

Nuevas tecnologías en traducción: generación automática de resúmenes y el proceso de documentación

María Cristina TOLEDO BÁEZ
Universidad de Málaga

Como citar este artículo:

TOLEDO BÁEZ, María Cristina (2008) «Nuevas tecnologías en traducción: generación automática de resúmenes y el proceso de documentación», en PEGENAUTE, L.; DECESARIS, J.; TRICÁS, M. y BERNAL, E. [eds.] *Actas del III Congreso Internacional de la Asociación Ibérica de Estudios de Traducción e Interpretación. La traducción del futuro: mediación lingüística y cultural en el siglo XXI. Barcelona 22-24 de marzo de 2007*. Barcelona: PPU. Vol. n.º 2, pp. 345-358. ISBN 978-84-477-1027-0. Versión electrónica disponible en la web de la AIETI:
<http://www.aieti.eu/pubs/actas/III/AIETI_3_MCTB_Tecnologias.pdf>.



Nuevas tecnologías en traducción: generación automática de resúmenes y el proceso de documentación

María Cristina Toledo Báez
Universidad de Málaga

Introducción

Tal y como señala Lavid (2005), vivimos en la era de la información, en la «tercera ola», parafraseando el título del libro de Alvin Toffler (1996), que alude a la revolución de la información y del conocimiento. De esta forma, es posible señalar que la sociedad de este nuevo milenio es la sociedad de la información, caracterizada por la introducción de las tecnologías de la información y la comunicación (TIC) en todos los ámbitos de nuestras vidas.

En el marco de las TIC, Internet se presenta como una herramienta de trabajo imprescindible para el traductor, pues le ofrece de forma conjunta nuevas formas de comunicación y transmisión del conocimiento científico, además de facilitarle y mejorar sensiblemente sus labores de búsqueda documental. Asimismo, como apunta Pinto Molina (2000: 2), la «revolución documental» permite al traductor abordar una mayor cantidad de información en un menor período de tiempo, lo cual contribuye a aumentar sensiblemente su rendimiento. Sin embargo, el traductor debe poseer un apropiado nivel de pericia en el empleo de las técnicas documentales más usuales, conociendo las estrategias necesarias para el procesamiento de los documentos. En consecuencia, el traductor ha de ser capaz de discernir la información fiable de la no fiable, pues, si bien la Red constituye una herramienta valiosísima, adolece de una serie de inconvenientes, entre los que destaca la credibilidad o fiabilidad de la información.

En este sentido, el resumen constituye una ayuda muy atractiva para el traductor profesional, puesto que, en su labor documental, un resumen orientativo e indicativo le lleva a seleccionar un determinado tipo de información. Habida cuenta del escaso tiempo del que normalmente disponen los traductores, el resumen aparece como una solución que, al reducir la información de documentos electrónicos o artículos, ofrece múltiples ventajas y facilita su labor traductológica en gran medida.

De esta forma, en el presente artículo, se ha llevado a cabo una aproximación a la extracción automática de resúmenes en el par de lenguas francés-español mediante dos objetos de estudio:

1. Por un lado, un ejemplo de condiciones del contrato de transporte aéreo, tanto en francés como en español, extraídas, por una parte, del macrocorpus multilingüe *Turicor*, el cual se elaboró en el seno del proyecto *TURICOR: Compilación de un corpus de contratos turísticos (alemán, español, inglés, italiano) para la generación textual multilingüe y la traducción jurídica*¹ (MCYT, BFF2003-04616, 2003-2006) y, por otra, del proyecto de excelencia de la Junta de Andalucía *La contratación turística electrónica multilingüe como mediación intercultural: aspectos legales, traductológicos y terminológicos* (n.º de referencia HUM-892, 2006-2009), ambos dirigidos por la Dra. Corpas Pastor.
2. Por otro lado, dos sistemas gratuitos y multilingües de extracción automática: *Copernic Summarizer*, un programa comercial del cual emplearemos la versión

¹ Es posible acceder al proyecto Turicor a través de su URL: <http://www.turicor.com> [Consulta: 10 de enero de 2007].

demo, y *WebSumm*², un sistema creado por David House, Eric Bloedorn e Inderjeet Mani en el seno de la investigación financiada por MITRE³.

Una vez elaborados los extractos de forma automática, se analizaron los resultados prestando atención tanto a sus aspectos positivos como negativos en aras de estudiar su impacto en la fase documental del traductor profesional.

Ámbito de estudio: entre la GLN y la GR

En un principio, cabría pensar que la generación de lenguaje natural (GLN) y la generación de resúmenes (GR), ambas enmarcadas en el Procesamiento del Lenguaje Natural (PLN), constituyen una misma actividad, pues ambas aluden al hecho de generar lenguaje de forma automática y, de hecho, varios autores consideran que la GR se trata de una disciplina dentro de la GLN (Lavid 2005: 216). Con el fin de observar las diferencias entre ambas, procederemos a mostrar una definición de cada una. La GLN, de acuerdo con Reiter y Dale (1997: 57) es:

Natural language generation is the subfield of artificial intelligence and computational linguistics that is concerned with the construction of computer systems that can produce understandable texts in [...] human languages from some underlying non-linguistic representation of information.

Por su parte, la GR es definida por Acero *et al.* (2001: 282) de la siguiente forma:

Por generación automática de resúmenes de texto entendemos el proceso por el cual se identifica la información sustancial proveniente de una fuente (o varias) para producir una versión abreviada destinada a un usuario particular (o grupo de usuarios) y una tarea (o tareas).

Por tanto, la GLN y la GR comparten el mismo objeto, la generación de un texto, pero se diferencian en la fuente de información, pues mientras en GLN se parte de una representación abstracta del conocimiento, tal y como apuntaban Reiter y Dale (1997), en generación de resúmenes se toma como base un texto en lengua natural. En cualquier caso, consideraremos que la GR constituye una modalidad de la GLN y, por ende, entra de lleno dentro del campo de las tecnologías de la traducción (TT). Dicha GR goza de un auge indiscutible debido a la cantidad ingente de información que encontramos en Internet, lo cual va estrechamente vinculado a las aportaciones documentales que le suponen al traductor.

En España, la investigación más representativa acerca de la generación automática de resúmenes posee dos focos principales: por un lado, la Universidad de Alicante, donde destacan los trabajos de Acero *et al.* (2001) con la generación automática de resúmenes personalizados en el seno del proyecto HERMES⁴, en el cual se aúnan aplicaciones de acceso a información textual multilingüe mediante la utilización de técnicas de ingeniería lingüística; por otro, el IULA-UPF, donde da Cunha (2005) ha llevado a cabo investigaciones acerca de la creación de un programa de generación automática de resúmenes para artículos médicos tomando como base la *Rhetorical Structure Theory* (RST) de Mann y Thompson (1988).

² Accesible en <http://complingone.georgetown.edu/~linguist/summarizer.html> [Consulta: 10 de enero de 2007].

³ MITRE constituye una asociación estadounidense sin ánimo de lucro cuyo interés principal es la tecnología. Su URL es la siguiente: <http://www.mitre.org/> [Consulta: 10 de enero de 2007].

⁴ La URL del proyecto HERMES es <http://nlp.uned.es/hermes/index.html> [Consulta: 10 de enero de 2007].

El resumen automático

Los sistemas de resúmenes automáticos, como señalan numerosos autores⁵, se dividen en dos grandes categorías: *extract* y *abstract*. Por un lado, como señalan Teufel y Moens (2002) un *extract* o extracto es un conjunto de pasajes (desde palabras sueltas a párrafos enteros) extraídos del texto de entrada literalmente, que unidos forman un resumen. La mayoría de las aproximaciones al resumen automático toman esta concepción (Luhn 1959; Edmundson 1969; Paice 1990), ya que resulta bastante útil en el campo de la recuperación de información para hacerse una idea del contenido de un texto, aunque los resultados del texto como tal no sean de una gran calidad.

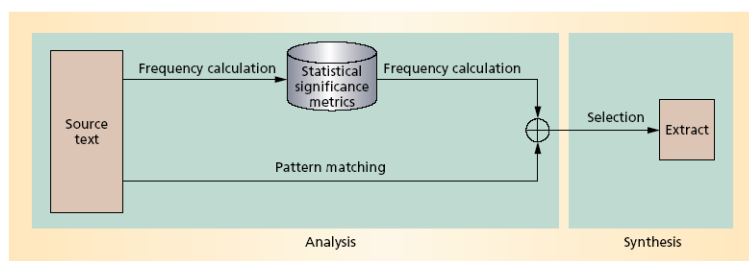


Figura 1. Representación de la extracción (Hahn y Mani 2000: 30)

Por otro lado, un *abstract* o resumen es un texto generado nuevamente, producido por alguna representación interna que resulta después del análisis del texto de entrada. Esta idea surge con el motivo de mejorar la calidad final del texto del resumen, y para ello se lleva a cabo un proceso posterior, acortando, fusionando o revisando el material (Marcu 2000).

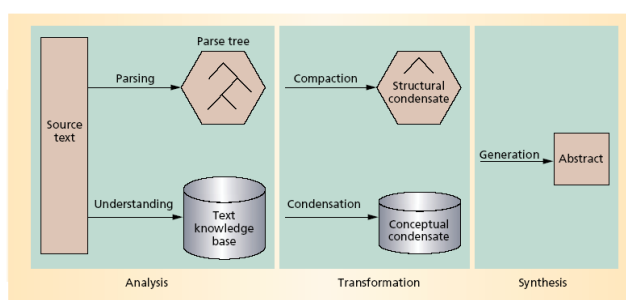


Figura 2. Representación del resumen propiamente dicho (Hahn y Mani 2000: 31)

En el presente trabajo, procederemos a analizar con mayor detenimiento la extracción automática, pues se trata de la opción disponible en los sistemas de resumen automático disponibles en Internet de forma gratuita.

Evaluación de sistemas de extracción automática

A continuación, pasaremos a analizar y evaluar dos sistemas de extracción automática disponibles en Internet de forma gratuita. La mayoría de los sistemas gratuitos que hallamos en la Red se limitan a realizar un *extract* del texto en cuestión, si bien hemos de tener en cuenta que existen sistemas comerciales o de pago⁶ que sí llevan a cabo

⁵ Consúltese las obras de Mani y Maybury (2001); Mani (2001); Endres-Niggemeyer (1998); Radev McKeown y Hovy (2002); Hahn y Mani (2000); Paice y Jones (1993); Da Cunha (2005); Teufel y Moens (2002).

⁶ En la Red hallamos numerosos sistemas de generación de resumen automático de pago, ya que se trata de productos comerciales. Entre dichos sistemas destacan los siguientes: *Brevity Document Summarizer* (disponible en <http://www.lextek.com/brevity/>), *Intelligent Miner for Text* (disponible en [3](http://www-</p></div><div data-bbox=)

resúmenes propiamente dichos, entre los que cabría resaltar *Summarist* (Hovy y Lin 1999), *DimSum* (Aone *et al.* 1999) y SCISOR (Lisa, Rau, Jacobs y Zernik 1989).

Categorización del documento de entrada

Como se ha comentado anteriormente (*cf.* introducción), el *input* estará constituido por un ejemplo de condiciones del contrato de transporte aéreo en las lenguas francesa y española extraídas, por un lado, del macrocorpus multilingüe *Turicor*, compilado en el seno del proyecto del MCYT BFF2003-04616 y, por otra, del proyecto de excelencia de la Junta de Andalucía HUM-892. Cabe apuntar que ambos proyectos poseen un carácter interdisciplinar que los sitúa en el ámbito de las nuevas TIC aplicadas a la traducción turística y jurídica.

Respecto al documento de entrada propiamente dicho, hemos de apuntar que constituyen las condiciones de transporte aéreo especificadas por la compañía francesa Air France y, aparte de en *Turicor*, pueden consultarse en línea en la URL [http://www.airfrance.de/double6/passage.nsf/\(LookupPublishedWeb\)/esPRSETLegalConditions?OpenDocument](http://www.airfrance.de/double6/passage.nsf/(LookupPublishedWeb)/esPRSETLegalConditions?OpenDocument).

Sistemas de extracción automática disponibles en Internet

Precisamos sistemas gratuitos⁷ disponibles en Internet que resuman textos en español y francés (y, si es posible, en otros idiomas) con el fin de analizar su funcionamiento así como sus resultados y carencias. La mayoría se basa en métodos estadísticos, entre los que destacan la posición en el texto (Edmundson 1969), la frecuencia de palabras y oraciones (Luhn 1958) o determinadas oraciones claves (por ejemplo, «es importante tener en cuenta») (Edmundson 1969). Asimismo, dichos sistemas recurren con frecuencia a la medida *tf.idf* (*term-frequency times inverse document frequency*), empleada con mucha frecuencia en la recuperación de información, (véanse Salton y McGill 1983; Teufel y Moens 1999; Kupiec, Pedersen y Chen 1995; Hovy y Lin 1999).

Una vez expuestas las características comunes de dichos programas, especificaremos que, aparte de *Copernic Summarizer* y de *WebSumm*, existen otros sistemas igualmente interesantes, entre los que destacan:

1. Extractor
2. GistSumm
3. Resumidor de Microsoft Word
4. SweSum

Análisis y evaluación de resultados

A continuación, mostraremos los extractos elaborados en ambas lenguas por cada uno de los sistemas con el fin de analizarlos posteriormente. En primer lugar, comenzaremos con *Copernic Summarizer* y, posteriormente, mostraremos los resultados extraídos mediante *WebSumm*.

Extracto elaborado por *Copernic Summarizer* en lengua española:

306.ibm.com/software/data/iminer/), *Inxight Summarizer* (en <http://www.inxight.com>), *Sinope Summarizer* (disponible en <http://www.sinope.info/en/Download>) y *TextAnalyst* (disponible en <http://www.megaputer.com/products/ta/index.php3>).

⁷ Al hablar de sistemas gratuitos, aludimos también a programas comerciales que facilitan la versión *demo*, como es el caso de *Copernic Summarizer*.

Concepts:

transportista, billete, equipaje, tarifas, condiciones, pasajero, Convenio, aviso, contrato, responsabilidad, condiciones generales, modificar, limitación, aplicables, documento.

Summary:

- El Convenio de Varsovia rige, y en la mayoría de los supuestos, limita la responsabilidad del transportista en caso de muerte o lesiones corporales, así como en caso de pérdida o robo de equipajes.

- Ver también el Aviso a Pasajeros Internacionales respecto a la limitación de responsabilidades y el Aviso de Limitación de responsabilidad en materia de equipajes.

- **CONDICIONES DEL CONTRATO DE TRANSPORTE AVISO:** El transportista se reserva el derecho de negarse a embarcar toda persona que se haya procurado un billete en violación de las leyes vigentes o de las tarifas o reglamento del transportista.

- 1. A los efectos de este contrato, la palabra «billete» significa billete de pasaje y talón de equipaje, o el Itinerario / Recibo de un billete electrónico, en su caso, del que forman parte las presentes Condiciones y Avisos adjuntos; la palabra «transportista» significa toda compañía aérea que transporte o se comprometa a transportar al pasajero y su equipaje en cumplimiento del contrato de transporte celebrado con el pasajero en una de las formas mencionadas; la palabra «billete electrónico» significa el Itinerario / Recibo emitido por o en nombre del transportista, los Cupones electrónicos al que hace referencia, y en su caso, un documento de embarque; «Convenio de Varsovia» significa el Convenio para la Unificación de ciertas Reglas relativas al transporte aéreo internacional, firmado en Varsovia el 12 de octubre de 1929, o dicho Convenio tal como fue modificado en La Haya, el 28 de septiembre 1955, según aplicación del uno o del otro.

Extracto elaborado por *Copernic Summarizer* en lengua francesa:

Concepts:

transporteur, bagages, billet, passager, Convention de Varsovie, responsabilité, conditions, Limitation de Responsabilité, internationaux, transporteur aérien, règles, billet électronique, Itinéraire/Reçu, territoire, réglementation applicable.

Summary:

- AVIS : La Convention de Varsovie peut être applicable si le voyage du passager comporte une destination finale ou une escale dans un autre pays que le pays de départ.

- La Convention de Varsovie régit, et dans la plupart des cas, limite la responsabilité du transporteur en cas de mort ou de lésions corporelles, ainsi qu'en cas de perte ou d'avarie de bagages.

- Se référer également à l'Avis aux Passagers Internationaux concernant la Limitation de Responsabilité et à l'Avis de Limitation de Responsabilité en matière de Bagages.

- AVIS: Le transporteur se réserve le droit de refuser d'acheminer toute personne qui s'est procuré un billet en violation des lois en vigueur ou des tarifs ou règlement du transporteur.

- 1. Au sens du présent contrat, le mot «billet» désigne le billet de passage et le bulletin de bagages, ou l'Itinéraire/Reçu d'un billet électronique, selon le cas, dont les présentes conditions et les avis joints font partie intégrante; le mot «transporteur» désigne toute compagnie aérienne qui transporte ou s'engage à transporter le passager et ses bagages en exécution du contrat de transport conclu avec le passager sous une des formes sus-visées; le mot «billet électronique»

désigne l'itinéraire/Reçu émis par le transporteur ou pour son compte, tout document électronique s'y rapportant et, le cas échéant, un document d'embarquement; «Convention de Varsovie» désigne la Convention pour l'unification de certaines règles relatives au transport aérien international, signée à Varsovie le 12 octobre 1929, ou cette même Convention amendée à La Haye, le 28 septembre 1955, selon que l'une ou l'autre est applicable.

Extracto elaborado por *WebSumm* en lengua española:

AVISO: El Convenio de Varsovia puede aplicarse cuando el viaje del pasajero efectúa un destino final o una escala dentro de un país distinto al de origen. El Convenio de Varsovia rige, y en la mayoría de los supuestos, limita la responsabilidad del transportista en caso de muerte o lesiones corporales, así como en caso de pérdida o robo de equipajes.

Ver también el Aviso a Pasajeros Internacionales respecto a la limitación de responsabilidades y el Aviso de Limitación de responsabilidad en materia de equipajes.

CONDICIONES DEL CONTRATO DE TRANSPORTE

AVISO: El transportista se reserva el derecho de negarse a embarcar toda persona que se haya procurado un billete en violación de las leyes vigentes o de las tarifas o reglamento del transportista.

En el caso de daño al equipaje en transporte internacional deberá presentarse por escrito la oportuna reclamaciones al transportista inmediatamente después de descubrirse el daño, y a lo sumo, dentro de los siete días siguientes a la fecha de entrega.; en caso de retraso, la reclamación deberá presentarse dentro de los veintiún días siguientes a la fecha en que fuera entregado el equipaje.

La tarifa del transportista objeto de las presentes es susceptible de modificación antes de iniciarse el viaje.

El transportista no asume la responsabilidad de garantizar los enlaces.

Ciertas tarifas están sujetas a condiciones particulares de venta y de transporte.

Las limitaciones de dimensiones y de peso señaladas son aplicables en las líneas de AIR FRANCE y de los asociados SKYTEAM

Extracto elaborado por *WebSumm* en lengua francesa:

La Convention de Varsovie régit, et dans la plupart des cas, limite la responsabilité du transporteur en cas de mort ou de lésions corporelles, ainsi qu'en cas de perte ou d'avarie de bagages.

Se référer également à l'Avis aux Passagers Internationaux concernant la Limitation de Responsabilité et à l'Avis de Limitation de Responsabilité en matière de Bagages.

CONDITIONS DU CONTRAT DE TRANSPORT

La durée de validité de ce billet est d'un an à compter de sa date d'émission, sauf disposition contraire dudit billet ou des tarifs, conditions de transport et règlements applicables du transporteur.

Le tarif du transport, objet des présentes est susceptible de modification avant le commencement du transport.

Le transporteur est en droit de refuser le transport si le tarif applicable n'a pas été payé.

Le transporteur s'engage à faire de son mieux pour transporter le passager et les bagages avec une diligence raisonnable.

Certains tarifs sont soumis à des conditions particulières de vente et de transport.

Como puede observarse, ambos sistemas ofrecen resultados distintos, de forma que cada uno de ellos aporta elementos positivos y negativos. Respecto a los extractos elaborados por *Copernic Summarizer* en español y francés, cabe destacar los siguientes rasgos positivos:

- Facilita una lista de conceptos claves, lo cual resulta sumamente interesante si se desea conocer el contenido de un documento.
- Las oraciones aparecen completas y guardan coherencia.
- Constan algunos de los elementos más importantes en el campo del transporte aéreo, como *transporteur, bagages, billet, passager, Convention de Varsovia*, etc.

En cuanto a los aspectos negativos, hemos de precisar los siguientes:

- Se repiten ciertos elementos, tal y como puede comprobarse en las palabras claves en lengua francesa, donde hallamos tanto *responsabilité* como *limitation de responsabilité*.
- Aparece información que no guarda gran relevancia y que, por ende, no resultaría pertinente en un resumen como es el caso de las definiciones del apartado 1 de ambos extractos.

En lo que atañe a los textos extraídos por *WebSumm*, resulta oportuno recalcar las siguientes aportaciones:

- Las oraciones aparecen completas y con coherencia.
- Mantiene los títulos principales como es el caso de *conditions du contrat de transport*.
- Ambos textos poseen aproximadamente la misma extensión.

No obstante, cabe apuntar los siguientes aspectos negativos:

- Al igual que ocurre con *Copernic Summarizer*, aparece información que no ha de constar en un resumen al no guardar gran relevancia, como es el caso de «Ciertas tarifas están sujetas a condiciones particulares de venta y de transporte» en español y «Certains tarifs sont soumis à des conditions particulières de vente et de transport» en francés.
- Los resúmenes carecen de coherencia al no mostrar una estructura global lógica, si bien guardan consistencia en lo que respecta a la cohesión.⁸

Aplicaciones prácticas del resumen automático en el proceso traductológico

Basándonos en parte en la utilidad del resumen según Pinto Molina (*apud* Pinto y Cordon 1999: 25), mostramos a continuación las aplicaciones prácticas del extracto automático a la hora de enfrentarse a una traducción determinada, en este caso la traducción de condiciones de transporte aéreo:

- Facilita una lista de conceptos claves en dos lenguas, lo cual resulta sumamente interesante si se desea conocer el contenido de un documento así como los términos de uso más frecuente. De esta forma, si el traductor desea conocer la terminología más empleada en los contratos hallará, en el caso del español, «transportista, billete, equipaje, tarifas, condiciones, pasajero, Convenio, aviso, contrato, responsabilidad, condiciones generales, modificar, limitación, aplicables, documento», etc. En cuanto a la terminología en lengua francesa, los

⁸ En el presente artículo nos centramos en las definiciones de *cohesión* y *coherencia* ofrecidas por Halliday y Hasan (1996).

términos más destacables son «transporteur, bagages, billet, passager, Convention de Varsovie, responsabilité, conditions, Limitation de Responsabilité, internationaux, transporteur aérien, règles, billet électronique, Itinéraire/Reçu, territoire, réglementation applicable», etc.

- Permite conocer la temática general del texto en cuestión. Gracias a los extractos automáticos, el traductor puede conocer, tras una rápida lectura, el tema central del texto y, de este modo, determinar si resulta o no de utilidad en su labor documental.
- En relación con la aplicación práctica anterior, cabe apuntar que los extractos ayudan a concluir sobre la conveniencia o no de consultar el texto completo, ya que, a partir del resumen automático, es posible determinar si el documento original resulta pertinente para la labor documental del traductor previa a la traducción en cuestión. En consecuencia, el traductor ve agilizado en gran medida el proceso documental, el cual, debido particularmente a la cantidad ingente de información disponible en la red, puede llegar a convertirse en una ardua y laboriosa tarea.
- Sirve como acceso temático en la recuperación de información que, por otra parte, es multilingüe. En consecuencia, el traductor podrá optar por emplear dicho texto como texto paralelo y como fuente de documentación o, por el contrario, buscar otro texto que posea mayor relevancia para una traducción determinada.
- Ofrece una fuente de actualización de conocimientos para el traductor acerca de las novedades en su campo de especialización.

No obstante, si bien los extractos ofrecen las aplicaciones prácticas anteriormente citadas, es preciso apuntar que, en aras de mejorar la calidad de los extractos, sería necesaria una «postedición»⁹ de los resúmenes a fin de que mejoren sus características, ya que, en la mayoría de los casos, carecen de coherencia y, por tanto, su efectividad se ve mermada en gran medida. Asimismo, sería aconsejable revisar la ortotipografía de los textos resultantes, pues aparecen diversos errores como falta de tildes, dobles espacios, etc.

Sin embargo, los sistemas de extracción automática, a pesar de sus carencias, aportan ciertos elementos positivos a la búsqueda documental. En cierto modo, puede resultar beneficiosa la inclusión de la extracción en la fase de documentación, pues el traductor se familiarizaría con determinados términos y construcciones sintácticas propias de los denominados textos paralelos y otras fuentes documentales de referencia.

Conclusiones

En el presente artículo se ha pretendido profundizar en la importancia de las nuevas tecnologías aplicadas a la Traducción, una disciplina que, por su carácter interdisciplinar, goza de las aportaciones de otras disciplinas afines o no y que, de esta forma, consolida su presencia y relevancia tanto en el ámbito nacional como internacional.

De este modo, basándonos en los dos ya citados proyectos de investigación desarrollados en la Universidad de Málaga, nos hemos aproximado al Procesamiento del Lenguaje Natural y, más concretamente, a la Generación del Lenguaje Natural y a la Generación de Resúmenes. Mediante la generación de resúmenes, y más en concreto de

⁹ Dicho término se emplea al hablar de la traducción automática y los lenguajes controlados y se refiere al proceso de revisar y corregir el texto de salida resultante de una traducción automática.

la extracción, hemos procurado mostrar la aportación de la GR al proceso documental del traductor, el cual, ante el acopio informativo, presente particularmente en Internet, precisa cribar las fuentes de documentación de las que dispone. Por tanto, el traductor debe poseer conocimientos de fuentes de información y técnicas de selección, evaluación, búsqueda y recuperación documental.

En este contexto, hemos mostrado que el resumen constituye un elemento sumamente útil para las labores documentales del traductor, y, muy especialmente, el resumen multilingüe, sobre todo por su doble función:

- El resumen constituye un elemento primordial en todas las áreas del saber, pues la tarea de redactar resúmenes es uno de los procedimientos que más facilita la acumulación de información de cara a la continua aparición de literatura científica.
- Se trata de una forma de acceder, en varias lenguas, a la información más relevante, pues supone una reducción a términos breves y precisos de lo esencial del contenido de un documento, cuyo propósito es facilitar al interesado la relevancia o no que tiene el documento original para su investigación.

Como conclusión, cabe afirmar que, a pesar de las patentes carencias existentes todavía en ciertos sistemas, gratuitos y multilingües, de resumen o extracción automáticos, la generación automática de resúmenes supone una ayuda para el traductor en su labor documental. Además, facilita en gran medida la consulta del usuario a la Red, no sólo de los profesionales de la Traducción, pues supone un nuevo paso hacia la optimización del tratamiento de la información en formato digital. Por último, el resumen y la extracción multilingües y automáticos constituyen un elemento clave a través del cual es posible fomentar la aplicación de las nuevas tecnologías en el campo de la didáctica y la práctica de la traducción profesional, al tiempo que sería un trasvase interesante, en tanto (novedosa) aplicación de la Traductología a la Lingüística Computacional y al Procesamiento del Lenguaje Natural.

Referencias bibliográficas

- Aceró, I., M. Alcojor, A. Díaz, J. M. Gómez y M. J. Maña (2001). «Generación automática de resúmenes personalizados». *Procesamiento del Lenguaje Natural* 27. 281-287. Disponible en <http://www.sepln.org/revistaSEPLN/revista/27/> [Consulta: 10 de enero de 2007].
- Aone, C. M., E. Okurowski, J. Gorlinsky y B. Larsen (1999). «A Trainable Summarizer with Knowledge Acquired from Robust NLP Techniques». En I. Mani y M. Maybury (eds.). *Advances in Automatic Text Summarization*. Cambridge: MIT. 71-80.
- Corpas Pastor, G. (ed.) (2003). Recursos documentales y tecnológicos para la traducción del discurso jurídico (español, alemán, inglés, italiano, árabe). Granada: Comares.
- Da Cunha, I. (2005). Hacia un modelo lingüístico de resumen automático de artículos médicos en español. Barcelona: Institut Universitari de Lingüística Aplicada- Universitat Pompeu Fabra. Disponible en <http://www.upf.edu/pdi/iula/iria.dacunha> [Consulta: 10 de enero de 2007].
- Edmundson, H. P. (1969). «New Methods in Automatic Abstracting». En I. Mani y M. Maybury (eds.). *Advances in Automatic Text Summarization*. Cambridge: MIT. 23-42.
- Endres-Niggemeyer, B. (1998). *Summarizing Information*. Hannover: Springer.
- Hahn, U. e I. Mani (2000). «The Challenges of Automatic Summarization». *IEEE Computer* 33 (11). 29-36.
- Hovy, E. y C. Lin (1999). «Automated Text Summarization in SUMMARIST». En I. Mani y M. Maybury (eds.). *Advances in Automatic Text Summarization*. Cambridge: MIT. 81-94.

- Kupiec, J., J. Pedersen y F. Chen (1995). «A Trainable Document Summarizer». En E. Fox, P. Ingwersen y R. Fidel (eds.). *Proceedings of the 18th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. Nueva York: ACM. 68-73. Disponible en <http://www.dcs.shef.ac.uk/~mlap/teaching/kupiec95trainable.pdf> [Consulta: 20 de diciembre de 2006].
- Lavid, J. (2005). *Lenguajes y nuevas tecnologías. Nuevas perspectivas, métodos y herramientas para el lingüista del siglo XXI*. Madrid: Cátedra.
- Lisa, F., P. Rau, S. Jacobs y U. Zernik (1989). «Information Extraction and Text Summarization using Linguistic Knowledge Acquisition». *Information Processing and Management* 25 (4). 419-428.
- Luhn, H. P. (1969). «The Automatic Creation of Literature Abstracts». En I. Mani y M. Maybury (eds.). *Advances in Automatic Text Summarization*. Cambridge: MIT. 15-21.
- Mani, I. y M. Maybury, M. (eds.) (2001). *Advances in Automatic Text Summarization*. Cambridge: MIT.
- Mann, W. C. y S. A. Thompson (1988). «Rhetorical Structure Theory: Toward a functional theory of text organization». *Text* 8 (3). 243-281.
- Marcu, D. (2000). *The Theory and Practice of Discourse Parsing and Summarization*. Cambridge: MIT.
- Paice, C. D. (1990). «Constructing Literature: Abstracts by Computer: Techniques and Prospects». *Information Processing and Management* 26 (1). 171-186.
- Pinto M. y J. A. Cordón (1999). *Técnicas documentales aplicadas a la Traducción*. Madrid: Síntesis.
- Pinto Molina, M. (2000). «Documentación para la Traducción en la sociedad de la información». En *XV Coloquio Association Internationale de Bibliologie*. Salamanca: AIB. Disponible en <http://www.mariapinto.es/web/mainframe.htm> [Consulta: 5 de abril de 2006].
- Radev, D., E. Hovy y K. McKeown (2002). «Introduction to the Special Issue on Summarization». *Computational Linguistics* 28 (4). 339-408.
- Reiter, E. y R. Dale (1997). «Building applied natural-language generation systems». *Journal of Natural-Language Engineering* 3. 57- 87.
- Salton, G. y M. McGill (1983). *Introduction to Modern Information Retrieval*. Nueva York: McGraw Hill.
- Teufel, S. y M. Moens (2002). «Summarizing scientific articles: Experiments with relevance and rhetorical status». *Computational Linguistics* 28 (4). 409-445.