión en el ámbito de la traducción, expondremos los criterios que pueden ser válidos para decidir acerca de la fiabilidad de los gestores de memorias de traducción de software libre disponibles. Por último, evaluaremos, de acuerdo con esos criterios, su grado de madurez, su estabilidad y su capacidad para ofrecer al traductor el soporte adecuado.

**Palabras clave:** traducción asistida, memorias de traducción, software libre.

## Abstract

Translation Memory Management Software is one of the main tools in Computer Assisted Translation (CAT), especially when handling highly repetitive texts. In this paper, we analyse free software translation memory programs and propose certain criteria to bear in mind when deciding which particular program to choose. After a brief introduction to the principles which inspire the free software movement and its impact in the field of translation, we describe the criteria by which the reliability of the different free software translation memory programs may be assessed. In accordance with these criteria, we then offer a review of the various programs available, focusing on their stability, maturity and capacity to provide adequate support for translators.

**Keywords:** computer-assisted translation, translation memories, free software.

## 1. Introducción

Los traductores profesionales, ya trabajen como *freelance* o formando parte de la plantilla de una empresa, y especialmente aquellos cuyo trabajo consiste en traducir textos que permiten un alto grado de reutilización de traducciones previas, utilizan con frecuencia memorias de traducción que les permiten optimizar el rendimiento

de su trabajo. A la hora de optar por un determinado programa para la gestión de las memorias, el traductor frecuentemente considera sólo aquellos programas que cuentan con una amplia implantación en el mercado, tales como *SDL-Trados*, *Wordfast*, *Dèjà-vu*, etc., y rara vez se plantean la utilización de gestores de traducción de software libre. En el presente trabajo expondremos las alternativas que el software libre ofrece al traductor profesional, sus posibles ventajas e inconvenientes frente a los programas de gestión de memorias de código cerrado, como los antes mencionados, y las cautelas que el traductor que opte por el software libre debe tener a la hora de elegir un programa de gestión de memorias de traducción concreto<sup>1</sup>. Por último, analizaremos, desde el punto de vista del traductor, los programas de software libre disponibles en *Sourceforge*, el mayor repositorio de software libre en la red. Nuestro análisis no se centrará en los aspectos de operatividad de los programas o en sus características, ni estableceremos comparaciones con los programas de software de propietario. Nuestro propósito es más bien ofrecer parámetros y criterios que permitan establecer el grado de madurez de los distintos proyectos<sup>2</sup>, sus medios de ayuda al usuario y sus posibilidades de continuar siendo una herramienta válida para el traductor a corto o medio plazo.

## 2. ¿Qué entendemos por software libre?

El concepto “software libre” abarca todos aquellos programas y aplicaciones que se distribuyen bajo una licencia que permite al usuario modificar, hacer copias y distribuir libremente un producto de software, con atribución o sin ella, mediante una licencia de código abierto (cf. Fogel, 2009). Aunque las licencias más comunes en el ámbito del software libre son las que se denominan GPL (*General Public License*), han surgido muchos otros tipos de licencias que difieren de éstas en diferentes aspectos, una lista completa de estas licencias puede consultarse en la *Open Source Initiative*<sup>3</sup>. La disponibilidad del código del programa para que el usuario pueda conocerlo y modificarlo en su caso es la característica que comparten todas estas licencias. Los programas distribuidos bajo estas licencias son frecuentemente, aunque no necesariamente, gratuitos.

1. En el presente artículo no expondremos las características generales de los gestores de memorias de traducción, por considerarlo innecesario para el propósito de nuestro trabajo, ya que las funciones que realizan los gestores de software libre no se diferencian en lo sustancial de las que realizan los programas de software de propietario.

2. A lo largo de este artículo utilizaremos términos tales como gestores/programas/herramientas cuando nos referimos a la aplicación informática en sí, y proyectos cuando nos referimos al conjunto de actividades (lanzamiento, actualizaciones, requerimientos de soporte, programación, evaluación...) que se generan alrededor de la aplicación.

3. <http://www.opensource.org/licenses/alphabetical>

Existen otras denominaciones tales como software abierto, FOSS (*Free, Open Source Software*), FLOSS (*Free/Libre Open Source Software*), *Copyleft*, etc., que ponen el acento en distintos aspectos del movimiento a favor del software libre.

Estas licencias deben distinguirse de las que se conocen como *Freeware* o *Shareware* o de *dominio público*.

Las *licencias freeware* se utilizan con frecuencia para distribuir programas de pequeñas aplicaciones puestas a disposición del público de manera gratuita pero sin revelar necesariamente su código y sin que respondan a la filosofía de desarrollo de una comunidad de usuarios, lo que significa que éstos podrán utilizar el programa pero no modificarlo o adaptarlo en modo alguno. En el ámbito de las memorias de traducción, un ejemplo de programa bajo licencia *freeware* sería *Wordfisher*, una aplicación desarrollada por Környei Tibor como complemento de *MS Word*.

Las *licencias shareware* permiten la descarga y libre distribución de un programa pero exigen el pago de una licencia si el programa se utiliza. Las licencias *shareware* con frecuencia establecen un límite temporal de prueba del programa, transcurrido el cual debe pagarse la licencia para seguir utilizándolo. El propietario del software no proporciona el código fuente del programa y mantiene un control absoluto sobre su uso.

Las *licencias de dominio público* son aquellas en las que el propietario de los derechos los ha puesto a disposición de la comunidad o aquellas en las que los derechos de autor han expirado con el transcurso del tiempo de acuerdo con las leyes de derechos de autor vigentes en cada país.

### 3. FSF y OSI: dos enfoques diferentes

Bajo la denominación genérica FOSS conviven dos filosofías diferentes que coinciden en la idea de cooperación como elemento clave en la producción, uso y difusión de programas informáticos y que se han desarrollado de modo paralelo en los últimos años, sobretodo a partir de la generalización del uso de las tecnologías de la información y las comunicaciones y en particular de Internet.

Su desarrollo se produce a partir de la cultura “*hacker*” por iniciativa de una pequeña comunidad de programadores y analistas que trabajaban en laboratorios informáticos de universidades de Estados Unidos, liderados por Richard Stallman y como reacción a la apropiación del software que se había venido desarrollando en esos laboratorios por parte de compañías comerciales. Stallman (1998) expone en profundidad la historia de la creación del movimiento y la filosofía que lo inspiró.

La *Free Software Foundation* (FSF) dirige su actividad a la defensa de la libertad del usuario de programas informáticos no sólo para utilizar esos programas con cualquier finalidad, hacer copias y redistribuirlas, sino también para estudiar su funcionamiento, adaptarlo a sus necesidades específicas o mejorarlo poniendo esas mejoras a disposición del público. Para que estas libertades puedan ejercerse de

manera efectiva, una condición previa es que pueda accederse libremente al código fuente de dichos programas y a sus sucesivos cambios.

La fuerza motriz de la propuesta de la FSF es la de la libertad de cooperación entre agentes diversos que se implican en un proyecto común con el convencimiento de que ese principio de cooperación redundará en beneficio de todos. La FSF pone el acento en la libre disposición del código fuente del software, no necesariamente gratis. A este respecto se ha hecho famosa la frase gráfica de Stallman “free as in free speech, not as in free beer”. La restricción más importante que impone el software propietario no va dirigida a los aspectos económicos que rigen la circulación de los programas, sino más bien al hecho de que el propietario del código dicta las condiciones bajo las que ese código puede usarse, lo que en la práctica significa la imposición de restricciones estrictas que limitan con frecuencia la propia funcionalidad del software.

La *Open Source Initiative* (OSI) propugna postulados similares, aunque desde una perspectiva ligeramente diferente, centrada en criterios de eficacia en el desarrollo y la mejora del software. La hipótesis de la OSI establece que si los programadores pueden disponer del código fuente de los programas, éstos lo mejorarán, depurarán y optimizarán a una velocidad mucho mayor que la que es habitual en software propietario, donde los cambios sólo se producirán cuando convenga a los planes comerciales de la empresa que los comercializa. Esta hipótesis se ha visto confirmada de manera espectacular en desarrollos tales como el software de servidor HTTP *Apache* o el navegador *Mozilla*, que representan ejemplos muy notables de la aplicación de esta filosofía.

La propuesta de la OSI está más centrada en la producción de software fiable que pueda ser utilizado de manera consistente en entornos profesionales o de negocios, mientras que la FSF se centra más en los aspectos filosóficos o en las implicaciones éticas de la protección del software propietario.

Dado que el software libre puede ser modificado y redistribuido sin limitación, una de las áreas en las que se ha experimentado una mayor actividad, y que es de especial relevancia para los traductores, ha sido la de la traducción y localización del software. En efecto, la localización del software libre y la traducción de la documentación asociada a ese software es un área tremendamente dinámica en la que trabaja un número importante de personas, traductores o no, en la mayoría de los casos de forma altruista. Como ejemplo, desde que en el año 2000 *Sun Microsystems* liberó la suite *Open Office*, se creó la *Native Language Confederation*<sup>4</sup> para albergar los proyectos de localización y traducción del software y su documentación a diferentes lenguas. Hasta el momento existen distintos niveles de localización, ayuda al usuario y control de calidad en más de ochenta lenguas, entre las que se incluyen no sólo lenguas con un elevado número de hablantes, sino también lenguas minoritarias para

4. <http://projects.openoffice.org/native-lang.html>

las que las compañías de software propietario no proporcionan versiones por no ser comercialmente rentable.

#### 4. Software libre y traducción

El uso del software libre entre los traductores no se ha podido cuantificar hasta el momento con precisión. Uno de los indicadores que podríamos utilizar para conocer, al menos de forma aproximada, este dato es la extensión de uso que puede tener *Linux*, el sistema operativo libre más extendido en el segmento de la informática de usuario y el que mejor representa el espíritu que subyace al movimiento de software libre, ya que otros que utilizan la denominación, como *Mac OS X*, incluyen módulos de código cerrado. El *TM Survey* (Lagoudaki, 2006) revela que un porcentaje mínimo de la muestra analizada utilizaba *Linux* (2%), mientras que una gran mayoría (92%) utilizaba *Windows*, y el resto declaraba utilizar *Mac OS* (4%) o una combinación de varios sistemas operativos (2%).

Al interpretar estos datos debe tenerse en cuenta el modo en que se llevó a cabo el estudio y el hecho de que un alto porcentaje de los encuestados declaraba tener un alto nivel de conocimientos informáticos y dedicarse preferentemente a la traducción de software técnico, ambas características sin duda favorecerían una mayor presencia de usuarios de software libre que en un grupo no marcado, con lo cual cabe deducir que probablemente la incidencia de *Linux* en el ámbito de los profesionales de la traducción sea aún menor. Se ha de tener en cuenta que una estimación del número de usuarios de *Linux* es especialmente difícil de establecer dada la libertad con que el software puede ser distribuido, en muchos casos sin conocimiento de empresa alguna o sin que tenga consecuencias de tipo comercial o económico. En los últimos años se ha producido un notable incremento del uso de software libre debido al interés de organismos oficiales e instituciones en tener un mayor control sobre la información y la seguridad de sus sistemas. Wheeler (2007) presenta una panorámica general del uso de software libre en relación con sistemas propietarios que revela que la presencia de software libre en la red es mucho mayor de lo que el usuario final puede percibir.

No parece, pues, aventurado afirmar que la presencia del sistema operativo libre en el ámbito de la traducción es similar a la que puede tener en general en la informática de usuario no específicamente dedicada a la traducción.

#### 5. Pros y contras del software libre

Llegados a este punto habremos de preguntarnos cuáles son las razones por las que un usuario, sea o no traductor, decide utilizar software libre. Las razones pueden variar enormemente de unos usuarios a otros dependiendo de su entorno profesional,

la naturaleza del trabajo, o incluso los gustos personales o las concepciones sociales. Entre ellas podemos destacar las siguientes:

- a) El control y la flexibilidad que el software libre proporciona tanto sobre el software en sí como sobre los datos que éste maneja. Esta característica hace que en los últimos años un número importante de organizaciones e instituciones se decanten por la opción del software libre ante la incertidumbre que el desconocimiento del código fuente puede inspirar en sistemas en los que la seguridad e integridad de los datos son factores primordiales.
- b) El coste es un factor significativo para muchos usuarios. Aunque el uso de software libre no haya de ser necesariamente gratuito, sí lo es en la mayoría de los casos y, cuando no lo es, su coste siempre es muy inferior al uso del software propietario equivalente, dado el alto precio de las licencias de los programas comerciales.
- c) El estilo de desarrollo de los productos de software libre responde a un modelo de evolución y no de revolución. Con el software propietario es frecuente que se den saltos en el desarrollo de un producto que dejan obsoleta una determinada versión del programa y obligan al usuario a migrar sus datos a un nuevo sistema o a una nueva versión. El software libre evoluciona por las aportaciones, en muchos casos mínimas, que contribuyen los desarrolladores, lo que tiende a evitar situaciones de cambios forzados de sistema.
- d) El soporte de los programas, que en ocasiones puede representar uno de los puntos débiles de esta opción, en la práctica rara vez constituye un problema real, siempre que se adopten las precauciones necesarias respecto a la elección del producto adecuado. Los productos bien establecidos y que mantienen un nivel de actividad sostenida a lo largo de los años ofrecen un soporte equivalente e incluso superior al del software propietario a través de los múltiples foros que generan y del contacto directo con los desarrolladores.
- e) El uso de software libre casi siempre garantiza la conformidad con estándares abiertos, lo que en todo momento hace posible un alto nivel de portabilidad de los datos y de interoperabilidad con otros sistemas, sean de propietario o no. La adopción de estándares abiertos de representación de datos por parte de las compañías de software es una de las grandes aportaciones del movimiento del software libre y una garantía de libertad para el usuario final.

Dicho esto, habremos de admitir que la opción por el software libre no está exenta de riesgos y de inseguridades para el usuario:

- a) En primer lugar, la percepción de confianza que puede ofrecer un gran fabricante de software no es comparable a la que puede ofrecer la alternativa

del software libre. Si bien programas tales como *Mozilla* o *Open Office* se acercan mucho o incluso superan a sus respectivas alternativas de software propietario en este aspecto, es justo reconocer que en el caso de aplicaciones menos desarrolladas, el usuario puede percibir una inseguridad provocada por la falta de respaldo y de garantías de estos productos.

- b) El uso de software libre suele requerir mayores conocimientos de informática, por parte del usuario, dado que en ocasiones la documentación de los programas no tiene el nivel de calidad y claridad requerido.
- c) Del mismo modo, el soporte del producto puede ser deficiente, o incluso inexistente, si el grupo que mantenía el software desaparece, o deja de tener actividad. En estos casos, aunque los foros de usuarios sigan ofreciendo soluciones a problemas concretos, el producto no evolucionará y acabará quedando obsoleto. Debe tenerse en cuenta en este sentido que los programas de software libre se obtienen “tal cual” (*as is*), lo que significa que los promotores del programa (los responsables iniciales y los que lo mantienen) y los desarrolladores (los que han contribuido a mejorarlo) no se hacen responsables de los fallos o posibles consecuencias tales como pérdida de datos.

## 6. Indicadores de fiabilidad

La adecuación o fiabilidad de un determinado programa de software libre para las tareas específicas que el traductor profesional requiere sólo puede establecerse mediante el método de prueba y error. Sin embargo, existen algunos indicadores de fiabilidad que nos pueden ayudar a decidir qué programa puede resultar más adecuado como opción para el traductor profesional que no sólo necesita que el programa funcione, sino que le ofrezca un cierto respaldo y unas garantías de permanencia a corto y medio plazo.

En primer lugar, la filosofía de desarrollo del software libre se basa en la colaboración de un número de personas en el diseño, difusión, mantenimiento y mejora del software. Cuanto mayor sea el número de desarrolladores y usuarios, mayor potencia tendrá el proyecto y mayores serán las posibilidades de desarrollo futuro y de ayuda en caso de errores o de búsqueda de información. Estos datos se manifiestan en el nivel de actividad de foros, listas de correos, novedades, etc.

En segundo lugar, y muy relacionado con el aspecto anterior, está la madurez del proyecto. Entendemos por madurez del proyecto el tiempo que éste lleva en activo y la tendencia creciente o decreciente de actividad. Un proyecto maduro sería aquel que se ha mantenido activo un número de años y que manifiesta una tendencia creciente en el número y actividad de los usuarios y desarrolladores.

El número de descargas que tenga un programa es sin duda un dato significativo que puede indicar la popularidad de un programa pero no es en absoluto significativo

del uso efectivo que ese programa tiene, ya que, al tratarse de programas gratuitos y de libre acceso, muchos usuarios los descargan sólo para probarlos o compararlos con otras alternativas, sin que eso signifique que se utilicen en la práctica. Sin embargo, sí es significativa la cantidad y el tiempo en que se descargan pequeñas mejoras o actualizaciones del programa, que sólo son útiles para aquellos que lo utilizan realmente.

El número de versiones y la frecuencia con la que se lanzan mejoras o actualizaciones del programa es también un indicador de la actividad del grupo y del esfuerzo por mejorarlo.

En nuestro análisis hemos utilizado los datos estadísticos de *Sourceforge* para establecer la fiabilidad de los distintos gestores de memorias de traducción de software libre.

## 7. Memorias de Traducción

Si la literatura acerca del uso de memorias de traducción es escasa (cf. Dillon y Fraser, 2006), aún lo es más la referente específicamente a memorias de traducción de software libre. Existen estudios referidos a aspectos específicos del uso de gestores de memorias de traducción, referidos en ocasiones a productos concretos (Benis, 2004a, b, 2005a; Davies, 2004). Iverson (2003) investiga las consecuencias del uso de esta tecnología en la gestión de procesos documentales. Varios estudios indican que la decisión de utilizar memorias de traducción en el trabajo se ve en ocasiones más como una imposición de los clientes que como una herramienta que los traductores adoptan voluntariamente porque les ayuda a hacer mejor su trabajo (Lebtahí, 2004). Fulford y Granell-Zafra (2005) analizan las tecnologías que usan los traductores abarcando tanto las de carácter general como las más específicamente relacionadas con la traducción, sin embargo se trata sólo de un estudio exploratorio referido a traductores *freelance* en el Reino Unido. Otros estudios manifiestan una cierta reticencia por parte de un número importante de traductores a utilizar memorias de traducción en su trabajo. Höcker y Wheatley (2003) establecen en sólo un 34% el número de traductores que utilizan memorias de traducción en una muestra de traductores en el Reino Unido y Alemania. En contraste con lo que Dillon y Fraser (2006) califican de falta de entusiasmo de los traductores, los proveedores de servicios lingüísticos y las empresas de traducción sí parecen mostrar gran interés en su adopción (cf. LISA 2000 y 2004) por razones de índole económica y de racionalización del flujo de trabajo.

Desde su irrupción en el ámbito de la traducción las memorias de traducción han experimentado un claro avance hasta el punto de que hoy día se han convertido en herramienta fundamental de un gran número de profesionales de la traducción y soporte básico en el trabajo para cualquier agencia de traducción. De hecho, en los últimos años la oferta de software que de un modo u otro gestiona memorias



de traducción se ha diversificado notablemente, si bien el mercado en cuanto a los traductores profesionales se refiere, sigue utilizando mayoritariamente los gestores más difundidos (*SDL-Trados*, *Dèjà-vu*, *Wordfast*)

Respecto a la proporción de usuarios que se decanta por el uso de memorias de traducción comerciales (software propietario) frente a los que utilizan herramientas de software libre, los datos del *TM Survey* de 2006 (Lagoudaki, 2006) indican que una mayoría muy amplia utiliza gestores de propietario<sup>5</sup>. De hecho, de los que se listan en el estudio, únicamente *Omega-T* aparece con un siete por ciento de los usuarios. Podría afirmarse que en el mercado de gestores de memorias de traducción la presencia del software libre respecto al software propietario es testimonial.

El mercado está dominado por *Trados* (51%), seguido de *Wordfast* (29%) *SDLTrados 2006* (24%), *Dèjà Vu* (23%), *SDLX* (19%), *Star Transit* (19%). La adquisición de *Trados* por parte de *SDL* coloca una porción muy significativa del mercado en manos de un solo proveedor. Ante esta situación la *GALA* (Globalization and Localization Association) llevó a cabo una encuesta entre sus asociados (empresas proveedoras de servicios lingüísticos tales como traducción o localización) que reveló que el 69% de las compañías encuestadas manifestaban su preocupación por el previsible establecimiento de un monopolio en el ámbito de la localización, dándose además la circunstancia de que *SDL*, un proveedor de servicios, se convertía en propietario de la herramienta mayoritariamente utilizada por la competencia. En este mismo informe, un número importante de compañías manifiesta su deseo de poder contar con herramientas de software abierto para poder hacer frente a una situación que consideran no deseable (cf. De Palma, 2005).

A continuación analizaremos los gestores de memorias de traducción de software libre. En nuestro estudio, incluiremos únicamente aquellos que tienen presencia en *Sourceforge.net*, el mayor repositorio de software libre<sup>6</sup> que alberga aplicaciones relacionadas con cualquier ámbito de la informática. El hecho de que un determinado producto tenga presencia en *Sourceforge.net* no garantiza en absoluto su fiabilidad, ni tan siquiera su operatividad. De hecho, una de las críticas que con frecuencia se hace de *Sourceforge.net* es que un alto porcentaje de los proyectos que alberga no han tenido nunca un nivel de desarrollo suficiente para ser tenidos en cuenta como alternativas válidas. Éste es uno de los factores que con frecuencia desaniman a los potenciales usuarios de software libre que, tras encontrar un proyecto que ofrece características que se adaptan a sus necesidades, comprueban que, o bien es obsoleto, o ni siquiera es utilizable debido a la falta de mantenimiento del mismo.

5. Tampoco existen datos acerca de los pares de lenguas en los que las memorias de traducción son más empleadas, ya que las memorias de traducción en sí son independientes de las lenguas, es decir, pueden utilizarse con pares de lenguas diversas.

6. <http://www.sourceforge.net>

Nos restringiremos al software específico de gestión de memorias de traducción, lo que significa que no consideraremos aplicaciones tales como *Translate Toolkit*<sup>7</sup> que, aunque contienen módulos para la gestión de memorias de traducción, éstos están integrados en un entorno más amplio concebido para tareas de localización de software. Nuestro análisis se centrará, entre otros aspectos, en el estado actual de los diferentes proyectos, su dinamismo, su difusión, su capacidad de respuesta a posibles incidencias y la operatividad del sistema en el ámbito de la traducción profesional. A partir de ese análisis, propondremos los criterios que el traductor *freelance* debe tener en cuenta a la hora de optar por utilizar gestores de memorias de traducción de software libre y las circunstancias en las que tal decisión podría ser adecuada.

### 7.1. OmegaT

*OmegaT* es un programa de traducción asistida distribuido bajo licencia GPL para la gestión de memorias de traducción. Inicialmente creado por Keith Godfrey y difundido desde *Sourceforge* y desde su propia web<sup>8</sup>, a su desarrollo ha contribuido un grupo de programadores que han ido aportando diversas porciones de código para mejorar aspectos de su funcionamiento o corregir errores.

*OmegaT* está desarrollado en *Java* (*Sun Microsystems*), que es un entorno de programación multiplataforma, lo que significa que puede ser ejecutado en cualquier sistema operativo (*Windows*, *Linux*, *Mac OS*, etc.) sin necesidad de operaciones específicas de instalación y puede ser transportado en cualquier sistema de almacenamiento para ser ejecutado en cualquier ordenador que tenga instalado el entorno de *Java* (*Java Runtime Environment*, *JRE*) De hecho, las últimas versiones incluyen como parte de la descarga el *JRE* que se instala de forma automática lo que simplifica las tareas de ejecución.

*OmegaT* incluye las funcionalidades habituales de los gestores de memorias de traducción, tales como la búsqueda de coincidencias exactas y difusas ('fuzzy'), mantenimiento de memorias de traducción, gestión de glosarios y filtros para diferentes formatos, así como la gestión simultánea de múltiples proyectos de traducción o de múltiples memorias. *OmegaT* fue concebido originalmente para manejar documentos en el formato estándar *Open Document Format (ODF)*, generados por los programas de la suite informática de código abierto *Open Office* (*Sun Microsystems*) o *Oracle Open Office*. Puede manejar también otros formatos abiertos tales como: *XHTML* y *HTML*; *Microsoft Office 2007*, *XML*, *XLIFF* (*Okapi*), *MediaWiki*, y *Plain text*. Los documentos producidos por otras aplicaciones con formatos de propietario no abiertos como los de *Microsoft Office* anteriores a 2007 (.ppt .doc, etc.) no son

7. <http://translate.sourceforge.net>

8. <http://www.omegat.org>

compatibles con *OmegaT*, por lo que para poder trabajar con ellos habrán de ser convertidos previamente.

Para la representación de las memorias de traducción, *OmegaT* utiliza el formato estándar *TMX*, lo que garantiza la interoperabilidad con otros programas de gestión de memorias de traducción tales como *Trados*, *Wordfast*, etc. También soporta *Unicode (UTF-8)* para su uso en lenguas de caracteres no latinos.

Las primeras versiones beta de *OmegaT* datan de 2004, desde entonces se han lanzado múltiples actualizaciones y dos versiones que incorporan nuevas características o mejoran determinados aspectos. Entre las mejoras cabe mencionar la segmentación por frases ('sentence segmentation') como característica que incluyen la versión 1.5 y posteriores. *OmegaT* proporciona versiones localizadas en un número importante de lenguas. La versión actual es la 2.05. update 3.

*OmegaT* proporciona además un conjunto de herramientas para llevar a cabo diferentes operaciones en el proceso de traducción. Entre estas aplicaciones adicionales pueden destacarse *bitext2tm* (aplicación para producir memorias de traducción compatibles con *OmegaT* a partir de traducciones ya realizadas) y un conjunto de herramientas y procedimientos elaborados por Samuel Murray para llevar a cabo diferentes procesos de tratamiento de texto o de compatibilidad de *OmegaT* con otros programas.

### 7.1.1. Uso y percepciones de Omega-T

El análisis de los datos relativos a *OmegaT* revela una comunidad de usuarios consolidada y que se muestra activa tanto en lo que respecta a la detección y solución de fallos en el sistema como a la respuesta a los requerimientos de ayuda o documentación adicional por parte de algunos usuarios. En la actualidad cuenta con siete responsables del proyecto y un número importante de usuarios avanzados que informan con asiduidad acerca de errores o sugieren mejoras. Esta actividad puede verse reflejada en un número constante de actividades de seguimiento del programa. Como ejemplo de la vitalidad del grupo, el foro (<http://tech.groups.yahoo.com/group/omegat/>) tiene actualmente más de mil trescientos usuarios y experimenta un flujo constante de actividad. Lo mismo ocurre con los foros de *Sourceforge* donde se informa de los posibles fallos y se sugieren constantes mejoras en el sistema, que con frecuencia conducen a actualizaciones o nuevas versiones. El foro dedicado a usuarios de *OmegaT* registró una media de más de 100 entradas por semana en los primeros seis meses de 2008; asimismo, los foros y listas de correo del sitio web de *OmegaT* registran un alto nivel de actividad tanto en lo referente a la detección y solución de problemas como a la mejora de la documentación o a la asistencia al usuario.

Un estudio de estos foros revela que los usuarios de *OmegaT* valoran positivamente los siguientes aspectos:

- Sencillez de uso como consecuencia de la simplicidad y la falta de sofisticación de su interfaz.
- Portabilidad: al tratarse de un desarrollo Java que funciona independientemente del sistema operativo (*Windows, Linux, Mac OSX*, etc) que se use, puede ser transportada a cualquier sistema con facilidad, lo cual resulta muy conveniente cuando se utilizan entornos de traducción diferentes.
- Su compatibilidad con otros gestores de memorias de traducción (*Trados, SDL, Wordfast*, etc) a través del estándar *TMX*.
- El dinamismo del grupo de desarrolladores y de los participantes en foros, lo que hace que, en la mayoría de los casos, los problemas puedan resolverse con rapidez y eficacia.

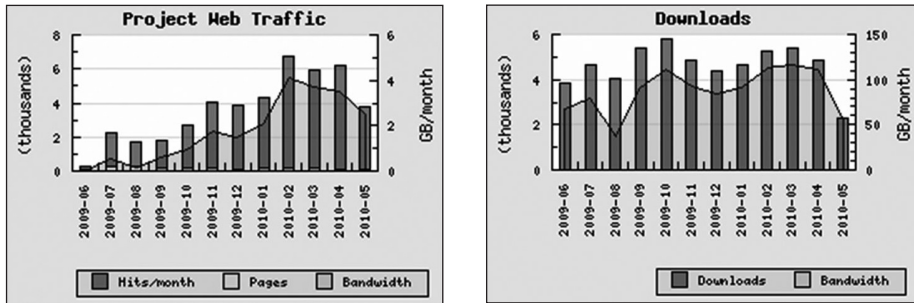
En el lado negativo se mencionan los siguientes aspectos:

- La simplicidad de las herramientas terminológicas, cuyas posibilidades están por debajo de la sofisticación que pueden ofrecer los programas de software propietario más extendidos.
- La falta de correspondencia que ocasionalmente puede producirse entre las diferentes actualizaciones y la documentación que de ellas se ofrece, dado que en ocasiones la introducción de algún cambio no siempre se ve reflejada en la documentación del programa.
- Los problemas de conversión que pueden surgir con la representación *TMX* generada por otros gestores de memorias de traducción. Estos problemas no son atribuibles a *OmegaT*, sino más bien a las conversiones que se llevan a cabo desde el lenguaje de representación utilizado por el software propietario y el estándar de *TMX*.

El siguiente gráfico<sup>9</sup> muestra la actividad del proyecto *OmegaT* desde junio de 2009 hasta mayo de 2010 y el número de descargas del programa realizadas entre esas fechas. Como puede observarse existe una regularidad tanto en el número de descargas como en el tráfico que la actividad del proyecto provoca en *Sourceforge*, que muestra una tendencia ascendente. Estos dos factores son indicativos de la consolidación del proyecto

9. Fuente: <<http://www.sourceforge.org>> [Consulta: 17-05-2010] Ha de tenerse en cuenta que la actividad en este gráfico está expresada en miles y en Gb., mientras que en otros gráficos aparece en unidades y Mb.

## Usage Statistics For OmegaT - multiplatform CAT tool



## 7.2. OmegaT+

*OmegaT+* es un proyecto que nace en 2005 tras haber surgido discrepancias, en el seno de *OmegaT*, relacionadas con las distintas concepciones del proyecto y especialmente con la atribución de autoría de ciertas secciones del código. Raimond Martin, que hasta entonces había contribuido a *OmegaT* con porciones de código y sobretodo con el programa *bitex2tm*, decide emprender un proyecto independiente basado en *OmegaT*, pero con una filosofía distinta de desarrollo y con el objetivo de mejorar la aplicación e integrar en ella ciertas funcionalidades que no estaban presentes en *OmegaT*. La primera versión de *OmegaT+* (1.0 Milestone 1. M1) se lanza en marzo de 2008 y desde entonces se han lanzado seis actualizaciones con diferentes mejoras o que incorporan funcionalidades nuevas, la última (*OmegaT+* 1.0 M3) se comenzó a distribuir en mayo de 2010.

La bifurcación del proyecto *OmegaT* es representativa del tipo de situaciones que puede producirse en el desarrollo de aplicaciones de software cuando existen divergencias significativas respecto a la idea y las líneas de actuación a seguir.

El nacimiento de *OmegaT+* no estuvo exento de polémicas entre los desarrolladores, que dieron lugar a no poca confusión entre los usuarios que se preguntaban qué diferencia existía entre ambos productos. Con frecuencia, la respuesta a esa pregunta derivaba en agria discusión de los desarrolladores de ambos proyectos acerca de la legalidad del uso de la denominación *OmegaT+*<sup>10</sup>. La propia utilización del nombre *OmegaT+* ha sido también objeto de críticas que han llevado a Raimond Martin a anunciar recientemente un posible cambio de denominación del proyecto a *OmegaTx*<sup>11</sup>.

10. Como ejemplo de uno de estos intercambios, véase [http://www.proz.com/forum/omegat\\_support/67948-omegat\\_vs\\_omegat%2B.html](http://www.proz.com/forum/omegat_support/67948-omegat_vs_omegat%2B.html)

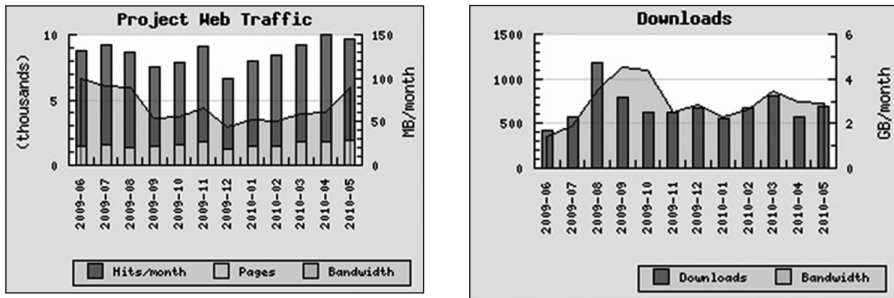
11. Cf. [http://groups.google.com/group/omegatplus/browse\\_thread/thread/ce45d40cc11923a0#](http://groups.google.com/group/omegatplus/browse_thread/thread/ce45d40cc11923a0#)

Desde la primera versión, *OmegaT+* ha intentado desarrollarse de forma independiente, si bien desde el punto de vista del usuario final son escasas las diferencias con *OmegaT*. El nuevo proyecto no ha logrado incorporar desarrolladores ni un grupo de usuarios suficientemente amplio como para hacerlo preferible al producto original, que da la impresión de estar más consolidado y ofrecer un mayor nivel de fiabilidad debido a su mayor índice de actividad.

Sería deseable que ambos proyectos confluyeran en sus concepciones sobre las líneas futuras de desarrollo del producto, pero esa colaboración no parece posible a juzgar por el nivel de enfrentamiento al que han llegado los desarrolladores<sup>12</sup>.

El siguiente gráfico muestra la actividad de *OmegaT+* en Sourceforge en los últimos doce meses.

Usage Statistics For OmegaT+ CAT Tools



Como puede observarse el nivel de actividad, tanto en descargas (alrededor de 500 descargas al mes, frente a las 4.000 de *OmegaT*) como en tráfico (alrededor de 150Mb frente a los 3 Gb de *OmegaT*) es indicativo del nivel de consolidación y, por tanto, de fiabilidad de *OmegaT*.

La comparación de rangos (el lugar que ocupa el proyecto en cuanto a su nivel de actividad en *Sourceforge*) se revela muy favorable a *OmegaT* (ver tabla comparativa).

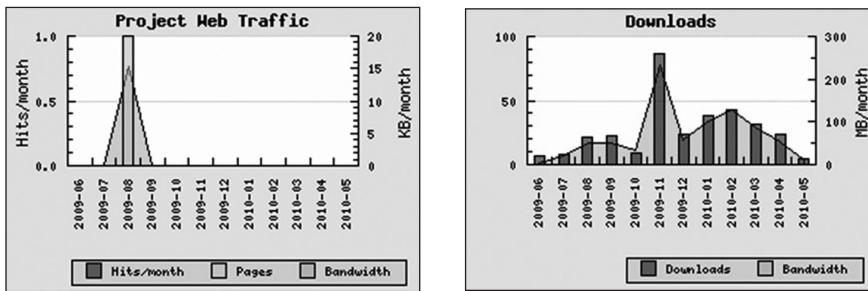
### 7.3 Frankenstein

De acuerdo con la escasa documentación que se ofrece en *Sourceforge*, *Frankenstein* es un proyecto centrado en el desarrollo de un conjunto de herramientas para la gestión de memorias de traducción, escrito en *Java* y bajo licencia *GPL*. Su creador y único responsable, Hans Sussenburger, lo lanzó en octubre de 2000 y

12. Cf. [http://groups.google.com/group/omegatplus/browse\\_thread/thread/156891bff047ab8b](http://groups.google.com/group/omegatplus/browse_thread/thread/156891bff047ab8b)

mantuvo desarrollos y actualizaciones hasta el 2003, año en que cesó toda actividad aparente en torno al proyecto<sup>13</sup>. El siguiente gráfico muestra el nivel de actividad desde mayo de 2009 hasta mayo de 2010.

Usage Statistics For Frankenstein translation Memory Suite



Más significativo aún que el escaso número de descargas es la absoluta falta de actividades de seguimiento (informes de errores, actividades de mantenimiento, etc.) desde 2004<sup>14</sup>.

#### 7.4. ForeignDesk

*ForeignDesk* es un ejemplo de proyecto con orígenes y objetivos radicalmente diferentes de los que subyacen en los proyectos ya mencionados. *ForeignDesk* es una marca registrada de *Lionbridge*, empresa de servicios lingüísticos establecida en 1966 en Walton (Mass). *Lionbridge* decidió en el año 2001, como estrategia de mercado de la compañía hacer público el código fuente de *ForeignDesk*, que hasta entonces había sido software propietario, e integrarlo como proyecto de software abierto en Sourceforge.net. Hasta marzo de 2003 se lanzaron un pequeño número de actualizaciones fundamentalmente dirigidas a la solución de problemas y a la compatibilidad con *TMX*. *ForeignDesk* se proponía como una herramienta de traducción integrada desarrollada en *C++* y *Visual Basic*, para ser utilizada en entornos *Windows*.

En abril de 2002 la página web de *Lionbridge* publicaba un artículo acerca de la enorme aceptación que esta estrategia había tenido entre los usuarios, lo que se acreditaba con el alto número de descargas que se habían producido hasta ese momento. En ese artículo, Jonathan Clark, responsable de desarrollo tecnológico de *Lionbridge* afirmaba:

13. Cf. <http://www.sourceforge.net>

14. Cf. <http://sourceforge.net/project/stats>

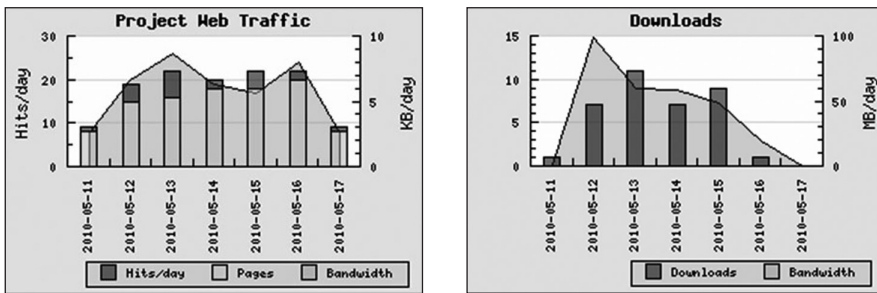
*“That ForeignDesk has received more than ten thousand downloads in such a short period of time underscores our belief that open-sourcing ForeignDesk was the right move for our customers and industry, (...). Clearly ForeignDesk is opening up tremendous opportunities for customers that consider translation memory tools cost prohibitive. It is not only benefiting the development community in the globalization sector but also in related industries such as content management where translation memory can be integrated into new and existing applications.”*<sup>15</sup>

Sin embargo, sorprendentemente, la compañía no parecía tener intención de continuar desarrollando el proyecto y, al no lograr captar el interés de desarrolladores independientes, quedó estancado, a pesar de los requerimientos de usuarios, uno de los cuales afirmaba en el foro de *Sourceforge.net* en agosto de 2006:

*“It is sad that there is no developing activity going on under this project! Too good software, with a huge potential, to leave it like that! It can be a real alternative to Passolo or even Catalyst if only upgraded a little bit”*.<sup>16</sup>

El siguiente gráfico<sup>17</sup> representa la actividad registrada en el proyecto en los últimos 12 meses.

#### Usage Statistics For ForeignDesk



Es significativo, sin embargo, y manifiesta el interés que existe en la comunidad de usuarios por los programas de memorias de traducción de software abierto que las descargas del programa siguen produciéndose aún a pesar de que la última versión que se ofrece hasta el día de hoy es la versión beta, que data de mayo de

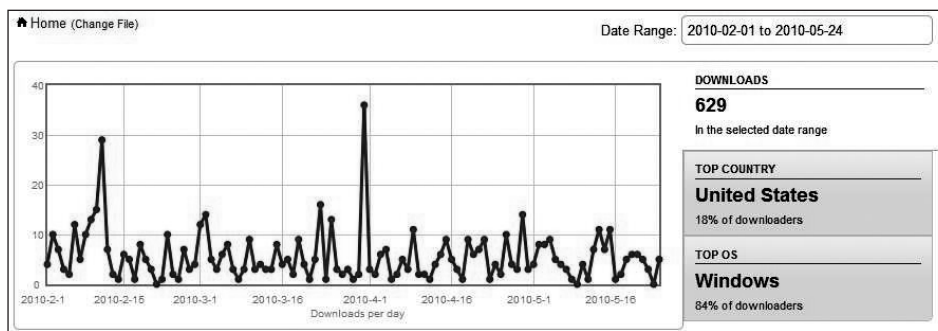
15. En <http://www.lionbridge.com/lionbridge/en-US/company/news/lionbridge-foreigndesk-is-rapidly-embraced-by-the-open-source-community-and-end-users.htm> [Consulta: 20-4-2010]

16. En [http://sourceforge.net/forum/forum.php?thread\\_id=1552581&forum\\_id=120442](http://sourceforge.net/forum/forum.php?thread_id=1552581&forum_id=120442) [Consulta: 20-4-2010]

17. [www.sourceforge.org](http://www.sourceforge.org)



2002 y que no ha sido actualizada desde entonces. El gráfico siguiente representa las descargas del programa desde febrero a mayo de 2010<sup>18</sup>.



*ForeignDesk* es un caso singular, aunque no único, de la utilización con fines exclusivamente comerciales de los principios que sustentan el movimiento del software libre.

### 7.7. Tabla comparativa de gestores de memorias de traducción de software libre<sup>19</sup>.

	<i>OmegaT</i>	<i>OmegaT+</i>	<i>Frankenstein</i>	<i>ForeignDesk</i>
Rango en <i>Sourceforge</i> (descargas: mayo 2010)	881	2.562	56.137	10.946
Fecha de inicio	2002	2005	2003	2002
Nº de versiones	2	1	1	1
Nº de responsables	8	1	1	2
Descargas: mayo 2009-mayo 2010	56.000	8.059	321	2.134
Total histórico de descargas	205.800	20.638	2.118	47.542
Media de descargas semanales.	1.089	170	4	37
Nº mensajes recibidos/Nº mensajes atendidos en tracker*	168/73	6/4	0/0	0/0
Foros externos <i>Sourceforge</i> **	1 <sup>+</sup>	1 <sup>++</sup>	0	0
Nº miembros foros externos.	1.328	67	-	-
Nº mensajes en foro: mayo 2009-mayo2010.	3.245	94	-	-

(Cont.)

18. [www.sourceforge.org](http://www.sourceforge.org)

19. Cf.: [www.sourceforge.org](http://www.sourceforge.org). [Consulta 20/05/2010]

	<i>OmegaT</i>	<i>OmegaT+</i>	<i>Frankenstein</i>	<i>ForeignDesk</i>
Nº de listas de correo en <i>Sourceforge</i>	4	1	0	2
Última actualización.	Mayo 2010	Mayo 2010	Octubre 2003	Mayo 2002
Localización nº lenguas	29	7	1	1
Plataformas	All (Java)	All (Java)	All (Java)	Windows

\*número de mensajes en el tracker del proyecto en *Sourceforge* que han sido atendidos (respondidos, o asignados para solucionar desde enero 2010 a mayo 2010.)

\*\* Existencia de foros externos a *Sourceforge*.

+ <http://tech.groups.yahoo.com/group/OmegaT/>

++ <http://groups.google.com/group/omegatplus>

## 8. Conclusiones

En el presente trabajo hemos hecho una evaluación de las memorias de traducción de software libre desde el punto de vista de su consolidación como herramientas fiables para el traductor profesional. Para ello hemos establecido una serie de parámetros y criterios, con arreglo a los cuales hemos valorado los proyectos presentes en *Sourceforge.net.*, con la finalidad de ofrecer al traductor una ayuda en el caso de que desee optar por la utilización de software libre en su trabajo. Esta opción no debe considerarse ni mejor ni peor que la opción por el software propietario, y no debe tomarse sólo en función de criterios económicos, aun siendo éstos importantes. La decisión habrá de tomarse considerando un amplio abanico de factores relacionados con las preferencias o circunstancias personales del traductor, con su nivel de conocimientos técnicos y con la naturaleza del trabajo que desarrolle en la práctica.

Como resultado de nuestro análisis, podemos afirmar que de los múltiples proyectos de software libre que se han ido generando en el ámbito de los gestores de memorias de traducción, sólo *OmegaT* ha logrado establecerse y conseguir una presencia prolongada como herramienta profesional para el traductor. De entre los demás, *ForeignDesk* y *Frankenstein* han sido abandonados progresivamente o han desaparecido porque sus autores han dejado de mantenerlos o porque no han logrado suscitar el interés de un número suficiente de desarrolladores que se implicasen en el proyecto y *OmegaT+* no parece haber suscitado el interés de los potenciales usuarios y presenta serias dudas respecto a los objetivos que persigue, las motivaciones que lo inspiran y las ventajas reales que puede ofrecer frente a *OmegaT*. Desde un punto de vista práctico, esto significa que el traductor profesional que desee utilizar este tipo de herramientas habrá de ser consciente de las consecuencias que para su trabajo puede tener una eventual cancelación del mantenimiento o actualizaciones del programa que utiliza. Esto no debería representar un problema desde el punto de vista de los datos, puesto que la compatibilidad con TMX garantizaría la importación de esos datos desde cualquier otro sistema, pero sí acarrearía las dificultades propias de la migración a un nuevo programa.

Cada producto debe ser evaluado cuidadosamente, ya que la presencia de éste en *Sourceforge.net* o en cualquier otro repositorio de programas de software libre no garantiza su operatividad ni su fiabilidad. Es aconsejable analizar a fondo los datos disponibles sobre el proyecto y realizar las pruebas necesarias para garantizar que el producto que se elige se ajusta a los requerimientos reales del usuario.

No es tarea fácil dilucidar si un determinado programa o suite de programas responderá satisfactoriamente a las necesidades del traductor profesional. Pero a la hora de elegir, si se opta por la alternativa válida del software libre, deben tenerse en cuenta los siguientes criterios:

- a) El índice de actividad del proyecto, que queda patente en la cantidad de canales de participación abiertos para los usuarios y desarrolladores, la frecuencia con que éstos utilizan esos canales y la prontitud con que se atienden demandas de ayuda o informes de fallos.
- b) La cantidad de desarrolladores y usuarios que participan en foros, listas de correo, etc.
- c) El historial de desarrollo del proyecto que se manifiesta en el lanzamiento de nuevas versiones o actualizaciones. Mientras más frecuentemente se produzcan esos lanzamientos, mayor índice de madurez tendrá el proyecto y menores posibilidades de caer en el abandono.

## Bibliografía

- Benis, M. (2004a). The leader's progress: Review of TradosTM 6.5.5 and TradosTM Freelance. *ITI Bull* (Sept/Oct), 29–32.
- (2004b). A mixed bag of tricks. *ITI Bull* (Nov/Dec), 26–32.
- (2005a). SDLXTM: The dark horse of translation memory. *ITI Bull* (Mar/Apr): 22–27
- De Palma, D. A. (2005). *SDL Trados. Language Service Provider Reaction to SDL Purchase of Trados*. Andover (MA): Globalization and Localization Association.
- Dillon, S. y Fraser, J. (2006). Translators and TM: An investigation of translator's perceptions of translation memory adoption. *Machine Translation* 20, 67-79.
- Fulford, H. y Granell-Zafra, J. (2005). Translation and technology: A study of UK freelance translators. *JoSTrans* 4, 2–17.
- Fogel, K. (2009). *Producing Open Source Software: How to Run a Successful Free Software Project*. Florida: University Press of Florida.
- Höcker, M. y Wheatley, A. (2003) *Translation Memory Survey. 2003*. En <[http://ecolore.leeds.ac.uk/downloads/2003.05\\_bdue\\_survey\\_analysis.doc](http://ecolore.leeds.ac.uk/downloads/2003.05_bdue_survey_analysis.doc)>. [Consulta: 03/11/2008]

- Iverson, S. (2003). Working with translation memory: When and how to use TM for a successful translation project. *MultiLingual Comput Technol* 14(7), 35–37.
- Lagoudaki, E. (2006). *TM Survey: Translation Memory Systems, enlightening users' perspective*. Imperial College London. En <<http://www.atril.com/docs/tm-survey.pdf>> [Consulta: 10-1-2010].
- Lebtahi, I. J. (2004). Traducteurs dans la société de l'information: Évolutions et interdépendances. *Meta* 49, 221–235.
- LISA (2002) *Translation memory survey 2002: Translation memory usage and trends*. Localisation Industry Standards Association (LISA), Romainmôtier, Switzerland. En <http://www.lisa.org/products/surveys/tm02survey.html>> [Consulta: 03/11/2008].
- (2004) *Translation memory survey 2004: Translation memory usage and trends*. Localisation Industry Standards Association (LISA), Romainmôtier, Switzerland. En <<http://www.lisa.org/products/surveys/tm02survey.html>> [Consulta: 03/11/2008].
- Stallman, R. (1998). The GNU Project. A Brief History of Free/Open Source Software Movement. En <<http://www.cisnet.com/glennmcc/foss/brief-open-source-history.html>> [Consulta: 15/11/2007].