

ISSN: 1579-9794

***The nuts and bolts of corpora: secuencia didáctica para el empleo de corpus en la enseñanza de la traducción científico-técnica***

***The nuts and bolts of corpora: a teaching sequence for using corpora in the scientific and technical translation classroom***

MÍRIAM PÉREZ-CARRASCO  
miriamperezcarrasco@gmail.com  
Universidad de Málaga

Fecha de envío: 01/12/2022

Fecha de aceptación: 09/02/2023

**Resumen:** Este trabajo analiza la aplicación de los corpus lingüísticos a la enseñanza de la traducción científico-técnica. En primer lugar, repasaremos algunos trabajos previos sobre experiencias con corpus en el aula de traducción especializada. Seguidamente, presentaremos la secuencia didáctica que hemos planteado en la asignatura Traducción Científico-Técnica I BA-AB (Inglés-Español/Español-Inglés) del Grado en Traducción e Interpretación de la Universidad de Málaga. El objetivo principal de la misma es familiarizar a nuestro estudiantado con los corpus lingüísticos y fomentar así su empleo como recurso para la traducción. Durante las cinco sesiones que hemos dedicado a este tema, el alumnado ha podido conocer mejor qué son los corpus y cómo pueden beneficiarse de ellos como traductores. Para ello, han trabajado, en primer lugar, con corpus de referencia disponibles en Internet, tanto en español (CREA y CORPES XXI) como en inglés (BNC y COCA); tras ello, han explorado los diferentes corpus a los que se puede acceder a través de Sketch Engine, así como las numerosas funciones que ofrece esta herramienta de análisis de corpus. Por último, han compilado un corpus *ad-hoc* técnicos que posteriormente han empleado para traducir un fragmento de manual de usuario de motocicletas. Esta experiencia docente sugiere que los corpus constituyen una herramienta muy eficaz para los estudiantes de traducción, pues les permite resolver dificultades lingüísticas muy variadas (terminológicas, conceptuales, gramaticales y estilísticas).

**Palabras clave:** Traducción científico-técnica, Corpus lingüísticos, Didáctica de la traducción científico-técnica, Empleo de corpus para la enseñanza de la traducción, Sketch Engine

**Abstract:** This paper illustrates the application of corpora in the scientific and technical translation classroom. First, we review some previous works on teaching experiences with corpora and specialised translation. Then, we will present the teaching sequence we have applied in the subject Scientific and Technical Translation I BA-AB (English-Spanish/Spanish-English) of the BA in Translation and Interpreting at the University of Malaga. The main objective of this teaching experience was to familiarise our students with linguistic corpora and thus encourage them to use corpora as a translation tool. During the five sessions we dedicated to this topic, the students were able to learn more about what corpora are and how they can benefit from them as translators. First of all, they worked with reference corpora available on the net, both in Spanish (CREA and CORPES XXI) and in English (BNC and COCA). Then, they explored the different corpora accessible from Sketch Engine, as well as the numerous functions offered by this corpus analysis tool. Lastly, they compiled an ad-hoc technical corpus which they later used to translate an extract of a motorbike user's manual. This teaching experience suggests that corpora are a very effective tool, as they enable translation students to solve a wide range of linguistic difficulties (terminological, conceptual, grammatical, or stylistic).

**Keywords:** Scientific and technical translation, Linguistic corpora, Scientific and technical translation teaching, Scientific and technical translation teaching with corpora, Sketch Engine

#### INTRODUCCIÓN

Tal y como subrayan diversos autores (Schäffner, 1998, p. 83; Gallego-Hernández, 2015, p. 113; Biel, 2017, p. 317), el empleo de textos paralelos<sup>1</sup> no es algo reciente. Sin embargo, hoy en día, y gracias a los avances tecnológicos, es posible potenciar los beneficios que estos documentos ofrecen, pues podemos emplearlos como complemento a algunas de las fuentes lexicográficas como son los diccionarios o glosarios. Así pues, desde el traductor profesional hasta el aprendiz de traducción, todos pueden acceder a través de Internet a una cantidad infinita de documentos útiles en su quehacer diario. Es aquí donde entra en juego la Lingüística de Corpus, pues permite recoger una muestra de textos, seleccionados siguiendo los distintos parámetros que hayamos establecido, que posteriormente puede ser analizada y explotada de forma (semi)automática a partir de los diferentes programas de gestión de corpus de los que disponemos hoy en día. De esta

---

<sup>1</sup> Se entiende como «texto paralelo» aquellos «relacionados con el texto de partida que aportan información sobre las convenciones de ese tipo de texto o las particularidades del uso lingüístico específico de ese campo: terminológicas, colocacionales, fraseológicas, sintácticas, etc.» (Sánchez-Gijón, 2009, p. 113).

forma, combinamos las ventajas para la documentación que nos ofrece la consulta de textos paralelos, con los beneficios que aportan los programas de gestión de corpus en cuanto a la cantidad de información que podemos analizar, el nivel de sofisticación de las búsquedas y el poco tiempo que nos requieren todas estas tareas que antes solo podían realizarse de forma manual. Como señala Borja Albi (2019, p. 23), actualmente podemos tanto compilar nuestros propios corpus, como acceder a miles de corpus de distintos tipos y sobre diferentes temáticas ya compilados y puestos a disposición del público para poder ser explotados mediante las diferentes herramientas de gestión de corpus.

Esta misma autora (Borja Albi, 2019, pp. 24-25) distingue cinco áreas de aplicación de la Lingüística de Corpus dentro de los estudios en Traducción. En primer lugar, encontramos trabajos —tales como los realizados por Baker (1995), Olohan (2004), Corpas Pastor (2008) o Laviosa (2010)— en los que la Lingüística de Corpus permite el estudio del comportamiento de los textos traducidos y la identificación de posibles normas de traducción. En segundo lugar, Borja Albi (2019) destaca las investigaciones dedicadas al empleo, en la práctica profesional, de lo que denomina *disposable corpora*<sup>2</sup>, con el fin de obtener información terminológica, fraseológica o textual útil para la traducción. Esto ha sido objeto de estudio para autores como Bowker (1998), Zanettin (2002), Bernardini y Castagnoli (2008) o Gallego-Hernández (2015). En tercer lugar, autores como Czulo y Hansen-Schirra (2017) han utilizado recursos y datos basados en el análisis de corpus lingüísticos para diferentes tareas relacionadas con el procesamiento del lenguaje natural, como la generación de memorias de traducción o el desarrollo de motores de traducción, entre otras. En cuarto lugar, Borja Albi (2019) destaca los trabajos de autores como Aston, Bernardini y Stewart (2004) o Sinclair (2004), pioneros en la compilación y el análisis de *language learner corpora* (o corpus de aprendices), es decir, aquellos integrados por textos producidos por los propios estudiantes de lenguas extranjeras; este tipo de corpus podrían ser útiles como material para la enseñanza de idiomas. La última aplicación de los corpus lingüísticos que propone Borja Albi (2019) es el empleo de corpus como recurso en el aula de traducción. Este tema ha sido objeto de estudio en los trabajos realizados por autores como Zanettin (1998), Aston (1999),

---

<sup>2</sup> Se trata de un corpus formado por textos que se encuentran disponibles en la red y que surge de la «necesidad perentoria de acceder a documentación específica sobre un tema concreto a la mayor brevedad posible», como subraya Corpas Pastor (2001, pp. 164-165). Otros autores proponen nombres diversos para esta misma tipología de corpus; a saber, *special purpose corpus* (Pearson, 1998), *customized corpus* (Austermühl, 2001), *do-it-yourself (DIY) corpora* (Zanettin, 2002), *corpus ad-hoc* (Aston, 1999; Corpas Pastor, 2001) o *corpus virtual* (Seghiri, 2006, 2011, 2017).

Varantola (2003), Rodríguez-Inés (2008), Bowker, McBride y Marshman (2008), Marco Borillo y Van Lawick (2009), Bernardini (2016), Rodríguez-Inés y Gallego-Hernández (2016) o Biel (2017). La aplicación de los corpus lingüísticos para la enseñanza de la traducción será también el tema de investigación del presente estudio.

Ya en la década de los 80, principio de los 90, algunos autores empezaban a indagar sobre las ventajas que podría suponer para los estudiantes trabajar con textos auténticos a través de datos lingüísticos (Biel, 2017, p. 317). Después de estos primeros acercamientos a la aplicación de los corpus en la enseñanza de idiomas en general, llegaron los experimentos en el aula de traducción en particular. Así, Bowker (1998) empleó con sus estudiantes un corpus monolingüe en lengua materna para realizar una traducción directa (FR>EN) de un texto especializado, en concreto un artículo sobre escáneres ópticos; a través de esta investigación, constató que el uso de este corpus había supuesto una mejor comprensión del texto, la terminología y las expresiones idiomáticas, lo cual aumentaba en gran medida la calidad de las traducciones propuestas por los aprendices. Posteriormente, Aston (1999) subrayaría las ventajas del uso de los corpus tanto para traducir como para enseñar a traducir; comenzaron así a celebrarse eventos científicos pioneros en esta temática (como la primera edición del Congreso Corpus Use and Learning to Translate, CULT); y se editaron los primeros volúmenes centrados en la aplicación de los corpus en el proceso de enseñanza-aprendizaje de la traducción (Bernardini y Zanettin, eds., 2000; Zanettin, Bernardini y Stewart, eds., 2003). Tras ellos, han sido numerosos los autores que han tratado de acercar al aula este tipo de recursos, especialmente cuando se trata de enseñar traducción especializada (Sánchez-Gijón, 2004; Monzó-Nebot, 2008; Rodríguez-Inés, 2008; López-Rodríguez y Buendía-Castro, 2011; Kübler, 2011; Krüger, 2012; Gallego-Hernández, 2015; Biel, 2017; Borja Albi, 2019).

Apoyándonos en los resultados mayoritariamente positivos de estas investigaciones, en el presente trabajo se recoge una secuencia didáctica llevada a cabo con alumnado de la asignatura Traducción Científico-Técnica I (BA-AB) Inglés-Español/Español-Inglés, del tercer curso del Grado en Traducción e Interpretación de la Universidad de Málaga durante el curso 2019/2020. En este caso, proponemos introducir a nuestro estudiantado en la Lingüística de Corpus para, posteriormente, poder emplear diferentes corpus (de referencia, especializados y *ad-hoc*) en sus correspondientes tareas de traducción especializada (científico-técnica), de forma que podamos analizar qué beneficios pueden aportar estos recursos tanto en su proceso de aprendizaje como en su futuro desempeño profesional.

## 1. CONTEXTO Y TEMPORALIZACIÓN

Como se ha mencionado anteriormente, la secuencia didáctica presentada en el presente trabajo se ha llevado a cabo en el aula de la asignatura Traducción Científico-Técnica I (BA-AB) Inglés-Español/Español-Inglés. Dentro del actual Plan de Estudios del Grado en Traducción e Interpretación de la Universidad de Málaga, el alumnado empieza a cursar asignaturas de traducción general en sus lenguas A y B durante el primer y segundo semestre del segundo curso (Traducción General BA-AB I y II). Una vez que ya se han familiarizado con los principios básicos de la traducción y con la traducción de textos generales, en el tercer curso del Grado se enfrentan a cuatro asignaturas obligatorias de traducción especializada, siempre en sus lenguas A y B. Durante el primer semestre, abordan las asignaturas Traducción Humanística BA-AB (I) y Traducción Científico-Técnica BA-AB (I); y, durante el segundo semestre, Traducción Jurídica y Socioeconómica BA-AB (I) y Traducción Audiovisual BA-AB (I). Para terminar, en el cuarto y último curso, los estudiantes deben elegir uno de los tres itinerarios propuestos: interpretación de conferencias, traducción general o traducción especializada. En estos dos últimos itinerarios, los alumnos cursan materias sobre traducción especializada, tanto en lengua B como C.

Son varias las razones que nos han llevado a aplicar la Lingüística de Corpus en la asignatura de Traducción Científico-Técnica (I). Por una parte, creemos que, temporalmente, el estudiantado se encuentra en un momento óptimo para acercarse a nuevos recursos que puedan ayudarles en sus tareas de traducción. Siendo una materia del primer semestre del tercer curso, los estudiantes ya están familiarizados con los principios básicos de la traducción general, así como con algunos recursos informáticos y terminológicos aplicados a la traducción, por lo que cuentan con nociones suficientes para sacar provecho del empleo de los corpus sin que estos se conviertan en un obstáculo más.

Por otra parte, hemos tenido en cuenta cómo está estructurada la asignatura en la que se ha aplicado la secuencia didáctica. En este sentido, cabe mencionar que la asignatura que aquí nos ocupa, Traducción Científico-Técnica (I) está organizada en tres grandes bloques temáticos, ordenados como sigue:

- Bloque temático 1. Este bloque sirve de introducción a la asignatura. A partir de los conocimientos previos con los que cuenten los estudiantes, se plantea la metodología de enseñanza y aprendizaje con la que se trabajará en clase. Los objetivos de este bloque son: familiarizarse con los conceptos básicos de la traducción especializada, con especial referencia

a la traducción científico-técnica; conocer el mercado de trabajo (analizando la demanda laboral y preparando plantillas para presupuestos y facturas que simulen modelos profesionales reales), así como localizar y manejar fuentes documentales genéricas, tanto electrónicas como en papel.

- Bloque temático 2. A lo largo de este bloque temático, los estudiantes se aproximan por primera vez a la traducción técnica, tanto desde un punto de vista teórico como práctico. Los objetivos que se plantean son los siguientes: conocer fuentes documentales, tanto electrónicas como en papel, para la traducción técnica; caracterizar el discurso técnico y sus realizaciones textuales en español y en inglés; manejar herramientas básicas de traducción en diferentes formatos y soportes; conocer y solucionar las dificultades que plantea la traducción técnica y realizar simulaciones de encargos de traducción técnica en la combinación lingüística inglés-español, llevando para ello a cabo tareas de documentales y terminológicas, tareas de gestión, traducción, revisión y control.
- Bloque temático 3. Los objetivos que se plantean a lo largo de este último bloque son idénticos a los del bloque 2, pero aplicados a la traducción científica, en lugar de técnica.

En concreto, la secuencia didáctica que se plantea en el presente trabajo se desarrolló durante la última semana del bloque temático 2. Hasta dicho momento, los estudiantes ya habían realizado dos simulaciones de encargo de traducción sobre textos técnicos. Cabe destacar asimismo que, para la realización de cada una de estas tareas, el alumnado se organizó autónomamente en grupos de cuatro o cinco miembros, quienes irían desempeñando, de forma rotatoria, los diferentes roles que debían adoptar en cada encargo de traducción, intentando reproducir así una futura realidad profesional. Estos roles eran: gestor de proyectos, terminólogo, traductor y revisor. A grandes rasgos, el cometido de cada uno de los roles era el siguiente: el gestor de proyectos debía crear un hilo de mensajes en el que repartiese las tareas al resto del grupo, incluyendo las instrucciones y la fecha límite de cada una de ellas; por su parte, el terminólogo debía entregar una tabla terminológica, en la que se recogiesen los términos claves del TO, su definición, el equivalente (o equivalentes) que hubiesen encontrado en español y la fuente (o fuentes) que hubiesen empleado para encontrar dichos equivalentes; el traductor (o traductores, dependiendo del tamaño del grupo) entregaría la traducción del TO, así como las dificultades surgidas al traducir

cada término y cómo se habían resuelto; por último, el revisor se encargaría de ayudar al traductor o traductores, además de revisar el TM y maquetar la versión definitiva del mismo. Cada grupo contaba, asimismo, con un foro privado dentro de la página Moodle de la asignatura, de forma que la docente pudiese seguir en todo momento el desarrollo del trabajo y la comunicación que estaba teniendo lugar en cada caso.

Creemos que esta metodología de trabajo beneficia al alumnado, pues lo acerca a la realidad profesional, en la que o bien se enfrentará a uno de estos roles (si trabaja para una agencia o empresa de traducción) o a todos o parte de ellos (si trabaja como traductor autónomo). Además, el desempeño de tareas diferentes en cada encargo no solo ameniza en cierta forma su labor —pues el alumno suele mostrarse curioso por llevar a cabo un rol diferente cada vez—, sino que además favorece el trabajo cooperativo, el respeto a las fechas de entrega para permitir trabajar al resto del grupo y la colaboración entre todos los miembros mediante el foro privado. Por otra parte, y como se ha mencionado anteriormente, la secuencia didáctica que se presenta en este trabajo se ubica al final del bloque temático 2, justo después de los dos primeros encargos de traducción de la asignatura y antes de comenzar el último bloque de la misma. Con esta temporalización, pretendemos, por un lado, que el estudiantado ya tenga cierta práctica con los textos técnicos, de forma que no unan el desconocimiento y el «temor» a un género nuevo (y en la mayor parte de los casos, a temáticas científico-técnicas ajenas a su experiencia o, incluso, intereses), con la inseguridad que suele generar el empleo de un recurso que posiblemente no hayan manejado anteriormente, como son los corpus. Por otro lado, el hecho de que aún tengan por delante otro bloque temático dentro de la asignatura posibilita que pueden empezar a aplicar todo lo aprendido sobre los corpus lingüísticos con otros tipos de textos especializados, así como a plantear en clase las dudas, beneficios o dificultades que hayan ido surgiendo mientras tratan de integrar este tipo de herramientas en su flujo habitual de trabajo.

## 2. EL EMPLEO DE LOS CORPUS LINGÜÍSTICOS EN EL AULA

Son numerosos los investigadores y docentes de traducción que han tratado de acercar a su alumnado a los corpus lingüísticos. Calzada Pérez (2004, p. 121) incluso considera los estudios de traducción basados en corpus una de las actuales siete tendencias en el mundo de la didáctica de la traducción. De acuerdo con Borja Albi (2019, pp. 29-32), este tipo de recurso en particular no solo ayudaría a los estudiantes a familiarizarse con la lengua que estén estudiando, sino que fomentan otras cuatro competencias que pueden resultar fundamentales durante su proceso de aprendizaje. En primer lugar, Borja Albi (2019) considera que el empleo de los corpus lingüísticos mejoraría la competencia estratégica y tecnológica de los aprendices, puesto

que, para trabajar con este tipo de recursos, los estudiantes deben estar ya familiarizados con el uso de otras herramientas como son las hojas de cálculo, bases de datos, sistemas de gestión para traductores, convertidores de formato, extractores terminológicos, buscadores y otras herramientas útiles durante la compilación de los corpus. Coincidimos con esta autora en que realizar labores tecnológicas tan variadas durante su formación ayudará a los estudiantes a estar mejor preparados para salir a un mundo laboral en el que cada vez surgen nuevos perfiles profesionales, muchas veces relacionados con las nuevas tecnologías. La segunda competencia que destaca Borja Albi (2019) es la relacionada con los aspectos extralingüísticos y conceptuales. Aunque su estudio se centre en la traducción jurídica, este aspecto también se aplicaría a la enseñanza de la traducción técnica, como es nuestro caso, pues, tratándose de una área de conocimiento tan especializada, el hecho de que los estudiantes puedan adquirir ciertos conocimientos conceptuales a través de los corpus —o al menos que puedan familiarizarse con dichos conceptos, normalmente desconocidos y alejados de su realidad e intereses— se verá reflejado en una mejor comprensión del TO y, por consecuencia, en un TM más preciso y de mejor calidad. Coincidimos también con el tercer punto que subraya Borja Albi (2019): la competencia lingüística. Sin duda, al trabajar con corpus comparables, los estudiantes podrán extraer la terminología, así como las unidades fraseológicas más frecuentes del género textual con el que estén trabajando, lo cual conllevará una traducción más natural del TO. Además, estarán expuestos a textos reales, lo cual les ayudará a conocer de primera mano las diferentes convenciones textuales de las respectivas lenguas y géneros. Por último, Borja Albi (2019) destaca los beneficios del empleo del corpus en el aula para el desarrollo académico de los estudiantes y su futuro interés como investigadores. Como apunta, cada vez son más los estudiantes de posgrado que investigan sobre lingüística de corpus. Personalmente consideramos que acercarlos a esta disciplina durante su etapa predoctoral podría ayudarlos a conocer desde el inicio de su actividad académica e investigadora todo el potencial que tienen los recursos basados en corpus, fomentando así tanto la investigación en este campo, como su aplicación en el ámbito profesional de la traducción.

Por otra parte, estamos de acuerdo con otros autores (Bernardini, Stewart y Zanettin, 2003, p. 5; Monzó Nebot, 2008, p. 247; Biel, 2017, p. 326; Borja Albi, 2019, p. 35) con respecto a los distintos beneficios que la aplicación de los corpus lingüísticos puede ofrecer desde un punto de vista pedagógico, pues su carácter experimental motiva a los estudiantes a tener un papel más activo, así como una mayor autonomía, dentro de su proceso de aprendizaje.

Algunos investigadores, sin embargo, señalan ciertos inconvenientes o dificultades para trabajar con corpus en el aula de traducción. Varantola (2003), por ejemplo, apunta al tiempo que se requiere para compilar y explotar un corpus; sin embargo, consideramos que los programas de gestión de corpus son cada vez más rápidos y fáciles de usar, aun por quien no sea especialmente ducho en la materia. Si el problema es que el empleo de los corpus no nos aporta soluciones inmediatas, sino resultados que debemos en gran parte interpretar para que puedan resultar de interés, esto significa que estos recursos requieren de un cierto pensamiento crítico a la hora de explotarlos; en este sentido, compartimos la opinión expuesta por Bernardini et al. (2003, p. 11), cuando subrayan lo siguiente: «corpora can and should be employed to *problematise* rather than *simplify* the task of (future) translators. The greatest pedagogic value of the instrument lies, we suggest, in its *thought-provoking*, rather than *question-answering*, potential». Es decir, a pesar de que saber explotar un corpus puede conllevar cierto tiempo, esto repercutirá en el desarrollo del pensamiento crítico de nuestro alumnado, por lo que no debemos verlo como un problema, sino como un gran beneficio. Por otra parte, entendemos que, si acercamos a los futuros traductores a este tipo de recursos desde su etapa como estudiantes, estos no solo adquirirán cada vez una mayor habilidad para emplearlos, sino que además los verán como un complemento a las fuentes de documentación tradicionales en lugar de como un obstáculo o una pérdida de tiempo. Por último, y como indica también Carratalá-Puertas (2015), a veces los traductores —tanto en formación como profesionales— indican no conocer qué son los corpus. Coincidimos en que esto se debe a la indefinición existente acerca de los corpus en el ámbito profesional de la traducción, ya que son muchos los traductores que, a pesar de reconocer que emplear recursos disponibles en línea, como *Linguee* o *Reverso Context*<sup>3</sup>, desconocen que estos están precisamente basados en corpus. Así pues, la formación en materia de corpus no solo promueve el empleo de estos, sino que intenta evitar el desconocimiento y el consecuente rechazo a los mismos que aún hoy sigue existiendo entre aprendices y profesionales de la traducción.

A continuación, presentaremos la secuencia didáctica que se ha aplicado en el aula de traducción científico-técnica con el fin de demostrar cómo podemos acercar a nuestro alumnado a los corpus lingüísticos de forma sencilla y práctica y qué beneficios puede suponer este tipo de actividades tanto para su proceso de aprendizaje, como para sus competencias como futuros traductores.

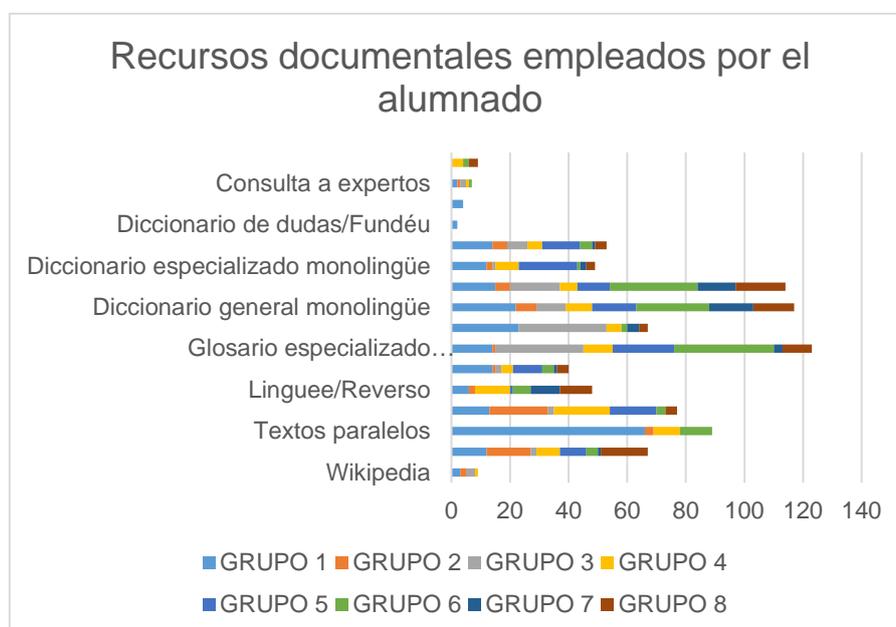
---

<sup>3</sup> Los sitios *Linguee* y *Reverso Context* ofrecen traducciones en contexto a partir de ejemplos reales y están disponibles en los siguientes enlaces: <<https://www.linguee.es/>> y <<https://context.reverso.net/traduccion/>>.

3. SECUENCIA DIDÁCTICA PARA APLICAR LOS CORPUS LINGÜÍSTICOS EN LA ASIGNATURA TRADUCCIÓN CIENTÍFICO-TÉCNICA BA-AB (INGLÉS-ESPAÑOL/ESPAÑOL-INGLÉS)

Como se ha mencionado anteriormente, la secuencia didáctica que se presenta se desarrolló con el alumnado de tercer curso del Grado en Traducción e Interpretación, concretamente en el seno de la asignatura Traducción Científico-Técnica BA-AB (Inglés-Español/Español-Inglés). Los estudiantes ya habían completado el primer bloque, de introducción teórico-práctica a la materia, y estaban acabando el segundo bloque, dedicado a la traducción técnica, al que seguiría el tercer y último bloque, centrado en la traducción médica.

Para la consecución de cada encargo de traducción, cada grupo de trabajo debía entregar, junto con su propuesta de traducción, una tabla terminológica. En esta recogerían los términos que no conociesen de cada texto, su definición, qué equivalentes habían encontrado, en qué fuente o fuentes habían localizado estos equivalentes y cómo habían resuelto las dificultades de traducción que se les había presentado a lo largo del texto. En el Gráfico 1, se incluye el número de ocasiones en que cada grupo manifiesta en su tabla terminológica haber empleado cada tipo de recurso durante su traducción. A partir del mismo, concluimos, en primer lugar, que, para el segundo encargo de traducción técnica propuesto durante la asignatura, nuestro alumnado empleó una gran variedad de recursos documentales. En concreto, el tipo de recurso más habitual fueron los glosarios especializados bilingües o plurilingües (para encontrar posibles equivalentes de términos técnicos), seguidos de cerca por los diccionarios generales monolingües (especialmente para dar con una definición de los términos) y los diccionarios generales bilingües (para llegar a posibles equivalentes de términos que, aun siendo técnicos, no eran de un alto nivel de especialización). Otros recursos empleados por la totalidad de los grupos, aunque en mayor o menor proporción, fueron los diccionarios especializados tanto monolingües como bilingües o plurilingües; los glosarios especializados monolingües; los sitios web especializados; y recursos basados en corpus como *Linguee* o *Reverso Context*. Seguidos de estos recursos especialmente populares entre nuestros estudiantes, encontramos otros como textos paralelos (destaca el gran uso que hizo de ellos el grupo 1), obras de referencia especializadas, el índice de frecuencia en buscadores, bases terminológicas, Wikipedia o la consulta a expertos. Por último, cabe destacar que solamente uno de los grupos manifestó haber empleado algunos de los corpus de la RAE, así como el *Diccionario Panhispánico de las Dudas* para realizar su último encargo de traducción.



**Gráfico 1. Recursos documentales empleados por el alumnado**

Fuente. Elaboración propia

A pesar de la variedad de recursos documentales con la que demostró haberse manejado nuestro alumnado —hecho que nos sorprendió positivamente—, también pudimos comprobar que el único grupo que hizo un uso consciente de los corpus lingüísticos para sus traducciones fue el grupo 1, pues emplearon dos de los corpus que pone a disposición del público la RAE: el CREA y el CORPES XXI. Por su parte, 7 de 8 grupos se sirvieron de *Linguee* o *Reverso Context* para su tarea de documentación; sin embargo, y como veremos posteriormente, en clase reconocieron no saber que estas herramientas están basadas en corpus. Asimismo, la mitad de los grupos (4 de 8) también afirmó haber empleado textos paralelos para su traducción; mientras que 6 grupos incluyeron como recurso documental el índice de frecuencia que arrojan los buscadores. Así pues, podemos concluir que 7 de los 8 grupos se sirvieron de textos reales, ya fuese a partir de los buscadores, de herramientas como *Linguee* o *Reverso Context* o de textos paralelos. Por este motivo, pensamos que los estudiantes sí reconocen, e incluso pretenden beneficiarse, del potencial que tienen los textos reales a la hora de enfrentarse a una traducción, simplemente no conocen qué son los corpus lingüísticos (los cuales les facilitarían en gran medida el empleo de este tipo

de textos) o, si los conocen, no saben o no se atreven a emplearlos mientras traducen.

Este escaso o nulo empleo —al menos de forma consciente— de los corpus lingüísticos por parte de los estudiantes contrasta con la opinión de los investigadores, quienes siguen subrayando los numerosos beneficios que pueden aportar este tipo de recursos para los traductores, tanto por la variada naturaleza de problemas que pueden solucionar —a saber, problemas temáticos, terminológicos, textuales o estilísticos, entre otros—, como por tratarse de un tipo de herramienta económica, rápida y fiable (Seghiri, 2011, 2017). Por esta razón, consideramos que la única forma de poder incrementar el empleo de estos recursos entre nuestro estudiantado es introduciéndolos en la materia tanto de manera teórica, para que entiendan cómo funcionan los corpus lingüísticos y que están realmente detrás de otros recursos que ya conocen, como de manera práctica e instrumental, fomentando que experimenten por sí mismos con estas herramientas, que conozcan cada vez mejor las funcionalidades y beneficios que estos presentan y que reflexionen sobre cómo pueden aplicarlos en su día a día como estudiantes y futuros profesionales de la traducción.

### 3.1. Estructura, objetivos y contenidos

Basándonos en las necesidades observadas tanto en clase como a partir de las tareas presentadas por los estudiantes, y tomando como referencia los trabajos realizados anteriormente por otros autores (Rodríguez-Inés, 2008; Biel, 2017; Borja Albi, 2019), planteamos una secuencia didáctica organizada como se muestra en la siguiente tabla:

- Título: *The nuts and bolts of corpora*: introducción a los corpus lingüísticos y a Sketch Engine para traductores científico-técnicos
- Número de sesiones: 5 sesiones presenciales de 90 minutos cada una + trabajo autónomo (tiempo estimado: 10 h)
- Objetivos de aprendizaje:
  - Al final de la presente secuencia didáctica, el alumnado será capaz de lo siguiente:
    - Conocer los principios básicos de los corpus lingüísticos.
    - Identificar cómo pueden emplear los corpus lingüísticos para resolver problemas de traducción.
    - Manejar herramientas de análisis de corpus.
    - Emplear corpus ya disponibles.
    - Compilar y emplear corpus *ad-hoc*.
    - Resolver problemas de traducción empleando herramientas de análisis de corpus.

<ul style="list-style-type: none"> <li>• Contenidos y secuenciación: <ul style="list-style-type: none"> <li>○ Sesión 1: principios básicos de los corpus lingüísticos. <ul style="list-style-type: none"> <li>▪ Conocimientos previos.</li> <li>▪ Qué son los corpus lingüísticos.</li> <li>▪ Aplicaciones habituales de los corpus lingüísticos.</li> <li>▪ Tipos de corpus.</li> </ul> </li> <li>○ Sesión 2: corpus disponibles en español e inglés. <ul style="list-style-type: none"> <li>▪ Corpus en español (CORDE, CREA, CORPES XXI, CDH, Corpus del Español).</li> <li>▪ Corpus en inglés (BNC, COCA).</li> </ul> </li> <li>○ Sesión 3: herramientas de análisis de corpus (Sketch Engine). <ul style="list-style-type: none"> <li>▪ Sketch Engine: <ul style="list-style-type: none"> <li>• Búsqueda avanzada de corpus: <ul style="list-style-type: none"> <li>○ Corpus generales.</li> <li>○ Corpus especializados.</li> <li>○ Corpus paralelos.</li> </ul> </li> <li>• <i>Word Sketch</i>.</li> <li>• <i>Word Sketch Difference</i>.</li> <li>• <i>Thesaurus</i>.</li> <li>• <i>Keywords</i>.</li> </ul> </li> </ul> </li> <li>○ Sesión 4: construcción de un corpus <i>ad hoc</i>. <ul style="list-style-type: none"> <li>▪ Recogida de textos: <ul style="list-style-type: none"> <li>• Repositorios de textos gratuitos en Internet.</li> <li>• La web como corpus.</li> </ul> </li> <li>▪ Conversión de formato. <ul style="list-style-type: none"> <li>• Formatos más habituales.</li> <li>• <i>Software</i> para la conversión de formato.</li> </ul> </li> <li>▪ Gestión y almacenaje: <ul style="list-style-type: none"> <li>• Nombrar y ordenar lógicamente los archivos de un corpus.</li> <li>• <i>Software</i> para nombrar de forma automática los archivos.</li> </ul> </li> <li>▪ Explotación de un corpus <i>ad-hoc</i>.</li> </ul> </li> <li>○ Sesión 5: resolución de problemas de traducción mediante corpus. <ul style="list-style-type: none"> <li>▪ Selección del o de los corpus más adecuados.</li> <li>▪ Aplicación del corpus durante las tareas de documentación y de traducción.</li> <li>▪ Reflexión sobre los beneficios y desventajas del trabajo con corpus para traductores.</li> </ul> </li> </ul> </li> </ul>
--

**Tabla 2. Esquema de la secuencia didáctica**

Fuente. Elaboración propia

### 3.2. Metodología y resultados

A continuación, por cuestiones de espacio, resumiremos algunas de las actividades y resultados más destacados de cada sesión. Así, la sesión 1, dedicada a los principios básicos de los corpus lingüísticos, comenzó con la puesta en común de los conocimientos previos con los que contaba nuestro alumnado. Para ello, empleamos una presentación interactiva creada a partir del sitio web *Mentimeter*, el cual permite al alumnado acceder a la presentación a través de un código y, a partir de ahí, responder a las diferentes preguntas. Cabe destacar que sus respuestas se comparten de forma anónima y en directo con el resto de participantes, lo cual favorece la interacción y la participación incluso de aquellos estudiantes a los que les supone un mayor esfuerzo dar a conocer su opinión delante del resto de la clase.

En la Imagen 1, podemos apreciar algunas de las respuestas dadas por nuestro alumnado a la pregunta *¿Qué entiendes por «corpus lingüístico»?* Queda patente que parte de nuestro alumnado sí contaba con conocimientos previos acerca de los corpus lingüísticos, en gran parte porque, tal y como explicaron durante la sesión, algunos de ellos habían asistido solo unas semanas antes a un curso de formación (voluntario y de pago) ofrecido en el seno del departamento y enfocado precisamente en el empleo de los corpus lingüísticos para la traducción de textos especializados, en concreto, textos museísticos.



**Imagen 1. Muestra de las respuestas dadas a la pregunta «¿Qué entiendes por corpus lingüístico?»**

Fuente. Mentimeter

En cuanto a la segunda pregunta, *¿Conoces algún corpus lingüístico?*, los resultados vuelven a ser, en parte, positivos, pues nuestros estudiantes mencionaron numerosos corpus. Como se puede observar en la nube de palabras creada por *Mentimeter* y que se incluye en la Imagen 2, los que enumeran con mayor frecuencia son CORPES XXI, CORDE y CREA, es decir, los corpus creados por la RAE. Otro corpus académico que también señalan conocer, aunque con una frecuencia mucho menor, es el CDH, también de la RAE. Asimismo, los estudiantes han incluido entre sus respuestas el corpus OPUS2 (accesible a través de Sketch Engine) y el sitio web Corpus del Español, a partir del cual se accede a diversos corpus, entre

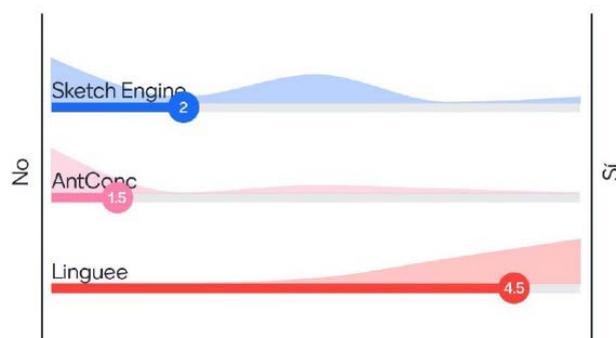
ellos uno de los que también han enumerado, el NOW Corpus. A pesar de que son varios los corpus que parecen conocer, ciertas respuestas sacan a la luz el desconocimiento aún existente entre el alumnado. En concreto, podemos apreciar cómo algunos estudiantes mencionaron IATE o TERMIUM Plus, Index Translationum, TERMITUR y Sketch Engine, cuando ninguno de estos recursos se trata de un corpus lingüístico en sí, sino de dos bases de datos terminológicas, una base de datos de traducciones de libros, un diccionario inteligente terminológico sobre turismo y una herramienta de análisis de corpus, respectivamente.



**Imagen 2. Nube de palabras generada a partir de la pregunta «¿Qué corpus lingüísticos conoces?»**

Fuente. Mentimeter

La última pregunta de la presentación sobre conocimientos previos consistía en responder cuál de las tres herramientas propuestas habían empleado alguna vez y con qué frecuencia. Tal y como se observa en la Imagen 3, casi la totalidad de los estudiantes habían empleado *Linguee* (o herramientas similares, como *Reverso Context*), y además lo hacían con frecuencia, mientras que en el caso de Sketch Engine y AntConc (ambas herramientas de análisis de corpus), la mayor parte de los estudiantes reconocían no haberlas empleado nunca y, aquellos que sí lo habían hecho, lo hacían con muy poca frecuencia, especialmente AntConc.



**Imagen 3. Respuestas a la pregunta «¿Has empleado algunas de estas herramientas alguna vez? ¿Con qué frecuencia?»**

Fuente: Mentimeter

Por su parte, la sesión 2 de esta secuencia didáctica estuvo dedicada a los corpus lingüísticos ya compilados y accesibles a través de Internet, como son los corpus de la RAE u otros corpus en inglés como el BNC o el COCA. Para ello, se presentaron dichos recursos, partiendo de los que ya parecían conocer con anterioridad, como era el caso de los distintos corpus ofrecidos por la RAE. Además, en todo momento se iniciaba con ejercicios aplicados a la lengua general para luego pasar a tareas en los que usaban los corpus para buscar lenguaje técnico que hubiésemos trabajado en los encargos de traducción técnica ya realizados. Esto les ayudaba a reflexionar cómo estas herramientas podrían haberles sido de ayuda a la hora de llevar a cabo su rol de terminólogos o traductores. Podemos citar como ejemplo la actividad que se muestra en la Imagen 4, la cual consistía en analizar las colocaciones del verbo *unscrew* dentro del BNC. Los resultados arrojados les hicieron comprender cómo, a través de los corpus, podrían conocer mejor las diferentes colocaciones de los verbos en cuestión. En este caso concreto, la traducción inmediata que todos propusieron a viva voz en clase para el verbo *unscrew* fue *desatornillar*; sin embargo, como se aprecia en los resultados del BNC, este verbo en inglés no solo aparece junto a *screw*, sino con otros complementos como *cap*, *tap*, *cover* o *nut*, casos en los que el verbo empleado en español no sería *desatornillar*, sino *desenroscar* o *destapar*.



The screenshot shows the British National Corpus (BNC) interface. The top navigation bar includes 'SEARCH', 'FREQUENCY', 'CONTEXT', and 'ACCOUNT'. Below the navigation bar, there are several utility buttons: 'ON CLICK', 'CONTEXT', 'TRANSLATE (??)', 'ENTIRE PAGE', 'GOOGLE', 'IMAGE', 'PRON/VIDEO', and 'BOOK (HELP)'. The main content area displays a table with columns for 'HELP', 'FREQ', 'ALL', '%', and 'MI'. The table lists 14 collocates for the word 'unscrew', with 'CAP' having the highest frequency (6) and 'VALVE' having the lowest (2). Each row also includes a star icon and a horizontal bar chart representing the frequency of the collocate.

HELP	FREQ	ALL	%	MI
1	6	2000	0.30	9.52
2	4	893	0.45	10.10
3	4	2017	0.20	8.93
4	4	10763	0.04	6.51
5	4	47613	0.01	4.37
6	3	637	0.48	10.20
7	3	3863	0.08	7.58
8	3	7193	0.04	6.68
9	3	22271	0.01	5.05
10	2	155	1.29	11.63
11	2	280	0.71	10.78
12	2	563	0.36	9.77
13	2	709	0.28	9.44
14	2	731	0.27	9.39

**Imagen 4. Colocaciones de *unscrew***

Fuente. British National Corpus (BNC)

En definitiva, a través de esta sesión práctica, pudieron, por una parte, conocer algunos de los corpus de referencia disponibles a través de Internet, tanto en inglés como en español; por otra, las diferentes actividades realizadas en clase ayudaron a nuestros estudiantes a comprender que los resultados de los corpus no son inmediatos, sino que sirven como punto de partida de todo un análisis lingüístico que puede resultar muy fructífero. En concreto, este tipo de actividades no solo demuestran cómo podemos emplear los corpus lingüísticos a la hora de traducir, sino que además pueden mejorar la competencia lingüística de nuestro alumnado. A diferencia de cuando emplean un glosario (el cual, basándonos en sus tablas terminológicas parece ser su recurso favorito), los corpus no solo nos ofrecen posibles equivalentes directos y descontextualizados, sino toda una serie de datos lingüísticos a partir de los cuales podemos profundizar cuanto queramos y descubrir, por ejemplo, colocaciones, frecuencia de uso, términos en contexto y otros muchos datos siempre basados en textos reales.

La siguiente sesión, la tercera de cinco, estaría dedicada a Sketch Engine. Se propuso emplear esta herramienta de análisis de corpus, y no otras como AntConc, por varios motivos. En primer lugar, Sketch Engine es una herramienta en línea, por lo que se puede acceder a la misma desde cualquier navegador, con cualquier sistema operativo y sin tener que descargar ningún tipo de *software*. En segundo lugar, nuestros estudiantes contaban con acceso libre gracias a la licencia ofrecida por la universidad. En tercer lugar, Sketch Engine permite tanto crear y explotar nuestros propios corpus *ad-hoc*, como analizar corpus ya compilados (generales y especializados, en diferentes lenguas, etc.) disponibles a través de la propia

herramienta. En cuarto lugar, Sketch Engine ofrece numerosas funcionalidades de especial interés para traductores. En quinto lugar, consideramos que esta plataforma posee una interfaz cada vez más atractiva y sencilla de emplear, lo cual motiva al alumnado y facilita su empleo, sobre todo si consideramos el tiempo limitado con el que contamos para presentar esta herramienta en clase. Por último, los conocimientos previos sobre análisis de corpus con los que partían nuestros estudiantes eran escasos, pero algunos de ellos ya conocían Sketch Engine gracias a su participación en un curso sobre corpus lingüísticos unas semanas antes, por lo que de esta forma podrían poner en práctica lo aprendido, ayudar a los compañeros que no conociesen en absoluto esta herramienta y animarse a emplearla en sus futuros encargos.

Así pues, durante la tercera sesión los estudiantes accedieron a Sketch Engine y, mediante una serie de ejercicios prácticos, pudieron conocer de primera mano las múltiples funcionalidades que ofrece esta herramienta y cómo pueden beneficiarnos como traductores. Las primeras actividades consistieron en localizar, mediante la búsqueda avanzada de Sketch Engine, diferentes corpus que pudiesen serles de especial interés. Así, los estudiantes localizaron, en primer lugar, el corpus con el mayor número de palabras disponible a través de esta herramienta<sup>4</sup>, para luego debatir sobre la importancia del tamaño del corpus frente a otras cuestiones sobre el diseño del corpus, como la temática o el tipo de corpus. Tras ello, pasamos a buscar corpus especializados, por ejemplo, un corpus de lenguaje médico como el Medical Web Corpus, pues esta sería la temática sobre la que traducirían en el próximo bloque de la asignatura; o corpus paralelos, en la combinación EN-ES (como el ya mencionado OPUS2 English) o en otras combinaciones con las que trabajasen (como las versiones del OPUS2 en francés/español o italiano-español). Una vez familiarizados con los diferentes corpus existentes y ya disponibles para explotar gracias a Sketch Engine, pasamos a realizar ejercicios con las siguientes funciones de esta herramienta de análisis de corpus: *Word Sketch*, *Word Sketch Difference*, *Thesaurus* y *Keywords*. Como en ocasiones anteriores, empezábamos con actividades sobre lenguaje general, para luego pasar a ejemplos con lenguaje técnico y, por último, dar la oportunidad al alumnado de pensar en algún caso que se les ocurriese, bien fuese sobre lenguaje técnico o sobre cualquier otro tipo de cuestión. La idea principal es que utilizarasen la herramienta y reflexionasen sobre su posible aplicación en problemas de traducción reales que se les hubiese

---

<sup>4</sup> En el momento en el que se desarrolló la actividad que aquí se describe, se correspondía con el Timestamped JSI web corpus (2014-2019) English, con un total de 40 669 324 757. En el momento de redactar el presente trabajo, este corpus se sigue manteniendo como el de mayor número de palabras accesible desde Sketch Engine, aunque en su versión 2014-2021 ya cuenta con más de 60 billones de palabras (60 409 480 489).

planteado mientras traducían. Las Imágenes 5 y 6 plasman dos de los ejercicios propuestos.

Así, en la Imagen 5, se incluyen los resultados ofrecidos por *Word Sketch Difference* tras comparar los modificadores que aparecen con mayor frecuencia junto a *engine*, frente a los más frecuentes con *motor*. Gracias a esta función de Sketch Engine, los estudiantes constataron que, a pesar de que sendos términos se traduzcan como *motor* en español, en inglés *engine* parece usarse más para los motores de combustión interna (de ahí a que aparezca con modificadores como *steam*, *diesel*, *fire*, *petrol* o *combustión*), mientras que el término en inglés *motor* se emplea, en la mayoría de las ocasiones, para motores eléctricos o híbridos (lo cual explica la frecuencia de los modificadores *electric* o *hybrid*).



modifiers of "engine/motor"			
steam	233	0	...
diesel	121	0	...
fire	174	0	...
diesel	70	0	...
petrol	79	0	...
combustion	62	0	...
outboard	6	29	...
electric	0	59	...
starter	0	21	...
variable-reluctance	0	20	...
stepping	0	33	...
hybrid	0	42	...

**Imagen 5. Resultados de *Word Sketch Difference* para *engine* vs. *Motor***  
Fuente: Sketch Engine

Para concluir la sesión, se pidió como tarea autónoma a los estudiantes pensar, a partir de lo que se había comentado y trabajado en el aula, en cinco formas de aplicar Sketch Engine en la asignatura de Traducción Científico-Técnica. De hecho, en la Imagen 6, se puede apreciar una de las posibles aplicaciones propuesta por una estudiante. En concreto, ella proponía emplear la función *Keywords* de Sketch Engine para la generación de un glosario (semi)automático sobre términos del cáncer. Para ello, había usado un corpus médico disponible en Sketch Engine (en concreto, el Medical Web Corpus) e, investigando de forma autónoma, había encontrado la posibilidad de filtrar los resultados de la función *Keywords*, lo cual permite elegir los

textos del corpus que procedan de un dominio en concreto. En esta ocasión, seleccionó, dentro del Medical Web Corpus, el dominio *Cancerdictionary*. De esta forma, consiguió generar el listado de términos clave que se incluye en la Imagen 6, a partir del cual pudo compilar un pequeño glosario propio, útil tanto para la traducción de textos médicos, en concreto sobre oncología, como, por ejemplo, para la interpretación sobre esta misma temática especializada. Creemos que la propuesta de esta estudiante, como tantas otras propuestas realizadas por otros discentes y que no podemos incluir por cuestiones de espacio, ilustran a la perfección los beneficios que ofrecen a nuestro estudiantado la secuencia didáctica que presentamos en este trabajo, con la que no solo se fomenta el empleo de corpus para su labor como traductores, sino que se promueve el trabajo autónomo de los estudiantes, quienes se esfuerzan por encontrar aplicaciones reales, motivados por sus propias necesidades.

Word	Word	Word	Word
1 carcinoma ***	11 lymphocyte ***	21 tumor ***	31 autologous ***
2 mastectomy ***	12 neoplasm ***	22 pectoralis ***	32 cytotoxic ***
3 antigen ***	13 deoxyribonucleic ***	23 mammary ***	33 antibody ***
4 ductal ***	14 leukocytes ***	24 biopsy ***	34 dcis ***
5 axillary ***	15 epithelial ***	25 intraductal ***	35 radiotherapy ***
6 chromosome ***	16 allele ***	26 cytoplasm ***	36 malignant ***
7 situ ***	17 autosomal ***	27 histological ***	37 neoplastic ***
8 alkylating ***	18 excision ***	28 noncancerous ***	38 dna ***
9 lymph ***	19 hormone-responsive ***	29 papilloma ***	39 tomography ***
10 na ***	20 lobular ***	30 recessive ***	40 polymerase ***

**Imagen 6. Listado de términos clave sobre el cáncer generado mediante la función *Keywords***  
Fuente. Sketch Engine

Llegados a la cuarta sesión, y ya familiarizados tanto con los corpus lingüísticos como con Sketch Engine, el objetivo era que los estudiantes compilasen un corpus *ad-hoc*. Antes de comenzar la tarea, explicamos la importancia que tiene marcar desde el inicio de la compilación del corpus los criterios de diseño con los que guiaremos nuestra labor. En concreto, planteamos los criterios de diseño propuestos por Bowker y Pearson (2002); es decir: a) propósito de la compilación, b) tamaño del corpus, c) medio, d) temática, e) tipo textual, f) autoría, g) fecha de publicación y h) lengua o lenguas de redacción.

En este caso, el *propósito* era compilar un corpus que les ayudase durante la traducción del próximo texto técnico con el que trabajarían (es decir, un fragmento de un manual de usuario de una motocicleta) y cuya temática no encontramos corpus ya compilados y disponibles. El tamaño del

corpus sería reducido, porque el fin de la actividad no era crear un corpus de grandes dimensiones, sino familiarizarse con los pasos básicos para poder compilar su propio corpus. En este sentido, cabe destacar que, a pesar de que Sketch Engine permite trabajar directamente con los textos en formato .pdf, decidimos que nuestros estudiantes compilarían su corpus con textos en formato plano (.txt), dado que este es el formato aceptado por la mayoría de herramientas de análisis de corpus (incluida la herramienta gratuita AntConc). Cabe destacar que los manuales de usuario originales suelen presentarse en formato .pdf, así como con una maquetación (textos en columnas y numerosos gráficos, tablas e ilustraciones) que no suele facilitar la conversión de formato. Así pues, para que la tarea resultase más sencilla, dado que era la primera vez que compilaban un corpus, y se ajustase al limitado tiempo con el que contábamos, propusimos que cada grupo trabajase únicamente con cuatro o cinco textos, uno por cada miembro del equipo; no obstante, podían decidir de forma voluntaria si aumentar el tamaño de su corpus según lo confiados que se sintiesen con las diferentes herramientas. El *medio* del corpus sería escrito; la *temática*, como acabamos de mencionar, las motocicletas; y, *el tipo textual*, los manuales de usuario. Para garantizar que fuesen de *autoría* fiable, bastaba con descargar los textos desde las páginas web de los propios fabricantes. Además, si se trataba de modelos actualmente en venta, se garantizaba que la *fecha de publicación* fuese reciente. Por último, la *lengua* del corpus sería el español, idioma hacia el que traducirían en esta ocasión.

Una vez establecidos los criterios de diseño del corpus, seguiríamos el protocolo de compilación de cinco fases —búsqueda de información, descarga, formato, codificación y almacenamiento— propuesto por Seghiri (2006, 2011, 2017). En concreto, en casa realizarían las fases de búsqueda de información y descarga, siguiendo las instrucciones dadas anteriormente; mientras que en clase nos encargaríamos de trabajar en el formato, la codificación y el almacenamiento de los textos. Para todas esas tareas, se propusieron herramientas gratuitas, de forma que pudiesen emplearlas sin problemas. Así, los estudiantes emplearon el sitio web *PDF to Text* para la conversión de formato y el *software* denominado *LupasRename* para la codificación automática de los textos.

En definitiva, el resultado de dicha sesión fue la compilación de un corpus sobre manuales de usuario de motocicletas en español, de tamaño limitado, pero representativo cualitativamente (Seghiri 2006, 2011 y 2017) y siguiendo en todo momento los criterios de diseño que pueden aplicar en un futuro para compilar corpus de mayor tamaño, o sobre cualquier temática o lengua. Además, para solventar la cuestión del tamaño del corpus, se puso a disposición de los estudiantes un paquete de textos ya formateados y listos

para ser añadidos a los que cada grupo había preparado, de forma que tuviesen un corpus de unas dimensiones adecuadas para poder llevar a cabo la próxima tarea. Esta consistiría en traducir, fuera del aula y por grupos, como habían hecho hasta ahora, un fragmento de un manual de usuario de una motocicleta de inglés a español. Para ello, no debían emplear el corpus como único recurso documental, sino que tenían que integrarlo dentro de sus recursos habituales, buscando así los beneficios que podían aportarles los corpus en combinación con los glosarios, diccionarios y el resto de herramientas que acostumbraban a utilizar hasta la fecha.

De esta forma, la secuencia didáctica que se presenta en este trabajo, concluyó con una última sesión cuyo objetivo era corregir la propuesta de traducción de cada grupo, haciendo hincapié en cómo se habían servido del corpus para solventar las diversas dificultades surgidas mientras trabajaban con el texto en cuestión. Tras esta sesión en clase, cada estudiante, de forma individual y como tarea para casa, debía entregar una reflexión personal sobre la utilidad que había tenido el empleo de un corpus durante su último encargo de traducción, incluyendo tanto los aspectos que consideraban más positivos como los negativos. A continuación, expondremos las conclusiones a las que hemos llegado a partir de los comentarios y las reflexiones realizadas por el alumnado.

#### 4. ANÁLISIS CUALITATIVO DE LA OPINIÓN DEL ALUMNADO

Para valorar los resultados del presente estudio, se pidió a los estudiantes que entregaran una reflexión personal y de carácter cualitativo al terminar la secuencia didáctica aquí presentada. En esta debían reflejar cómo habían empleado Sketch Engine (y sus diferentes funcionalidades) en su proceso de documentación y traducción, así como su opinión personal sobre la utilidad de implementar este tipo de contenidos y herramientas sobre gestión de corpus en una asignatura como Traducción Científico-Técnica I (BA-AB) Inglés-Español/Español-Inglés. A continuación, se incluyen algunos de los comentarios realizados por los estudiantes.

En primer lugar, cabe destacar que, en palabras de una estudiante, Sketch Engine es «una herramienta que permite consultar las palabras y colocaciones o frases en su contexto de forma rápida y muy avanzada». Asimismo, las funciones que el alumnado más mencionó haber empleado durante su tarea fueron *Concordance*, *Word Sketch* y *Word Sketch Difference*. En palabras de una de las estudiantes, estas funciones le resultaron «muy útiles en muchos aspectos de la traducción, como concordancias, terminología, contexto [...]». En especial, valoran encontrar «los términos en su contexto». Otro de los estudiantes subraya lo siguiente: «una de las ventajas es poder entender cómo funcionan los términos que no

conozco, o sobre los que no sé mucho, en textos reales, y no de forma aislada en un glosario del que me tengo que fiar». Así pues, parece que los estudiantes aprecian la fiabilidad de una herramienta basada en textos reales, frente a otras donde los términos y sus equivalentes aparecen de manera descontextualizada. Por último, cabe destacar que otro de los estudiantes incluso afirmaba en su reflexión final sobre la secuencia didáctica lo siguiente: «Sketch Engine es la herramienta más útil para un traductor que he visto en mi vida [...]. Me parece muy efectivo para el futuro de un traductor que vaya a trabajar en diferentes ámbitos y estoy seguro de que voy a seguir utilizándola».

Si analizamos los comentarios negativos por parte de los estudiantes, encontramos, sobre todo, críticas sobre el hecho de no contar ya con corpus disponibles a través de Sketch Engine sobre temáticas técnicas, pues, a tenor de los estudiantes, «lleva tiempo compilar tu propio corpus». Asimismo, otros alumnos añadieron que el corpus compilado «no podía solucionar todas las dudas». Consideramos, no obstante, que ningún recurso es capaz de contener la solución a todos los problemas, sino que el buen traductor es aquel que combina los diferentes recursos, aprovechando las ventajas y las respuestas que nos ofrece cada uno de ellos. Otra crítica que consideramos estar causada por el desconocimiento por parte del alumnado es que «el corpus consumía casi la totalidad del espacio disponible en Sketch Engine», pues, aunque fuese el caso, los estudiantes cuentan con la posibilidad de solicitar mayor espacio de almacenamiento en la herramienta si lo necesitan.

Otros comentarios negativos sí que nos ayudan, sin embargo, a plantearnos mejoras para el futuro. Por ejemplo, un estudiante apuntaba como crítica que le «hubiera gustado buscar sinónimos del término *portafusibles* mediante la función *Thesaurus*. Sin embargo, mi corpus no ha encontrado ninguno», mientras que otro señala que el programa «no permite incluir más idiomas». En estos casos, los problemas están de nuevo basados en su desconocimiento: por un lado, es normal no encontrar sinónimos para términos tan técnicos y, por otro lado, la herramienta permite tanto añadir corpus en otros idiomas, como compilar un corpus paralelo en dos idiomas. Por último, otro estudiante lamenta «no haber tenido más tiempo para seguir explorando otras funciones más avanzadas, sobre todo para crear nuestros propios corpus»; esta opinión nos ayuda a plantear posibles mejoras en la temporalización o metodología de nuestra secuencia didáctica.

## CONCLUSIONES

A menudo es complicado encontrar la forma y, especialmente, el tiempo para incluir nuevos contenidos en nuestras asignaturas. Sin embargo,

creemos que es fundamental no solo actualizar continuamente nuestra forma de trabajar, sino también llevar parte de nuestra investigación a nuestras propias aulas.

En esta ocasión, hemos querido plasmar la secuencia didáctica que aplicamos en la asignatura Traducción Científico-Técnica I (BA-AB) Inglés-Español/Español-Inglés, cuyo objetivo era acercar a nuestros estudiantes de manera teórica, pero sobre todo práctica, a la lingüística de corpus, para animarlos no solo a emplear corpus lingüísticos ya compilados, sino además a compilar los suyos propios, y que estos puedan servirles como recurso documental a la hora de enfrentarse a una traducción especializada (en este caso, de un texto técnico), con la esperanza de que no solo les sea útil como estudiantes de traducción, sino también como futuros profesionales del sector. En concreto, los estudiantes han empleado diferentes corpus de referencia disponibles en internet para la lengua española (CREA, CORPES XXI) e inglesa (BNC, COCA). Asimismo, han conocido la herramienta de análisis de corpus Sketch Engine y todas las funcionalidades que esta ofrece, ya sea para el análisis de los corpus disponibles a través de esta plataforma como para compilar y analizar un corpus *ad-hoc*.

Los alumnos han demostrado, a través de sus comentarios y sus ejemplos de uso reales, que las herramientas de análisis de corpus como Sketch Engine pueden ser de gran utilidad a la hora de resolver las dudas de corte terminológico, conceptual, gramatical o estilístico que se le plantean al traducir textos técnicos. No obstante, también han expresado críticas constructivas sobre la temporalización o la metodología seguida. Atendiendo a las mismas, creemos conveniente para ocasiones futuras dedicar más tiempo a explorar las diferentes funciones de Sketch Engine, así como mejorar la coordinación con el resto de docentes del Grado para poder introducir al estudiantado a las herramientas dedicadas al análisis de corpus lingüísticos durante la asignatura sobre recursos informáticos aplicados a la traducción y a la interpretación que forma parte del plan de estudios del Grado en Traducción e Interpretación.

Con respecto al tiempo que podemos dedicar a Sketch Engine, y dado que en clase siempre estamos limitados, una opción podría ser preparar videotutoriales con los que los estudiantes pudiesen seguir reforzando o ampliando desde casa los conceptos trabajados en el aula. Otra idea surgida a partir del *feedback* de los estudiantes es poner a disposición del alumnado distintos materiales que les permita ampliar los contenidos vistos en clase si así lo desean. Algunos de los contenidos de ampliación podrían ser, por ejemplo, una breve introducción teórica y práctica al *software* para el cálculo de la representatividad de un corpus o un tutorial para el etiquetado de textos o la alineación de textos, entre otros.

En definitiva, podemos concluir que la aplicación de los corpus en el aula de traducción especializada ha servido como herramienta para motivar a nuestro alumnado y descubrirles un tipo de recurso documental que, o bien no conocían, o conocían de manera superficial y no se atrevían aún a emplearlo durante sus traducciones. En general, creemos que se ha cumplido este objetivo y que esta secuencia didáctica ha servido para que los estudiantes mejoren su capacidad de análisis lingüístico y de trabajo autónomo a partir de la exploración de una herramienta que esperamos que pueda serles de interés tanto académico como profesional.

En un futuro, nos planteamos proponer el mismo fragmento de texto técnico con dos grupos docentes: en el grupo de control antes de haber implementado esta secuencia didáctica basada en los corpus lingüísticos y, en el grupo experimental, después de la misma. Posteriormente, compilaremos un corpus de aprendices a partir de las traducciones propuestas por los propios estudiantes, de forma que podamos confrontar los resultados de sendos grupos y analizar así en qué medida afecta el empleo de corpus lingüísticos como recurso documental en la calidad de las traducciones realizadas por nuestro alumnado. Por último, siguiendo lo realizado en Pérez-Carrasco y Miriam Seghiri (2022), este nuevo corpus de aprendices podría servirnos para analizar los errores más frecuentes y plantear así distintas propuestas para reforzar aquellos puntos más débiles entre nuestro alumnado.

#### AGRADECIMIENTOS

La presente investigación ha sido realizada en el seno de la red temática T3 de la Universidad de Málaga, así como en el marco de los proyectos INMOCOR (Ref. P20\_00109, PAIDI-Junta de Andalucía), VIP II (Ref. PID2020-112818GB-I00), TRIAJE (Ref. UMA18-FEDERJA-067) y El uso de TICs y corpus aplicado al discurso jurídico (UVa). Asimismo, este estudio se ha realizado gracias a un contrato FPU (Formación del Profesorado Universitario), código FPU18/00206, concedido por el Ministerio de Ciencia, Innovación y Universidades a Míriam Pérez-Carrasco. Por último, este trabajo se ha desarrollado en el marco de las actividades del grupo de investigación LexyTrad y del Instituto Universitario de Investigación de Tecnologías Lingüísticas Multilingües de la Universidad de Málaga (IUITLM), de la que la autora es miembro.

#### REFERENCIAS BIBLIOGRÁFICAS

Aston, G. (1999). Corpus use and learning to translate. *Textus*, 12, 289-314. <https://godzilla.sslmit.unibo.it/~guy/textus.htm>.

- Aston G, Bernardini, S. y Stewart, D. (2004). *Corpora and Language Learners*. Amsterdam, John Benjamins.
- Austermühl, F. (2001). *Electronic Tools for Translators*. St. Jerome.
- Baker, M. (1995). Corpora in Translation Studies: An Overview and Some Suggestions for Future Research. *Target*, 7(2), 223-243. <https://doi.org/10.1075/target.7.2.03bak>.
- Bernardini, S. y Castagnoli, S. (2008). Corpora for translator education and translation practice. En Yuste Rodrigo, E. (Ed.), *Topics in Language Resources for Translation and Localisation* (pp. 39-55). John Benjamins.
- Bernardini, S., Stewart, D. y Zanettin, F. (2003). Corpora in Translator Education. An Introduction. En Zanettin, F., Bernardini, S. y Stewart, D. (Eds.), *Corpora in Translator Education* (pp. 1-13). St. Jerome.
- Bernardini, S. y Zanettin, F. (Eds.) (2000). *I corpora nella didattica della traduzione. Corpus and learning to translate*. CLUEB.
- Bernardini, S. (2016). Discovery Learning in the Language-for-Translation Classroom: Corpora as Learning Aids. *Cadernos de Tradução*, 36 (n.º especial 1), 14-35. <https://doi.org/10.5007/2175-7968.2016v36nesp1p14>.
- Biel, Ł. (2017). Enhancing the communicative dimension of legal translation: comparable corpora in the research-informed classroom. *The Interpreter and Translator Trainer*, 11(4), 316-336. <https://doi.org/10.1080/1750399X.2017.1359761>.
- Borja Albi, A. (2019). How corpora can assist legal translation learners: The GENTT TransTools Corpora platform and Sketch Engine. *Quaderns de Filologia: Estudis Lingüistics*, XXIV, 21-38. <https://doi.org/10.7203/QF.24.16297>.
- Bowker, L. (1998). Using Specialized Monolingual Native-Language Corpora as a Translation Resource: A Pilot Study. *Meta: Translators' Journal*, 43(4), 631-651. <https://doi.org/10.7202/002134ar>.
- Bowker, L., McBride, C. y Marshman, E. (2008). Getting more than you paid for? Considerations in integrating free and low-cost technologies in translator training programs. *Redit*, 1, 26-47. [https://riuma.uma.es/xmlui/bitstream/handle/10630/11319/redit1\\_2008\\_segundoart%C3%ADculo.pdf](https://riuma.uma.es/xmlui/bitstream/handle/10630/11319/redit1_2008_segundoart%C3%ADculo.pdf).

- Calzada Pérez, M. (2004). Applying Translation Theory in Teaching. *Perspectives*, 12(2), 119-133. <https://doi.org/10.1080/0907676X.2004.9961495>.
- Carratalá-Puertas, I. (2015). Corpus y traducción profesional, una relación tan omnipresente como invisible. IV Congreso Internacional CULT (Corpus Use and Learning to Translate), mayo de 2015, Alicante.
- Corpas Pastor, G. (2001). Compilación de un corpus ad hoc para la enseñanza de la traducción inversa especializada. *TRANS: Revista de Traductología*, 5, 155-184. <https://doi.org/10.24310/TRANS.2001.v0i5.2916>
- Corpas Pastor, G. (2008). Investigar con corpus en traducción: los retos de un nuevo paradigma. Peter Lang.
- Czulo, O. y Sansen-Schirra, S. (2017). Crossroads between contrastive linguistics, translation studies and machine translation: TC3-II. Translation and Multilingual Natural Language Processing Series. Language Science Press.
- Gallego-Hernández, D. (2015). The Use of Corpora as Translation Resources: A Study Based on a Survey of Spanish Professional Translators. *Perspectives: Studies in Translatology*, 23(3), 375-391. <https://doi.org/10.1080/0907676X.2014.964269>.
- Krüger, R. (2012). Working with corpora in the translation classroom. *Studies in second language learning and teaching*, 4, 505-525. <https://doi.org/10.14746/ssl.t.2012.2.4.4>.
- Kübler, N. (2011). Working with Corpora for Translation Teaching in a French-speaking Setting. En Frankenberg-Garcia, A., Flowerdew, L. y Aston, G. (Eds.), *New Trends in Corpora and Language Learning* (pp. 62-80). Bloomsbury Academic.
- Laviosa, S. (2010). Corpora. En Gambier, Y. y van Doorslaer, L. *Handbook of Translation Studies*, 1 (pp. 80-86). John Benjamins.
- López-Rodríguez, C. y Buendía-Castro, M. (2011). En busca de corpus online a la carta en el aula de traducción científico y técnica. *Trans-kom*, 41, 1-22. [http://www.trans-kom.eu/bd04nr01/trans-kom\\_04\\_01\\_01\\_Lopez\\_Buendia\\_Corpus.20110614.pdf](http://www.trans-kom.eu/bd04nr01/trans-kom_04_01_01_Lopez_Buendia_Corpus.20110614.pdf).
- Marco Borillo, J. y Van Lawick, H. (2009). Using corpora and retrieval software as a source of materials for the translation classroom. En Beeby, A., Rodríguez, P. y Sánchez-Gijón, P. (Eds.), *Corpus Use and Translating* (pp. 9-28). John Benjamins.

- Monzó-Nebot, E. (2008). Corpus-Based Activities in Legal Translator Training. *The Interpreter and Translator Trainer*, 2(2), 221-252. <https://doi.org/10.1080/1750399X.2008.10798775>.
- Olohan, M. (2004). *Introducing Corpora in Translation Studies*. Routledge.
- Pearson, J. (1998). *Terms in Context, Studies in Corpus Linguistics*. John Benjamins.
- Pérez-Carrasco, M. y Seghiri, M. (2022). ESTECNICOR: explotación de un corpus de aprendices para la detección y clasificación de errores colocacionales en la traducción de textos de automoción (inglés-español). En Fernández-Quesada, N. y Rodríguez-Rubio Mediavilla, S. (Eds.), *Detección y tratamiento de erratas: un diagnóstico para el siglo XXI* (pp. 158-184). Editorial Dykinson.
- Rodríguez-Inés, P. y Gallego-Hernández, D. (2016). Corpus Use and Learning to Translate, almost 20 years on. *Cadernos de Tradução*, 36 (n.º especial 1), 9-13. <https://doi.org/10.5007/2175-7968.2016v36nesp1p9>.
- Rodríguez-Inés, P. (2008). *Uso de corpus electrónicos en la formación de traductores (inglés-español-inglés)* [Tesis doctoral. Universitat Autònoma de Barcelona].
- Sánchez-Gijón, P. (2004). *L'ús de corpus en la traducció especialitzada: compilació de corpus ad hoc i extracció de recursos terminològics*. Barcelona, España: Edicions a Petició S. L, Documenta Universitaria.
- Sánchez-Gijón, P. (2009). Developing documentation skills to build do-it-yourself corpora in the specialised translation course. En Beeby, A., Rodríguez-Inés, P. y Sánchez-Gijón, P. (Eds.), *Corpus use and translating* (pp. 109-127). John Benjamins.
- Schäffner, C. (1998). Parallel Texts in Translation. En Bowker, L., Cronin, M. y Kenny, D., *Unity in Diversity? Current Trends in Translation Studies* (pp. 83-90). St. Jerome.
- Seghiri, M. (2006). *Compilación de un corpus trilingüe de seguros turísticos (español-inglés-italiano): aspectos de evaluación, catalogación, diseño y representatividad*. [Tesis doctoral. Universidad de Málaga].
- Seghiri, M. (2011). Metodología protocolizada de compilación de un corpus de seguros de viaje: aspectos de diseño y representatividad. *Revista de lingüística teórica y aplicada*, 49(2), 13-30. <http://dx.doi.org/10.4067/S0718-48832011000200002>.

- Seghiri, M. (2017). Metodología de elaboración de un glosario bilingüe y bidireccional (inglés-español/español-inglés) basado en corpus para la traducción de manuales de instrucciones de televisores. *Babel*, 63(1), 43-64. <https://doi.org/10.1075/babel.63.1.04seg>.
- Sinclair, J. McH. (2004). *How to use corpora in language teaching*. John Benjamins.
- Varantola, K. (2003). Translators and disposable corpora. En Zanettin, F., Bernardini, S. y Stewart, D. (Eds.), *Corpora in Translator Education* (pp. 55-70). St. Jerome.
- Zanettin, F. (1998). Bilingual Comparable Corpora and the Training of Translators. *Meta*, 43(4), 613-630. <https://doi.org/10.7202/004638ar>.
- Zanettin, F. (2002). Corpora in Translation Practice. En *Proceedings of the LREC Workshop. Language Resources for Translation Work and Research*, 10-14.
- Zanettin, F., Bernardini, S. y Stewart, D. (Eds.) (2003). *Corpora in Translator Education*. St. Jerome.

#### Recursos electrónicos

- Anthony, L. (2022). AntConc (Version 4.2.0) [Software]. <https://www.laurenceanthony.net/software>
- Davies, M. (2004). British National Corpus (BNC) (from Oxford University Press). <https://www.english-corpora.org/bnc/>
- Davies, M. (2008). The Corpus of Contemporary American English (COCA). <https://www.english-corpora.org/coca/>
- Davies, M. (s.f.). El Corpus del Español. <https://www.corpusdelespanol.org/>
- Gobierno de Canadá (2023). TERMIUM Plus. <https://www.btb.termiumplus.gc.ca/tpv2alpha/alpha-eng.html?lang=eng>
- Lexical Computing CZ (s.f.). Sketch Engine. <http://sketchengine.eu>
- LexyTrad, Universidad de Málaga (s.f.). TERMITUR. Diccionario inteligente terminológico para el sector turístico. <http://www.lexytrad.es/termitur/termitur.php>
- Mentimeter (s.f.). Mentimeter: Interactive presentation software. <https://www.mentimeter.com/es-ES>

- Real Academia Española (s.f.). Corpus de referencia del español actual (CREA). <https://www.rae.es/banco-de-datos/crea>
- Real Academia Española (s.f.). Corpus del Español del Siglo XXI (CORPES XXI). <https://www.rae.es/banco-de-datos/corpes-xxi>
- Real Academia Española (s.f.). Corpus diacrónico del español (CORDE). <https://www.rae.es/banco-de-datos/corde>
- Real Academia Española (s.f.). Diccionario panhispánico de dudas (DPD). <https://www.rae.es/dpd/>
- Real Academia Española. (s.f.). Diccionario de la lengua española, 23.<sup>a</sup> ed., [versión 23.6 en línea]. <https://dle.rae.es>
- UNESCO (s.f.). Index Translationum. <https://www.unesco.org/xtrans/bsform.aspx>
- Unión Europea (2023). IATE. European Union terminology. <https://iate.europa.eu/home>