

Necesidades formativas emergentes: corpus virtuales como recurso para la formación en Traducción e Interpretación en los Servicios Públicos¹

Emerging training needs: virtual corpora as resources for Public Service Interpreting and Translation training

María del Mar Sánchez Ramos

Universidad de Alcalá
España

ONOMÁZEIN 38 (diciembre de 2017): 169-187
DOI: 10.7764/onomazein.38.07



María del Mar Sánchez Ramos: Departamento de Filología Moderna, Facultad de Filosofía y Letras, Universidad de Alcalá (Madrid), España. | Correo electrónico: mar.sanchezr@uah.es

Fecha de recepción: agosto de 2016
Fecha de aceptación: diciembre de 2016

Resumen

La creciente necesidad de una mediación lingüística y cultural entre los agentes de los servicios públicos y los usuarios que no dominan la lengua de las administraciones públicas ha hecho posible la aparición de una disciplina dentro de los Estudios de Traducción llamada Traducción e Interpretación en los Servicios Públicos (TISP), que engloba distintos ámbitos especializados de actuación como el sanitario o administrativo, entre otros. La documentación en estos casos se presenta como necesaria para que el traductor e intérprete adquiriera las habilidades y los conocimientos adecuados a la hora de producir el texto meta. Nuestro artículo describe cómo se ha aplicado la lingüística de corpus a la traducción de textos especializados en el aula de TISP como recurso lexicográfico y documental adicional. Para ello, se detalla el proceso de compilación de un corpus monolingüe virtual especializado realizado por alumnos de posgrado y el uso de un programa de concordancias para analizar el corpus creado como fuente de información conceptual, terminológica y fraseológica.

Palabras clave: Traducción e Interpretación en los Servicios Públicos (TISP); corpus especializados; compilación de corpus; programas de concordancias.

Abstract

The increasing need for a linguistic and cultural mediation between public service providers and users who do not have sufficient command of the official language of the public authorities has contributed to the development of a discipline within Translation Studies—Public Service Interpreting and Translation Studies (PSIT)—, covering different areas such a health-

-
- 1 La investigación descrita en este artículo se enmarca dentro del grupo de investigación “FITIS-Pos - Formación e Investigación en la Traducción e Interpretación de los Servicios Públicos” (REF CCHH2006/R04), financiado por la Universidad de Alcalá (Madrid, España). De igual forma, dicha investigación forma parte de un proyecto de innovación educativa llamado “NOVATIC: Integración de nuevas herramientas TIC basadas en corpus en el aula de traducción especializada” (Ref. PIE-145), proyecto financiado por la Universidad de Málaga. Por último, la propuesta descrita en el artículo se ha desarrollado también como actividad investigadora dentro del grupo de innovación docente Herramientas Informáticas para la Traducción e Interpretación y su docencia - HITID (UAH-GI14-71), Universidad de Alcalá (Madrid, España).

care or administrative settings. It seems that documentation is essential at these stages so that the translator/interpreter can acquire the appropriate skills and knowledge to be able to produce the target text. This paper describes how we have used a corpus methodology as an additional lexicographical and documentation resource for the translation of healthcare texts in the PSIT classroom. We show the corpus compilation process our postgraduate students followed and how they worked with a well-known concordance program in order to analyze their corpus as a source of conceptual, terminological and phraseological information.

Keywords: Public Service Interpreting and Translation (PSIT); specialized corpora; corpus compilation; concordance programs.

1. Introducción

Estamos siendo testigos de una necesidad imperiosa de comunicación entre los proveedores de servicios públicos (hospitales, juzgados, comisarías de policía, administraciones públicas) y los usuarios de estos servicios que no dominan de forma fluida la lengua de uso (Mikkelsen, 1996). Ante esta situación, y el cada día más acuciante movimiento migratorio, se hace imprescindible la figura del traductor e intérprete en el contexto de los servicios públicos con el fin de asegurar una comunicación exitosa. Todo ello ha hecho posible el desarrollo de la llamada Traducción e Interpretación en los Servicios Públicos (TISP²), cuya principal característica es la de ser el medio de comunicación entre las entidades públicas y una minoría cultural y lingüística que no domina (o desconoce) la lengua (Valero Garcés, 2005). Buen ejemplo de esta necesidad lingüística y de la formación que se precisa en estos ámbitos es el proyecto de investigación FITISPos (Formación e Investigación en la Traducción e Interpretación en los Servicios Públicos, Universidad de Alcalá, Madrid, España) en el que se enmarca este trabajo³. De esta necesidad formativa se deriva una instrucción específica de la que los docentes de este ámbito somos conscientes. Si bien la formación en TISP ha sido centro de atención debido, principalmente, a la escasez de programas formativos, el creciente interés de esta disciplina y la necesidad de dotar a la sociedad de profesionales cualificados en un entorno de TISP han sido las principales razones por la que hoy en día se cuenta con programas de formación en distintos ámbitos de TISP (Abril y Martín, 2008; Angelelli, 2006; Vargas Urpí, 2016). Prueba de ello es el programa de posgrado de la Universidad de Alcalá (Madrid, España). Bajo el título de Máster en Comunicación Intercultural, Traducción e Interpretación en los Servicios Públicos⁴, esta propuesta docente engloba módulos específicos centrados en la traducción y la interpretación sanitaria y jurídico-administrativa, con el fin de cubrir las necesidades que los usuarios de los servicios públicos demandan.

-
- 2 El trabajo que se describe a continuación se centrará en el ámbito de la traducción dentro de TISP. Somos conscientes de que la interpretación, tanto la formación de intérpretes (Abril Martí, 2006) como el uso de la metodología de corpus en la interpretación en los servicios públicos (Lázaro Gutiérrez y Sánchez Ramos, 2015), despierta gran interés, pero un análisis exhaustivo del uso de la metodología de corpus en la interpretación sobrepasaría los límites de nuestro trabajo.
 - 3 Puede encontrarse más información sobre este grupo en http://www3.uah.es/traduccion/grupo_fitispos.html. En esta línea, es de destacar también la iniciativa Red Comunica (<http://red-comunica.blogspot.com.es/>), que recoge información relevante (jornadas, congresos, grupos de investigación, etc.) sobre TISP.
 - 4 Este programa de formación de posgrado forma parte de la prestigiosa red *European Commission's European Master's in Translation* (EMT). Ofrece formación en distintos pares de lenguas, como son español e inglés, francés, árabe, chino, ruso o rumano. Puede encontrarse más información en el siguiente enlace: <http://www3.uah.es/master-tisp-uah/presentacion/>.

No hay duda de que la documentación es esencial en el día a día del traductor e intérprete, y que se constituye como uno de los factores esenciales para garantizar el éxito del proceso de trasvase y del producto final. Es por esta razón por la que la formación en documentación debe estar presente en los programas de enseñanza en TISP. Así, y teniendo en cuenta las necesidades específicas que el traductor e intérprete que trabaja en el ámbito de los servicios públicos debe desarrollar, el presente trabajo describe cómo se ha integrado y desarrollado una metodología de corpus como parte de la formación documental del traductor e intérprete en TISP. En concreto, se detalla cómo ha sido la metodología para la compilación de un corpus monolingüe virtual en el aula de TISP como recurso documental dentro del ámbito sanitario. Nuestro objetivo es doble: de un lado perseguimos que el alumnado, con un bagaje académico diverso, sea consciente de la importancia de manejar herramientas documentales más allá de los convencionales diccionarios, glosarios o bases de datos a la vez que sean capaces de compilar y sacar el máximo partido a un corpus monolingüe virtual para una mejor comprensión de los textos especializados, y, de otro, el desarrollo de la competencia documental y tecnológica del alumnado.

2. La documentación en TISP

La documentación supone uno de los pilares sobre los que debe asentarse la formación del traductor e intérprete en general y, principalmente, debe reforzarse en ámbito de los servicios públicos, ya que traductores e intérpretes se mueven en contextos muy específicos y en situaciones que, en algunos casos, pueden ser muy delicadas (relación paciente-doctor, interpretación judicial, entre otras). El proceso de documentación, por tanto, será clave. Y es que la traducción, entendida como proceso informativo, se debe a una comprensión del texto original y una re-expresión como texto meta, donde los conocimientos lingüísticos y documentales son inexcusables. De una buena documentación dependerá que el traductor e intérprete adquiera los conocimientos temáticos, terminológicos, fraseológicos y textuales necesarios para llevar a cabo el proceso de trasvase (Hurtado Albir, 2001). Este proceso de adquisición de conocimientos durante el proceso traductor se conoce de forma generalizada como competencia documental⁵ (Hurtado Albir, 2001; Pinto Molina y Sales Salvador, 2008). Dentro de TISP, conocer las distintas fuentes de documentación, saber utilizarlas y seleccionar las más adecuadas de acuerdo a la tarea asignada debe formar parte de la tarea diaria de un traductor e intérprete.

5 Como recoge Durán Muñoz (2014), esta competencia también se conoce como competencia heurístico-documental, denominada así por la norma europea EN-15038 (CEN, 2006), subcompetencia heurística (Corpas Pastor, 2001), competencia profesional e instrumental (Rabadán y Fernández, 2002) o infocompetencia (Sales Salvador, 2006).

No es la finalidad de este trabajo ofrecer un plan de documentación exhaustivo, pero sí creemos interesante (y muy útil) la distinción que la estudiosa Orozco Jutorán (2012: 220) realiza de los distintos tipos de documentación que el traductor especializado debe dominar: la documentación temática o conceptual, la terminológica y la contextual y co-textual. En primer lugar, y relacionado con la documentación temática, es labor del traductor e intérprete el adquirir los conocimientos necesarios de un determinado campo de especialidad para una adecuada comprensión del texto (Orozco Jutorán, 2012: 222). La autora menciona los tesauros y obras que varían en su grado de especialización (textos divulgativos, textos escritos por especialistas) entre los recursos que pueden ayudar en la documentación temática. En cuanto a la documentación terminológica, la identificación y comprensión de los términos especializados del texto origen es crucial, al igual que las palabras de uso general y que adquieren un significado especializado según su uso. Las obras que la autora propone son, en este caso, de corte lexicográfico (diccionarios generales, especializados, glosarios, bases de datos terminológicas o diccionarios visuales especializados). Por último, en la documentación contextual y co-textual, en palabras de la estudiosa, “el traductor debe comprobar que la macroestructura y la microestructura del TM sean adecuadas a la función del TM y que éste se lea de forma fluida y natural por parte del lector final [...]” (Orozco Jutorán, 2012: 225). Destaca, además, el uso de los textos paralelos como principal fuente documental.

La docencia en documentación y TISP viene a solventar los vacíos documentales que los alumnos puedan presentar⁶. En relación con la documentación y la traducción especializada, ha de destacarse las distintas habilidades y destrezas que el traductor debe poseer y desarrollar. Corpas Pastor (2004) habla de subcompetencia lingüística (diferencia del discurso científico de las dos lenguas), subcompetencia cultural y textual (conocimiento adecuado de la estructura retórica de los textos, normas de producción de textos en la lengua de origen y de llegada) y, por último, de subcompetencia documental. Esta última implicaría un conocimiento especializado sobre un tema en concreto (por ejemplo, ámbito sanitario), conocimiento de documentos especializados en distintas fuentes, conocimiento de fuentes información, uso adecuado de las mismas y evaluación de las mismas y rapidez en el proceso de trasvase y calidad del producto final. Con todo ello, parece innegable que el proceso de documentación supone como un gran reto en el proceso de traducción y cobra una especial importancia en los contextos especializados, por las particularidades de los mismos. Como ejemplo retomamos las palabras de Montalt y González Davies (2007), quienes señalan los vacíos de información que el traductor especializado en el ámbito sanitario debe cubrir: a) información temática:

6 Remitimos a los datos que Valero Garcés (2013) arroja sobre la variedad multicultural y de formación de los alumnos del Máster en Comunicación Intercultural, Traducción e Interpretación de los Servicios Públicos (Universidad de Alcalá, Madrid, España) y que inciden de forma substancial en la estructura curricular para que esté acorde a las necesidades del alumnado.

conocimiento del área de conocimiento especializado; b) información terminológica: conocimiento de términos especializados, abreviaturas, símbolos, etc., y c) información textual: conocimiento de los géneros textuales, convenciones textuales, estructurales y estilísticas de determinados textos

En esta línea, y muy de la mano de la llamada formación basada en competencias (Hurtado Albir, 2007) y de las directrices marcadas por el EEES (Valero Garcés, 2013), el Máster en Comunicación Intercultural, Traducción e Interpretación en los Servicios Públicos de la Universidad de Alcalá oferta un módulo llamado Técnicas y Recursos en TISP en el ámbito sanitario, que incluye, junto con formación en traducción asistida por ordenador (TAO) y terminología, contenidos de metodología de corpus. Los principales objetivos de dicho módulo son: a) buscar, seleccionar y evaluar fuentes de documentación para el ámbito sanitario en TISP; b) saber dominar recursos de consulta lingüística y terminológica en Internet; c) saber dominar herramientas de consulta (corpus, programas de concordancias) como recursos documentales. Es en este contexto docente donde se enmarca nuestra propuesta, que describiremos en las siguientes secciones.

3. Lingüística de corpus y traducción

Los llamados Estudios de Traducción Basados en Corpus han sido fuente de inspiración para estudiar diversos aspectos traductológicos (Baker, 1995, 1996; Kenny, 2001; Saldanha, 2004), el uso de corpus como recurso terminológico (Bowker, 2003; Zanettin y otros, 2003; Zhu y Wang, 2011), el uso de corpus en la traducción especializada (Sánchez Gijón, 2009; Vila Barbosa, 2013) o la aplicación de una metodología de corpus para fines pedagógicos (Calzada Pérez, 2005; Krüger, 2012; Kübler, 2011; Fantinuoli y Zanettin, 2015). En cuanto a la tipología de corpus, existe una amplia gama (Bowker y Pearson, 2002: 11), pero nos parece apropiada la distinción de Laviosa (2002: 36), que resume los principales corpus que se utilizan en los estudios de traducción. Esta académica divide los corpus monolingües en simples (recopilación de textos en una única lengua) y comparables (recopilación de textos originales en una lengua A y textos traducidos en esa misma lengua A); los bilingües quedan clasificados en: paralelos (textos originales en lengua A y sus traducciones en lengua B) y comparables (textos originales en lengua A y textos originales en lengua B); y los multilingües se agrupan en: paralelos (textos originales en lenguas diversas con sus respectivas traducciones), comparables (“bi/multilingual corpus made up to two or more sets from the same subject domain(s)” (Laviosa, 2002: 36)).

Junto con esta tipología, cada día está adquiriendo una mayor popularidad, debido principalmente a las necesidades docentes, la compilación de corpus de dimensiones más reducidas (monolingües, comparables o paralelos) como fuente de documentación en tareas de

traducción concretas o bien como recurso pedagógico, y cuyo material no es otro que textos electrónicos extraídos de Internet y que han seguido una evaluación de dichos textos y un criterio o protocolo de compilación concreto (Corpas Pastor, 2008; Seghiri, 2011). Se trata de los llamados *corpus virtuales*⁷, colecciones de textos extraídos de Internet y compilados “for the sole purpose of providing information—either factual, linguistic or field specific—for use in completing a translation task” (Sánchez Gijón, 2009: 115). El objetivo último es reunir la mayor información posible. La calidad no es la principal preocupación en estos casos, sino que lo es la selección de textos que documentan el corpus así como su evaluación tanto temática como textual en relación al texto origen desde el que se parte. Este tipo de corpus es considerablemente útil como herramienta de documentación en temas especializados, donde los recursos pueden llegar a ser escasos.

Los beneficios pedagógicos del empleo derivado del uso de corpus en traducción también han sido destacados (Bowker y Pearson, 2002: 10; Lee y Swales, 2006: 74; Corpas Pastor y Seghiri, 2009: 102). La mayoría de los autores se centran en describir las distintas estrategias y técnicas empleadas por los traductores en el análisis de los corpus creados, pero son menos las investigaciones que detallan cómo se traslada al aula la lingüística de corpus en forma de actividades que desarrollen las habilidades tecnológicas del alumnado (Rodríguez Inés, 2010). Una de las principales ventajas de esta línea que se plantea es el desarrollo de la llamada subcompetencia instrumental (PACTE, 2003: 53) o lo que el grupo de expertos de EMT (2009) denomina *information mining competence*. Estos últimos inciden en la necesidad de que, debido a los avances tecnológicos en los que vivimos inmersos, los traductores e intérpretes conozcan y manejen herramientas de concordancia. El uso de los programas de concordancias y el análisis de corpus electrónicos permite a los traductores e intérpretes la realización de una serie de tareas, como puede ser la identificación y verificación de patrones léxicos, fraseológicos y textuales (Neunzig, 2003) o el fomento de la autonomía del aprendizaje del alumno (Ulrych, 2000). Otros autores van más allá, como es el caso de Oster (2007), para quien uno de los retos docentes en materia de lingüística de corpus es el desarrollo de lo que ella llama *corpus linguistic awareness*, que supondría conocer los distintos tipos de corpus, desarrollar estrategias de búsqueda de textos para la compilación de corpus, identificar las características lingüísticas que un corpus puede ofrecer frente a un diccionario o la interpretación de los resultados, entre otras características, tareas que sin duda desarrollarían la competencia instrumental de los alumnos.

7 Este tipo de corpus recibe otras denominaciones, como bien señala Seguiri (2011). Junto con *corpus virtual* (Ahmad y otros, 1994; Seguiri, 2006), otros autores prefieren hablar de *corpus ad hoc* (Corpas Pastor, 2001), *disposable corpus* (Varantola, 2003) o *do-it-yourself (DIY) corpus* (Zanettin, 2002).

4. Compilación de un corpus virtual para su explotación en el aula de traducción especializada

Como se ha mencionado en párrafos anteriores, el uso de diversas fuentes de documentación y manejo de herramientas electrónicas es un aspecto de máxima importancia en TISP debido, en gran parte, al variado y diverso número de fuentes de documentación que el traductor e intérprete tiene que emplear para comprender de manera adecuada su tarea de traducción o interpretación. Por otro lado, en lo que se refiere al ámbito sanitario y TISP, la terminología especializada entraña una dificultad añadida y un escollo que el traductor e intérprete debe solventar para realizar con éxito el proceso de trasvase comunicativo. La elaboración de glosarios, por ejemplo, se constituye como una práctica frecuente en estas tareas con el objeto de llegar a comprender el texto que ha de traducirse y permitir la producción de textos en el ámbito sanitario (Gómez González-Jover y Vargas Sierra, 2004).

Es por esta razón que, teniendo en cuenta que la documentación en la formación de TISP ocupa un lugar destacado con el fin de asegurar la producción de un texto adecuado desde el punto de vista lingüístico y funcional, nos embarcamos en la tarea de formar a nuestros alumnos de posgrado de TISP en metodología de corpus.

La propuesta que aquí se describe se realizó por primera vez con un grupo de 38 alumnos (35 alumnos eran nativos españoles y 3 alumnos eran nativos ingleses), cuyo par de lenguas era inglés-español. La lengua inglesa, por tanto, era la lengua de trabajo. Se les presentó un texto en español, sobre el que tenían que documentarse para su posterior traducción al inglés, si bien la actividad se presentaba como primer paso de documentación antes de la fase de traducción.

Para cumplir nuestros objetivos (ver sección 2), creímos conveniente instruir en la compilación de un corpus monolingüe virtual especializado como fuente de documentación conceptual, terminológica y fraseológica para la traducción de un texto especializado que versaba sobre la llamada anemia drepanocítica (en inglés *sickle-cell anemia*). Describimos en las siguientes páginas cómo se llevó a cabo esta metodología basada en corpus y cómo nuestros alumnos de posgrado procedieron en las distintas etapas de compilación y análisis del mencionado corpus con la ayuda de un programa de concordancias (AntConc versión 3.4.3)⁸.

Previamente al proceso de compilación, se procedió a la fase del diseño del corpus. Se pidió a los alumnos el diseño de un corpus virtual monolingüe formado por textos especializados sobre la anemia drepanocítica escrito en inglés británico. Nuestro alumnado asistió

8 Este programa informático, desarrollado por Laurence Anthony y completamente gratuito, puede descargarse del siguiente enlace: <http://www.laurenceanthony.net/software.html>.

a dos sesiones formativas de un total de 8 horas de duración. En la primera sesión (4 horas), se presentaron los conceptos teóricos sobre Estudios de Traducción Basados en Corpus, los principales recursos documentales electrónicos en la TISP (diccionarios especializados, glosarios terminológicos, diccionarios visuales, bases de datos terminológicas, etc.) a la vez que se instruyó en las distintas estrategias de búsqueda (operadores básicos, expresiones booleanas y operadores filtro) para sacar el máximo partido a las búsquedas por Internet. Este último aspecto fue de especial relevancia, ya que una búsqueda precisa permitiría a nuestros estudiantes encontrar fuentes de información especializada para la creación de su corpus virtual de calidad en etapas posteriores.

Los alumnos también se formaron en el uso de programas de concordancias, como AntConc, y se familiarizaron con sus principales funciones (generador de listados de palabras frecuentes, uso de concordancias, recuperación de colocaciones y paquetes léxicos). En la segunda sesión (4 horas), se les pidió el encargo de compilar un corpus monolingüe virtual (inglés británico) como herramienta documental previa a la traducción de un texto especializado sobre la anemia drepanocítica. Como señala Seghiri (2011), es imprescindible establecer unos parámetros de diseño de un corpus o protocolo de compilación. Siguiendo muy de cerca la metodología establecida por esta autora, se les presentó una metodología de compilación de corpus a los discentes, que quedó dividida en tres fases: a) fase documental, b) fase de compilación y c) fase de análisis.

a) Fase de documentación

Una de las etapas iniciales, y podríamos afirmar que cruciales, es el proceso de documentación previo al diseño del corpus. Es fundamental que el alumnado se familiarice con el tema para poder realizar una búsqueda de textos de calidad⁹ que formen su corpus. Sin duda, Internet se ha convertido en lo que Kilgarriff y Grefenstette (2003) han llamado “a fabulous linguists’ playground”. Huelga decir que desarrollar estrategias de búsqueda permitirá una búsqueda mucho más efectiva y repercutirá en la calidad de del corpus. Según Austhermül (2001), tres son los tipos de búsqueda que deben seguirse para que nuestra búsqueda en Internet sea efectiva: búsqueda institucional, la que se realiza a través de distintos sitios web de organizaciones, instituciones o cualquier tipo de asociación; búsqueda temática, la que puede realizarse a través de portales especializados; y búsqueda por palabras clave, la que se realiza a través de buscadores. Al igual que Seghiri (2011), creemos que estos tres tipos de búsqueda pueden reducirse a dos: búsqueda institucional y búsqueda por palabras clave, puesto

9 Puede consultarse Seghiri (2014) para profundizar en el tema de diseño de corpus de calidad. Para los fines propuestos en nuestra investigación, no hemos creído conveniente extendernos más en este aspecto.

que los portales temáticos incluidos en la búsqueda temática pueden considerarse sitios web especializados. Tras ello, y con la ayuda de descriptores, búsquedas avanzadas que los alumnos habían adquirido en sesiones previas, los alumnos, tras documentarse sobre el tema proporcionado y la lectura de textos especializados, procedieron a buscar y seleccionar textos a través de palabras clave o *seed words* como *sickle-cell anaemia*, *hemoglobin*, *erythrocytes* o *Hb SS disease*, y la combinación de las mismas y búsqueda en sitios web institucionales. Junto con los operadores booleanos, los alumnos también emplearon los llamados operadores filtro para restringir los parámetros de dominio o sitio. Por ejemplo, si utilizamos la siguiente ecuación **sickle-cell site: www.nhlbi.nih.gov**, el buscador nos recuperará una selección de textos especializados dentro del sitio web señalado que contenga la palabra *sickle-cell*. Por el contrario, si empleamos **sickle cell site: co.uk**, nos ofrece páginas que contengan la palabra *sickle cell*, pero que se encuentren bajo el dominio co.uk, lo que delimitará diatópicamente nuestras búsquedas y responderá a uno de nuestros criterios de compilación (inglés británico).

TABLA 1

Algunos ejemplos de operadores filtro

| OPERADORES | DEFINICIÓN | EJEMPLO |
|------------|--|-------------------------------------|
| Site | Permite buscar palabras o frases dentro de un sitio web o un dominio concretos | sickle-cell site: www.nhlbi.nih.gov |
| Intitle | Permite buscar palabras en los títulos de las páginas web | intitle:sickle cell anemia |
| Related | Permite buscar sitios web que comparten características similares | related:www.nhlbi.nih.gov |
| Link | Permite buscar páginas web que vinculan a otras páginas en concreto | link:www.nhlbi.nih.gov |

b) Fase de compilación

Una vez ubicados los textos, los alumnos procedieron a la descarga, almacenamiento y conversión a texto plano (.txt) de los mismos. La descarga puede realizarse de forma manual, con la descarga de cada página web (CTRL+S), donde se nos guardará en formato .html o bien se pueden utilizar gestores de descarga de sitios web (Wget, WinHTTrack Website Copier¹⁰).

10 Se trata de dos programas muy útiles para la descarga de sitios web. El primero de ellos sirve tanto para la descarga de sitios web como páginas web individuales. La principal diferencia que es Wget está incluida en Linux, aunque también puede descargarse en Windows, y el segundo programa es multiplataforma.

En cuanto al almacenamiento y conversión, existen gestores para nombrar archivos en lote (Personal Renamer, FastFile Rename y Lupass Rename) y programas para la conversión de formatos (HTMLasText y PdftoText). Estos programas también permiten la conversión en lote de páginas web.

c) Fase de análisis

Uno de los principales objetivos que nos planteábamos a la hora de incorporar la metodología de corpus en el aula de TISP era que los alumnos tomaran conciencia del uso de corpus como recurso de documentación y fuente de adquisición conceptual, terminológica y fraseológica. En esta tercera fase se entra de lleno en la explotación del corpus creado. Hay diversos programas informáticos que pueden utilizarse en esta fase (e. g., Wordsmith Tools o Monoconc Pro). Hemos creído conveniente utilizar el programa de concordancias creado por Laurence Anthony, AntConc (versión 3.4.3), ya que se trata de un programa gratuito, con una interfaz sencilla y fácil de utilizar. Los alumnos comenzaron a utilizar su corpus virtual como herramienta terminológica. Los discentes generaron una lista de palabras de términos monoléxicos ordenados por orden alfabético y frecuencia de aparición (figura 1). A su vez, crearon lo que se llama *stoplist* para filtrar palabras como artículos, adverbios y preposiciones. De este modo, se evitaba que palabras irrelevantes o el llamado ‘ruido’ aparecieran como palabras más frecuentes en el corpus. Este tipo de listas, que pueden subirse al programa AntConc, ofrece la posibilidad de crearlas manualmente o bien subir alguna de las listas que pueden encontrarse por Internet¹¹.

El listado de unidades monoléxicas reveló el uso de palabras de uso general que adquieren un significado especializado en el corpus: *red, blood, treatment, results*, y vocabulario especializado relacionado con la temática principal del corpus (*oxygen, hemoglobin, haemolysis, erythrocytes, bilirubin*). El listado de palabras fue realmente útil también para trabajar con acrónimos (DNA, *Deoxyribonucleic acid*; SCD, *sickle-cell anaemia*; RCB, *red blood cells*) y abreviaturas (ml, *mililitre*; hb, *hemoglobin*; db, *decibel*; hbs, *hemoglobin S*). Los listados de palabras creados permitieron que el alumno explorara el comportamiento de dichas palabras en contexto gracias a la función de concordancia (*concordance*), que posibilitó la generación de un listado de concordancias. Otras funciones que los alumnos utilizaron fueron *collocates* para la creación de listados de colocaciones y patrones colocacionales así como los llamados agrupamientos léxicos (*clusters*), que permitió explotar en más detalle el corpus creado.

11 Algunas direcciones útiles para la descarga de las llamadas “stop lists” son: <http://www.lextek.com/manuals/onix/stopwords1.html>, <http://www.ranks.nl/stopwords>.

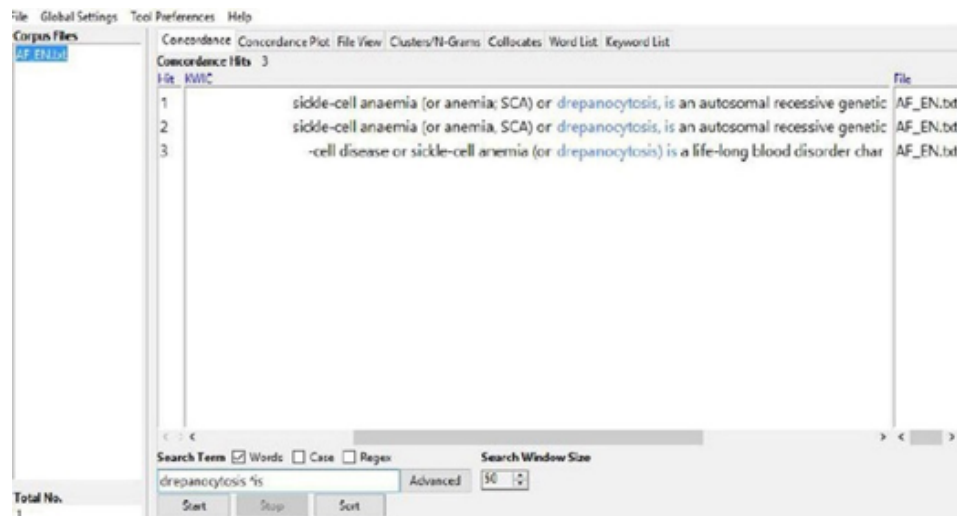
FIGURA 1

Listado de palabras más frecuentes del corpus

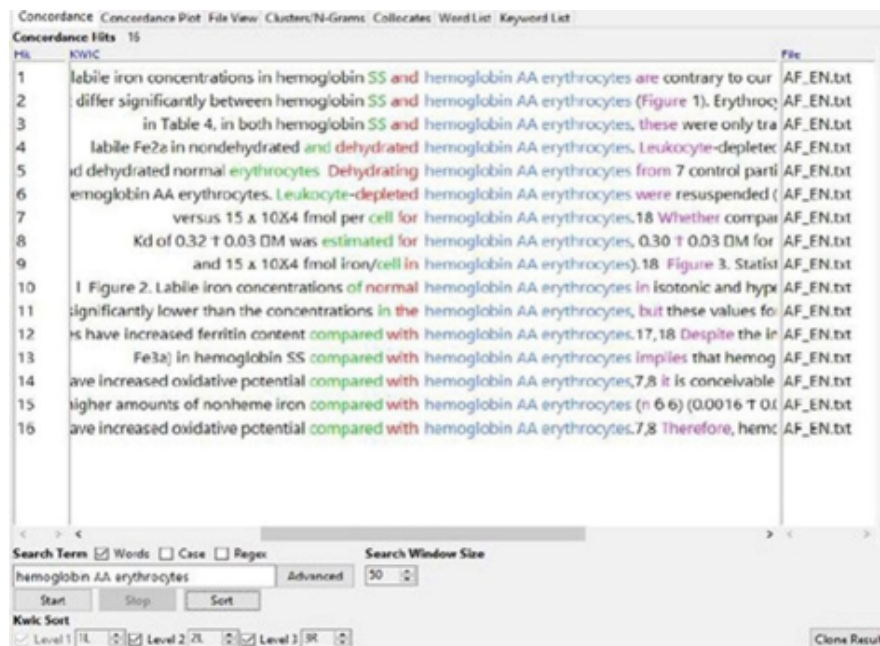
| Rank | Freq | Word | Lemmas Word Form(s) |
|------|------|--------------|---------------------|
| 1 | 618 | sickle | |
| 2 | 580 | cell | |
| 3 | 425 | blood | |
| 4 | 366 | hemoglobin | |
| 5 | 350 | cells | |
| 6 | 240 | disease | |
| 7 | 219 | red | |
| 8 | 209 | hb | |
| 9 | 208 | haemolysis | |
| 10 | 191 | iron | |
| 11 | 181 | patients | |
| 12 | 179 | haemoglobin | |
| 13 | 163 | anemia | |
| 14 | 157 | erythrocytes | |
| 15 | 129 | anaemia | |
| 16 | 127 | normal | |
| 17 | 108 | sickling | |
| 18 | 102 | oxygen | |
| 19 | 101 | hbs | |
| 20 | 100 | labile | |

AntConc ofrece también la posibilidad de consultar información conceptual y sinónímica. Para tal fin los alumnos emplearon comodines (*wildcards*), que refinaron las búsquedas. Una vez abierta la sección de *Concordance*, los alumnos introdujeron distintos patrones de búsqueda para buscar definiciones y sinónimos de la palabra *sickle-cell anemia*. Por ejemplo, emplearon los siguientes patrones de búsqueda: *sickle* is*, *sickle-cell anemia*or*, *sickle-cell disease**, *sicke* is called*, *sickle* is known as*. Y como resultado, se recuperaron los patrones definitorios y los siguientes sinónimos: *meniscocytosis*, *sickleemia*, *meniscocytosis*, o *drepanocytosis* y las siglas SCD y SCA (figura 2). Para consultar la definición de los sinónimos encontrados, bastaba con teclear el término en cuestión seguido de **is* (figura 2).

En lo referido a la fraseología, la búsqueda de patrones combinatorios como colocaciones (*collocates*) o conjuntos léxicos (*collocation clusters*) fue de especial utilidad para nuestros alumnos a la hora de identificar la frecuencia de aparición de patrones léxicos concretos. AntConc permite agrupar las colocaciones de cada término a la derecha o a la izquierda por orden de frecuencia, característica especialmente útil en las traducciones a la hora seleccionar un patrón léxico concreto. Por ejemplo, una de las palabras más frecuentes en el corpus compilado por los alumnos fue *erythrocytes*. Gracias a la función *N-Grams/Cluster*, se puede obtener información sobre distintos tipos de eritrocitos. Esta función permitió identificar

FIGURA 2Función *Concordance*

cuatro tipos de eritrocitos: *hemoglobin aa erythrocytes*, *hemoglobin ss erythrocytes*, *sickle erythrocytes* o *infected erythrocytes* además de conocer los patrones combinatorios de cada uno de ellos con la función *Concordance* (figura 3).

FIGURA 3Ejemplo de palabra en su contexto con la función *Concordance*

5. Resultados

Los alumnos realizaron un informe final en donde recogían la utilidad del uso de corpus así como las reflexiones finales. Esto nos sirvió para conocer sus opiniones y valoraciones finales. En términos generales, los comentarios proporcionados por los alumnos fueron positivos y nos alentaron a continuar con esta metodología. Los discentes apreciaron la utilidad de las distintas funciones que el programa de concordancias AntConc les ofrecía (y que en su mayoría desconocían). Compilar y analizar un corpus les proporcionó una mayor confianza en cuanto a sus destrezas técnicas y actitud ante diversos problemas traductológicos. Al mismo tiempo, los alumnos destacaron el uso de corpus especializados como recurso lexicográfico adicional a los recursos tradicionales (bases de datos, diccionarios especializados o glosarios especializados). Ofrecemos a continuación algunos de los comentarios emitidos por los alumnos:

Alumno 1: La realización de la presente tarea nos ayuda a comprender la importancia de los distintos recursos con los que ha de trabajar todo profesional de la Traducción e Interpretación que se precie.

Alumno 2: Gracias a todos estos recursos que nos permiten gestionar la terminología, el traductor optimiza los esfuerzos documentales y es capaz de mantener la consistencia y la coherencia a lo largo de sus diferentes encargos.

Alumno 3: Desconocía este tipo de recursos. Su utilidad es enorme para traducir, aunque hay que aprender a utilizarlo y se necesita tiempo para la compilación.

6. Conclusiones

Hemos querido con este artículo mostrar cómo se ha utilizado la metodología de corpus (creación de corpus y análisis con un programa de concordancias) en el aula de TISP (ámbito sanitario), en concreto en la asignatura de Técnicas y Recursos en TISP dentro del Máster en Comunicación Intercultural, Traducción e Interpretación en los Servicios Públicos (Universidad de Alcalá). Como ha quedado resaltado a lo largo del presente trabajo, el uso de un corpus virtual se muestra como una herramienta necesaria y útil para la enseñanza de terminología, patrones colocacionales, lenguaje formulaico y fraseológico. Además, se trata de un recurso poco explorado en el aula de la TISP, donde los alumnos proceden de contextos académicos diversos. No hay que olvidar que, si bien el corpus virtual como herramienta pedagógica es indiscutible, la formación es esencial: se requieren una serie de destrezas técnicas y una inversión de tiempo considerable. No obstante, una planificación didáctica adecuada, basada en una instrucción previa, y el diseño de un protocolo de compilación como el ofrecido en este trabajo (fase de documentación, fase de compilación y fase de análisis) pueden contribuir, sin duda, al empleo de esta herramienta en el aula de la TISP como andamiaje documental que facilite los procesos de trasvase.

La puesta en marcha de este tipo de actividades también ha servido como punto de reflexión para el alumno, ya que le ha permitido trabajar (y evaluar) una herramienta de documentación. La tarea descrita, además, ha permitido a los alumnos tomar conciencia de la utilidad de buscar y compilar textos para, posteriormente, trabajar con ellos y elaborar recursos, como, por ejemplo, glosarios o memorias de traducción.

7. Bibliografía citada

ABRIL MARTÍ, María Isabel, 2006: *La interpretación en los servicios públicos: caracterización como género, contextualización y modelos de formación. Hacia unas bases para el diseño curricular*. Tesis doctoral, Universidad de Granada.

ABRIL, María Isabel, y Anne MARTIN, 2008: “Profesionalizar en la práctica: la competencia en ISP como base para la formación de profesionales” en Carmen VALERO GARCÉS (ed.): *Investigación y práctica en traducción e interpretación en los servicios públicos*, Alcalá de Henares: Servicio de Publicaciones de la Universidad de Alcalá, 103-120.

ANGELELLI, Claudia, 2006: “Designing curriculum for healthcare interpreter education: A principles approach” en Cynthia ROY (ed.): *New approaches to interpreter education*, Washington D.C.: Gallaudet University Press, 23-46.

BAKER, Mona, 1995: “Corpora in translation studies: An overview and some suggestions for future research”, *Target* 7 (2), 223-243.

BAKER, Mona, 1996: “Corpus-based translation-studies: The challenges that lie ahead” en Harold SOMERS (ed.): *Terminology, LSP, and translation: studies in language engineering in honor of Juan C. Sager*, Ámsterdam/Philadelphia: John Benjamins, 175-186.

BOWKER, Lynne, y Jennifer PEARSON, 2002: *Working with specialized language: A practical guide to using corpora*, Londres: Routledge.

BOWKER, Lynne, 2003: “Corpus-based applications for translator training: Exploring possibilities” en Sylviane GRANGER, Jacques LEROT y Stephanie PETCH-TYSON (eds.): *Corpus-based approaches to Contrastive Linguistics and Translation Studies*, Ámsterdam / Nueva York: Rodopi, 169-183.

CALZADA PÉREZ, María, 2005: “Corpus electrónicos como herramientas de documentación y formación de traductores” en Dora SALES SALVADOR (ed.): *La biblioteca de Babel. Documentarse para traducir*, Granada: Comares, 163-199.

CORPAS PASTOR, Gloria, 2001. “La competencia traductora. A propósito del texto médico” en Carmen VALERO GARCÉS e Isabel DE LA CRUZ CABANILLAS (eds.): *Traducción y nuevas tecnologías. Herramientas auxiliares del traductor. Encuentros en torno a la Traducción*, Alcalá de Henares: Servicio de Publicaciones de la Universidad de Alcalá, 137-164.

CORPAS PASTOR, Gloria, 2004: “La traducción de textos médicos especializados a través de recursos electrónicos y corpus virtuales” en Luis GONZÁLEZ y POLLUX HERNÁNDEZ (eds.): *Las palabras del traductor en Actas del II Congreso Internacional El español, lengua de traducción, 20 y 21 de mayo, Toledo 2004*, Bruselas: Comisión Europea / ESLETRA, 137-164.

CORPAS PASTOR, Gloria, 2008: *Investigar con corpus en traducción: los retos de un nuevo paradigma*, Fráncfort: Peter Lang.

CORPAS PASTOR, Gloria, y Miriam SEGHIRI, 2009: “Virtual corpora as documentation Resources: Translating travel insurance documents (English–Spanish)” en Alison BEEBY, Patricia RODRÍGUEZ INÉS y Pilar SÁNCHEZ-GIJÓN (eds.): *Corpus use and translating. Corpus use for learning to translate and learning corpus use to translate*, Ámsterdam / Philadelphia: John Benjamins, 75–107.

EMT EXPERT GROUP, 2016: “Competences for professional translators, experts in multilingual and multimedia communication” [disponible en http://ec.europa.eu/dgs/translation/programmes/emt/key_documents/emt_competences_translators_en.pdf, fecha de consulta: 18 de junio de 2016].

FANTINUOLI, Claudio, y Federico ZANETTIN (eds), 2015: *New directions in corpus-based translation studies*, Berlín: Language Science Press.

GÓMEZ GONZÁLEZ-JOVER, Adelina, y Chelo VARGAS SIERRA, 2004: “Aspectos metodológicos para la elaboración de diccionarios especializados bilingües destinados al traductor” en Luis GONZÁLEZ y POLLUX HERNÁNDEZ (eds.): *Las palabras del traductor en Actas del II Congreso Internacional El Español, Lengua de Traducción, 20 y 21 de mayo, Toledo 2004*, Bruselas: Comisión Europea/ ESLETRA, 365-398.

HURTADO ALBIR, Amparo, 2001: *Traducción y traductología*, Madrid: Cátedra.

HURTADO ALBIR, Amparo, 2007: “Competence-based Curriculum Design for Training Translators”, *The Interpreter and Translator Trainer* 1 (2), 163-195.

KENNY, Dorothy, 2001: *Lexis and Creativity in Translation. A Corpus-based Study*, Manchester: St. Jerome.

KRÜGER, Ralph, 2012: “Working with corpora in the translation classroom”, *Special issue of Studies in Second Language Learning and Teaching* 2 (4), 505-525.

KÜBLER, Natalie, 2012: “Working with corpora for translation teaching in a French-speaking setting” en Ana FRANKENBERG-GARCÍA, Lynne FLOWERDEW y Guy ASTON (eds.): *New trends in Corpora and Language Learning*, Londres: Continuum, 62-80.

LAVIOISA, Sara, 2002: *Corpus-based Translation Studies. Theory, Findings, Applications*, Ámsterdam / Nueva York: Rodopi.

LEE, David, y John, SWALES, 2006: "A corpus-based EAP course for NNS doctoral students: Moving from available specialized corpora to self-compiled corpora", *English for Specific Purposes* 25, 56-75.

MONTALT RESURRECCIÓ, Vicent, y María GONZÁLEZ DAVIES, 2007: *Medical translation step by step. Learning by drafting*, Manchester: St. Jerome.

NEUNZIG, Wilhelm, 2003: "Tecnologías de la información y traducción especializada inversa" en Dorothy KELLY y otros (eds.): *La direccionalidad en Traducción e Interpretación. Perspectivas teóricas, profesionales y didácticas*, Granada: Atrio, 189-205.

OROZCO JUTORÁN, Mariana, 2012: *Metodología de la traducción directa del inglés al español. Materiales didácticos para traducción general y especializada*, Granada: Comares.

RABADÁN, ROSA, y Purificación FERNÁNDEZ NISTAL, 2002: *La traducción inglés-español: fundamentos, herramientas, aplicaciones*, León: Servicio de publicaciones de la Universidad de León.

PACTE GRUPO DE INVESTIGACIÓN, 2003: "Building a translation competence model" en Favio ALVES (ed.): *Triangulating translation: Perspectives in process oriented research*, Ámsterdam/Philadelphia: John Benjamins, 43-66.

PINTO MOLINA, María, y Dora SALES SALVADOR, 2008: "INFOLITRANS: A model for the development of information competence for translators", *Journal of documentation* 64, 413-437.

RODRÍGUEZ INÉS, Patricia, 2010: "Electronic corpora and other information and communication technology tools. An integrated approach to translation teaching", *The Interpreter and Translator Trainer* 4, 251-282.

SALDANHA, Gabriela, 2004: "Accounting for the exception to the norm: A study of split infinitives in translated English", *Language Matters, Studies in the Languages of Africa* 35, 39-53.

SALES SALVADOR, Dora, 2006: *Documentación aplicada a la traducción: presente y futuro de una disciplina*, Gijón: Ediciones Trea.

SÁNCHEZ GIJÓN, Pilar, 2009: "Developing documentation skills to build do-it-yourself corpora in the specialized translation course" en Alison BEEBY, Patricia RODRÍGUEZ INÉS y Pilar SÁNCHEZ GIJÓN, (eds.): *Corpus use and translating. Corpus use for learning to translate and learning corpus use to translate*, Ámsterdam/Philadelphia: John Benjamins, 109-127.

SÁNCHEZ RAMOS, María del Mar, y Raquel LAZARO GUTIÉRREZ, 2015. "Corpus-based Interpreting Studies and Public Service Interpreting and Translation Training Programs: The case of Interpreters Working in Gender Violence Context" en J. Romero TRILLO (ed.): *Yearbook of Corpus Linguistics and Pragmatics. Current Approaches to Discourse and Translation Studies*, Suiza: Springer International Publishing, 275-291.

SEGHIRI, Miriam, 2011: “Metodología protocolizada de compilación de un corpus de seguros de viajes: aspectos de diseño y representatividad”, *RLA. Revista de Lingüística Teórica y Aplicada* 49, 13-30.

ULRYCH, Margherita, 2000: “Teaching translation into L2 with the aid of multilingual parallel corpora: issues and trends”, *Miscellanea* 4, 58-80.

VALERO GARCÉS, Carmen, 2005: “La difícil tarea de documentarse en traducción e interpretación en los servicios públicos” en Dora SALES SALVADOR (ed.): *La biblioteca de Babel. Documentarse para traducir*, Granada: Comares, 99-121.

VALERO GARCÉS, Carmen, 2013: “Formación de traductores e intérpretes en una sociedad multicultural. El programa de la Universidad de Alcalá, Madrid”, *Cuadernos de Aldeu* 25, 215-238.

VARANTOLA, Kristan, 2003: “Translators and Disposable Corpora” en Federico ZANETTIN, Silvia BERNARDINI y Dominic STEWART (eds.): *Corpora in Translator Education*, Manchester: St. Jerome, 55-70.

VARGAS URPI, Miriam, 2016: “La difícil tarea de dar respuesta a las necesidades de formación en interpretación en los servicios públicos (ISP) en Cataluña: 10 años de avances y retrocesos”, *FITISPos International Journal* 3, 92-103.

VILA BARBOSA, María Magdalena, 2013: “Corpus especializados como recurso para la traducción: análisis de los marcadores de la cadena temática en artículos científicos sobre enfermedades neuromusculares en pediatría”, *Onomazéin* 27, 78-100.

ZANETTIN, Federico, Silvia BERNARDINI y Dominic STEWART (eds.), 2003: *Corpora in Translator Education*, Manchester: St. Jerome.

ZHU, Chunshen, y Hui WANG, 2011: “A corpus-based, machine-aided mode of translator training”, *The Interpreter and the Translator Trainer* 5 (2), 269-291.