



La traducción automática en las empresas de traducción

Juan José Arevalillo Doval
Hermes Traducciones y Servicios Lingüísticos, SL
juanjo.arevalillo@hermestrans.com

RESUMEN

La traducción automática tiene una presencia creciente en los procesos de traducción de muchos clientes y con ella la posesición de textos traducidos con motores de traducción. Esto supone nuevos retos en la adecuación de esta realidad a los procedimientos existentes y en la elaboración del perfil profesional de poseedor.

Palabras clave: traducción automática, posesición, empresas, motores, poseedor, TAO

RESUM (*La traducció automàtica a les empreses de traducció*)

La traducció automàtica té una presència creixent en els processos de traducció de molts clients i amb ella la postedició de textos traduïts amb motors de traducció. Això suposa nous reptes en l'adequació d'aquesta realitat als procediments existents i en l'elaboració del perfil professional de posteditor.

Paraules clau: traducció automàtica, postedició, empreses, motors, potedició, TAO

ABSTRACT (*Machine translation in translation companies*)

Machine translation is more and more present in many customers' translation processes, together with post-editing of texts translated with translation engines. This involves new challenges in fitting this reality to the existing procedures and the professional profile creation of post-editors.

Keywords: translation machine, post-editing, business, engines, post-editor, CAT



Introducción

La traducción ha sufrido en los últimos 20 años cambios significativos gracias a la aplicación de la tecnología, aunque no siempre con el ritmo que a todos en el sector les habría gustado. De ser un campo yermo en ayudas para el traductor que no fueran los diccionarios, bibliotecas o máquinas de escribir, desde la aparición del ordenador doméstico y sus primeros programas de procesamiento de textos se han venido sucediendo cambios muy significativos que han perfilado la figura y los procedimientos del traductor hasta puntos casi inimaginables en aquellos momentos. En este artículo se intentará explicar parte de esa evolución centrándose en la traducción automática aplicada a los procesos de traducción en conjunción con las memorias de traducción, y la presencia cada vez mayor de la posesición como servicio adicional en las empresas de traducción.

Desarrollo

A finales de los años 80 del siglo pasado —suenan lejano pero no hace tanto...— los principales fabricantes de informática lanzaban al mercado sus equipos y programas con gran profusión de actualizaciones de sus versiones, con lo que movían significativas cantidades de documentación traducida a muchos idiomas. En numerosas ocasiones la documentación entre una versión nueva y su precedente cambiaba muy poco, pero no era posible reutilizar el material traducido de la versión anterior con la suficiente fluidez. De ahí que dichos fabricantes hicieran patente la necesidad de desarrollar programas internos que facilitaran ese reaprovechamiento. Fruto de ese esfuerzo, el gigante IBM llegó a desarrollar un programa que usaba internamente para tal fin y que con el tiempo comercializó en el sector de la traducción con el nombre de IBM TranslationManager/2 (IBM TM/2) que se ejecutaba en la plataforma propietaria OS/2. El sector de la traducción no lo sabía, pero estaba asistiendo a un punto de inflexión que marcaría la forma de trabajar de los futuros traductores: las memorias de traducción.

En efecto, este programa era el primer esfuerzo serio por organizar el material reutilizable en una memoria de traducción; un concepto que en principio costó entender para muchos y que caló con muchas dudas y dificultades, pero que todo el sector fue reconociendo como algo propio con el tiempo. Se trataba de la primera aplicación tecnológica desarrollada por traductores y para traductores con un alto grado de operatividad; es decir, una herramienta solo aplicable al sector de la traducción. No debe olvidarse que había otros intentos pero que adolecían de un potencial insuficiente; por el contrario, IBM TM/2 sí tenía detrás el apoyo de un gigante de la informática para cubrir sus necesidades internas en primer lugar y para comercializarlo después.

Hasta este punto todos estaban de acuerdo en que era una clara mejora en las ayudas para el traductor —entre otras cosas porque no había prácticamente ninguna ayuda específica—, ya que lo más tecnológico que empleaban los traductores se limitaba a procesadores de texto cuyo uso y curva de aprendizaje, en la mayoría de las ocasiones, no era la aconsejable; no obstante, la nueva aplicación traía consigo una vertiente envenenada: la aplicación de descuentos por reutilización del material traducido; algo a lo que se aferraron con suma rapidez los clientes de los servicios de traducción... En este sector el uso de herramientas específicas de traducción tiene un trasfondo que no se produce en otros sectores: pensemos en cualquiera de ellos y constataremos que, si una empresa invierte un buen dinero en mejorar sus herramientas, que pueden permitirle una mayor productividad, rapidez e incluso calidad en la entrega, por lo general la tarifa de esa empresa es superior a la de otras que no han desembolsado una inversión similar. Por el contrario, si un traductor o una empresa invierte en tecnología específica se va a encontrar con que, siempre que aplique esas herramientas en la forma debida, podrá entregar sus proyectos con mayor rapidez, mayor coherencia terminológica y —¿por qué no?— mayor calidad, pero se va a encontrar con que el cliente le va a solicitar descuentos por tramos de reutilización del



material existente, lo que arrojará unos ingresos teóricamente menores en lo inmediato, pero no necesariamente en el medio o largo plazo...

Esta particularidad caló a gran velocidad en el sector. Corrieron ríos de tinta en foros y ciberforos, y prácticamente todos los profesionales estábamos de acuerdo en no aplicar esos descuentos: el sector era un clamor. Sin embargo, la realidad del mercado se fue imponiendo poco a poco, y ese furor inicial se fue suavizando hasta casi desaparecer: o los profesionales se adecuaban a estas nuevas prácticas impuestas por los clientes, o se podía perder a esos clientes. Con toda probabilidad esta fue la razón que nos movió a muchos a pasarnos de unas posturas inicialmente radicales; pero en mi opinión también tuvo que ver mucho la posibilidad de incrementar la productividad que los programas de TAO permitían al poder abarcar y procesar un número mayor de palabras por día si se daban las circunstancias correctas de reutilización y personalización de las memorias de traducción: como se explicaba antes, una ventaja relativa a medio plazo.

Como bien comentaba Rubén Rodríguez: «Translation Tools are not developed for translators, but for translation clients... That doesn't mean translators cannot take advantage of them» (Rodríguez, 2011), las memorias de traducción surgieron de las urgencias sufridas por los fabricantes del sector informático para lanzar la documentación que acompañaba a sus productos en el mercado, pero al mismo tiempo se dieron cuenta de que generaba un subyacente ahorro de costes, y esa herencia pecuniaria ha venido acompañando a estos programas en todo momento. Pero es de extrema importancia que los usuarios sepan beneficiarse también de las ventajas que encierran, sobre todo relacionadas con la personalización para clientes concretos y el mantenimiento correcto de esas memorias. Nadie duda de que el control directo de las memorias, sin combinarlas con otras memorias de mil y una procedencias y de calidad dispar, pueda permitir una mayor productividad y compensar económicamente los descuentos impuestos por los clientes o negociados con ellos en el mejor de los casos.

Este fue el panorama inicial y el aceptado hasta hoy de las memorias de traducción. Una vez ya bien asentada esta tecnología, existía otra que sigilosamente se asomaba al sector de la traducción: la traducción automática.

No es el objetivo de este artículo explicar cuáles fueron los inicios ni la historia de la traducción automática, sino encuadrarla en el ámbito del sector profesional. Las pautas de introducción en el sector son muy parecidas a las ya mencionadas, así como los sentimientos que trajo consigo, aunque mucho más radicales y viscerales como era de esperar. A nadie se le escapa que entre la traducción asistida y la automática existen relaciones y diferencias, pero hay una característica práctica que marca la frontera entre ambas: la traducción asistida busca reaprovechar materiales validados y automatizar las tareas repetitivas basadas en dicho reaprovechamiento; la traducción automática, en cambio, busca sustituir al traductor humano. Es una definición también muy radical, pero nos sitúa sin dudas en el terreno de juego. Este concepto ha marcado muy negativamente a la traducción automática, lo que ha reforzado posiciones muy defensivas en el sector, puesto que se la viene interpretando como el destructor del traductor humano. Pueden existir clientes potenciales desconocedores de la realidad profesional que piensen así, pero nadie duda de que esa sustitución es una falacia: el traductor humano siempre va a existir, porque en la traducción entran en juego unos procesos mentales, gramaticales, semánticos y contextuales difícilmente interpretables al 100 % por una máquina, o al menos con una garantía mínima de fiabilidad.

Un argumento habitual dentro de los foros de traductores radica en que la traducción automática quitará mucho trabajo a los traductores. Desde un examen superficial así puede parecer, pero si nos tomamos la molestia de analizar los datos del sector no solo no es así, sino que parece ser que el uso de la traducción automática da lugar a que se acometan proyectos de traducción que antes no eran posibles por los costes o cualesquiera otras razones. Echemos un vistazo a los datos fríos...



Common Sense Advisory en su informe *The Top 100 Language Service Provider* correspondiente al año 2012, que a su vez es un resumen de su *The Language Services Market: 2012*, presenta unos datos de facturación del sector de la traducción que, lejos de menguar, van creciendo progresivamente año tras año: de una facturación mensual en el 2011 de 29 885 millones de dólares, el año 2012 registrará una facturación próxima a los 33 523 millones (Common Sense Advisory, 2012, 7). Incluso su predicción para el año 2015 es de 47 317 millones de dólares. Este informe es la principal fuente estadística del sector mundial de la traducción y sus cálculos de año a año gozan de una enorme fiabilidad por la amplitud de sus datos, por sus fuentes y por sus técnicas de elaboración extensamente explicadas en el prólogo de los mencionados informes. Si nos atenemos a estas cifras, la teoría de que la traducción automática supone un menor trabajo para el sector se cae por su propio peso. Podrán aportar otros problemas al traductor profesional, pero no ese precisamente, porque los números hablan por sí solos.

Sin embargo, sí estoy de acuerdo en que la traducción automática no es la panacea como algunos clientes creen: solo si se aplica de modo controlado a determinados tipos de textos y con unos diccionarios en continuo mantenimiento. En un principio la traducción automática llevó una vida independiente de la traducción asistida, pero en los últimos años se ha producido una sinergia entre ambas y un gran número de clientes utilizan la primera como un complemento de la segunda —no como el eje principal de la traducción— con un porcentaje de aproximación determinado, y únicamente aplicable cuando en la memoria de traducción asistida no se encuentran segmentos válidos o aprovechables. Esta técnica de integración no es nueva, pues ya viene utilizándose desde hace años: la Dirección General de Traducción (DGT) de la Unión Europea lleva años utilizando este método en sus gigantescas memorias, en concreto en combinaciones entre idiomas romances (sobre todo, francés y español) en textos jurídicos. Asimismo, la posesición es un servicio que la DGT viene solicitando a sus proveedores de servicios de traducción (PST) también desde hace años.

Efectivamente, la posesición es el otro caballo de batalla de la traducción automática. De hecho, es un servicio que habitualmente muchos clientes piden a los PST desde hará unos tres o cuatro años. Tanta importancia ha cobrado este servicio, que existen una gran demanda de planes de formación para poseedores, un perfil cada vez más exigido en el sector, así como directrices que marquen la evaluación de este servicio: «The translation industry is adopting machine translation (MT) in increasing numbers. Yet a prerequisite for efficient adoption, the evaluation of MT output quality, remains a major challenge for all». (TAUS, 2012, 3)

Pero volvamos a comparar traducción asistida con la automática: la primera podía ser asumida sin ningún problema por traductores y empresas de traducción sin otros esfuerzos que no fuera la familiarización con la herramienta correspondiente. Sin embargo, la traducción automática no presenta un escenario tan sencillo: la producción de un sistema eficaz requiere una inversión en tiempo y dinero, lo que supone una barrera difícilmente franqueable para un traductor autónomo, puesto que precisa su tiempo para producir traducciones en su mayor parte. La dedicación para generar un motor de traducción de cierta fiabilidad supone el trabajo de una persona como mínimo con plena dedicación durante mucho tiempo, además del desembolso inicial para la adquisición de programas comerciales, cuya configuración tampoco es fácil... Incluso solo resulta asumible para empresas con cierta estructura. No obstante, la aparición de programas abiertos como Moses o Apertium ha dado un vuelco significativo a esta situación, puesto que ha permitido que incluso empresas con menores recursos dediquen parte de su tiempo a la familiarización con estos programas para preparar sus motores. También pueden optar a esta solución empresas especializadas que preparen estos motores a precios más competitivos. Sea cual sea la solución elegida, es innegable que deberá encomendarse el proyecto del motor como mínimo a una persona con los conocimientos debidos y con el tiempo —y tranquilidad— lógico para cosechar frutos en este proceso. Pero la historia no acaba ahí, ya que los diccionarios y la operatividad de estos sistemas requiere un mantenimiento y actualización continuos, con lo que la dedicación plena



sigue vigente incluso después de la generación de los motores en cuestión. Por si fuera poco, a esto hay que añadir los problemas que pueden dar los formatos diversos con que se presentan las traducciones y su compatibilidad con las etiquetas con que se gestionan estos formatos en los diversos programas: una compatibilidad que puede ser fuente de problemas en la traducción asistida, pero que se puede convertir en una pesadilla en la traducción automática.

A pesar de ello, una inversión inteligente empleada en conjunción con los sistemas existentes de traducción asistida puede suponer la disponibilidad de un motor fiable, que deberá completarse con un uso adecuado —y no universal— de esta tecnología y con la formación debida para los poseedores, una exigencia creciente de muchos profesionales. Se trata de un tipo de servicio que está aquí para quedarse, y por mucho tiempo. Son cada vez mayores los volúmenes de traducción recibidos, a pesar de los funestos presagios de algunas voces del sector.

Volvamos, no obstante, a la posesición. Como decía, la posesición lleva una vida paralela a la de la traducción automática, y en este contexto las empresas se enfrentan a dos posibilidades: en primer lugar, la traducción automática de textos sin más, que se pasan a los poseedores; en segundo lugar, la producción de traducciones mediante una combinación de la traducción asistida y la automática. Como resulta obvio, ambas requieren distintas estrategias. En el primer caso, la productividad del poseedor puede variar significativamente de un proyecto a otro en función del grado de desarrollo, actualización y especialización del motor empleado por el cliente. Para ello es preciso hacer un análisis inicial del material recibido, porque en muchos casos requiere su retraducción. Para esa valoración, las empresas desarrollan sus propios sistemas de medición del factor de distancia: la diferencia porcentual de cambios entre el texto inicial y el poseitado; factor este que se utiliza para calibrar el esfuerzo de posesición en tiempo y dinero. En el segundo caso, la integración de segmentos de traducción automática en un texto traducido con memorias de traducción facilita la labor, ya que suele tratarse de sistemas más desarrollados y depurados, por lo que el índice de aprovechamiento suele ser mayor.

Los dos parámetros básicos para valorar, en cualquier caso, un texto poseitable son en términos básicos la calidad a la que puede llegar el motor y la calidad final que espera el cliente: publicable o funcional para un fin determinado (TAUS, 2010, 2), si bien entre ambas puede haber una amplia escala que depende de los diversos conceptos de calidad que tienen los distintos clientes, y que el poseedor debe conocer con su experiencia o mediante las indicaciones específicas de los clientes, que pueden variar sensiblemente de uno a otro. Para más información, véase <http://www.translationautomation.com>.

Todo este proceso hace necesaria la presencia de un perfil cada vez más especializado: el poseedor, cuyo plan de formación no es estándar y se ajusta a las necesidades de cada empresa, e incluso a las de un cliente concreto. Si bien un traductor y un revisor pueden optar perfectamente a este perfil, tiene características específicas que lo diferencian, ya que es conveniente que tenga conocimientos que le permitan desarrollar macros que automaticen determinados cambios o que afecten a estructuras gramaticales que vayan surgiendo, que luego deben traspasarse al motor para que vaya aprendiendo... En el caso de que el motor sea externo, muchos clientes están abiertos a ese proceso de colaboración entre cliente y PST que permita refinar el estado de forma de ese motor, que requiere una actualización continua de sus diccionarios. Todo esto supone una carga mayor si cabe en los procesos de las empresas, que deben saber manejar. De hecho, en la actualidad son las propias empresas las que van formando a los poseedores de un modo eminentemente práctico y en ocasiones poco planificado por la propia variabilidad de la calidad del material poseitable. Por lo tanto, en el perfil del poseedor son bienvenidos conocimientos que le permitan ir más allá del simple proceso traductológico o de revisión, ya que tendrá que lidiar con cuestiones de ajuste técnico relacionadas con macros y estructuras de diccionarios compatibles con los motores que se usen. Es de esperar que esta exigencia redunde en un esfuerzo convergente



del sector para establecer una formación más o menos estándar a este respecto, así como unos sistemas de valoración comunes al estilo del LISA QA Model o del SAE J2450.

Conclusión

De lo anteriormente dicho, no queda lugar para la duda sobre la presencia cada vez mayor de la traducción automática y de la producción de textos para posesición en el sector. Los clientes exigen este servicio con mayor frecuencia y en proporción creciente, en paralelo con los servicios de la traducción tradicional. Las empresas de traducción están ajustando los perfiles de sus revisores y traductores a las tareas de posesición, y también remodelan sus procedimientos para dar cabida a la nueva tecnología en sinergia con la existente y para poder calibrar la calidad de su producción de textos traducidos con motores, como, por ejemplo, el factor de distancia, que mide los niveles de cambio desde el texto traducido automáticamente hasta el final posesitado.

El contexto de la traducción automática es algo habitual en la actualidad, pero no debe olvidarse que viene de muy atrás. Common Sense Advisory ya avisaba de lo siguiente en el 2006: «Our mantra: MT is a substitute for zero translation. We figure that 99.44% of information inside an organization never, ever gets translated [...] MT will not replace human translation but complement it» (Common Sense Advisory, 2006). Buenos conocedores del sector, ya aseguraban que había muchos falsos supuestos relacionados con la traducción automática y que entonces era un mercado pequeño pero con un gran potencial y sentenciaba lo siguiente: «Deployment success depends on doing the homework rather than defaulting to leaders» (Common Sense Advisory, 2006). Toda una premonición del panorama actual.

Bibliografía

- Common Sense Advisory (2012). The Language Services Market: 2012, CSA.
- Common Sense Advisory (2012). The Top 100 Language Service Providers, CSA.
- Common Sense Advisory (2006). Presentación «Evaluating MT Prior to Deployment», en 2006 Conference of the Association for Machine Translation in the Americas.
- Rodríguez, R. (2011). «Machine Translation DiY: A brief guide for translators to benefit from MT instead of being exploited by it», en ATA Annual Conference 2011.
- TAUS (2012). Advancing Best Practices in Machine Translation Evaluation, TAUS.
- TAUS (2010). TAUS-CNGL Machine Translation Postediting Guidelines. TAUS.
- TAUS. <http://www.translationautomation.com/guidelines/machine-translation-post-editing-guidelines>. Fecha de última actualización: 2010. Fecha de consulta de la página: 10.09.2012.