

SEGHIRI, Miriam (ed.). *La lingüística de corpus aplicada al desarrollo de la competencia tecnológica en los estudios de traducción e interpretación y la enseñanza de segundas lenguas*. Berlín: Peter Lang, pp. 293. ISBN: 978-3-631-74122-1.

Este volumen, editado por Miriam Seghiri y publicado por la editorial Peter Lang, supone una novedosa contribución a los estudios de la Lingüística de corpus, disciplina que en las últimas décadas ha revolucionado la forma en la que tradicionalmente se realizaban las investigaciones sobre la lengua. En la actualidad, el estudio y análisis de corpus, es decir, de colecciones de textos reales, tanto escritos como orales, en formato electrónico se aplica en numerosos ámbitos: enseñanza de lenguas extranjeras, lexicografía, procesamiento del lenguaje natural, traducción automática o asistida por ordenador, formación de traductores e intérpretes o recursos para el ejercicio profesional de la traducción e interpretación, por citar solo algunos de los más frecuentes.

La obra aquí reseñada se abre con el Prólogo de Pedro Mogorrón Huerta, Catedrático en Traducción e Interpretación de la Universidad de Alicante, en el que describe las virtudes de la obra y destaca «la variada naturaleza de los textos que componen este volumen», además de cómo las aportaciones incluidas consiguen mostrar «las numerosas posibilidades de desarrollo traductológico y didáctico de la lingüística de corpus», disciplina que califica como «una apuesta de futuro». Le sigue la Introducción, llevada a cabo por la propia editora de la obra, en la que nos revela que el propósito principal del trabajo es «entra[r] de lleno en el uso de las TAC [tecnologías del aprendizaje y el conocimiento] basadas en corpus en el aula de traducción, interpretación, así como de enseñanzas de segundas lenguas». Así pues, el volumen queda dividido en dos grandes bloques: uno centrado en el uso de corpus en las aulas de traducción especializada y de interpretación, y otro, orientado a la enseñanza de segundas lenguas a través de la lingüística de corpus. Cada uno de estos bloques contiene textos de características muy diversas, lo cual demuestra una vez más las numerosas aplicaciones que puede llegar tener la lingüística de corpus en estos ámbitos.

El primer bloque se abre con un capítulo de María-Teresa Ortego-Antón y Purificación Fernández-Nistal, titulado «Estudio contrastivo de la terminología de embutidos en inglés y en español con ParaConc y tICorpus a partir del corpus paralelo P-GEFEM y del comparable C-GEFEM». En este caso, las autoras estudian con éxito el comportamiento morfosintáctico de

léxico especializado y su fraseología en el ámbito de las industrias de los embutidos, para el par lingüístico inglés-español, a través de dos extensos corpus compilados a tal efecto, uno paralelo (P-GEFEM) y otro comparable (C-GEFEM). Asimismo, apuntan que los resultados de este estudio pueden ser aplicados tanto al desarrollo de una base de datos terminológica para futuras aplicaciones de procesamiento de lenguaje natural como al aula de traducción especializada. De hecho, y con base a los corpus compilados en este estudio, las autoras han desarrollado y patentado recientemente una herramienta de traducción multilingüe semiautomática con el fin de internacionalizar el sector agroalimentario de su región (Castilla y León), demostrando así la transferencia al ámbito empresarial de su investigación basada en corpus sobre el ámbito de los embutidos.

El siguiente capítulo, «Traducción de consentimientos informados y lingüística de corpus: una propuesta metodológica para el aprendizaje de la traducción de textos médico-jurídicos», de M.^a del Mar Sánchez Ramos y Raquel Lázaro Gutiérrez, se centra en el uso de corpus para la traducción de consentimientos informados (inglés-español) en la formación de traductores del ámbito sanitario. En concreto, el uso de WordSmith 4.0 les permite interrogar los corpus compilados y generar un listado de términos en inglés y español de cada uno de los campos semánticos de este particular género textual, que en su caso tiene a la apendicectomía como tema médico. Además, el estudio incluye un completo análisis de los patrones de combinación de los términos extraídos.

El siguiente capítulo, realizado por Lorena Arce Romeral y Míriam Seghiri, y titulado «Diseño de plantillas de redacción y traducción al inglés (variedades británica y estadounidense) de contratos de compraventa de viviendas basadas en corpus», tiene como objetivo la compilación y explotación de un corpus comparable y virtual de compraventa inmobiliaria para la creación de plantillas de redacción y traducción semiautomática de contratos de compraventa de viviendas. De este modo, tras describir pormenorizadamente la legislación aplicable en Reino Unido (haciendo especial hincapié en lo correspondiente a Inglaterra y Gales) y Estados Unidos (en concreto, el estado de Luisiana), presentan los criterios de diseño y compilación del corpus VENTACOR, con los cuales se garantiza su representatividad tanto cualitativa como cuantitativa. Posteriormente, a través del programa AntConc, las autoras identifican y extraen la macro y microestructura de este género textual en las variedades británica y estadounidense para luego realizar un estudio contrastivo de las mismas. Finalmente, y a través de los resultados del análisis micro- y macroestructural de dichos textos, implementan, en formato Excel, plantillas de redacción y traducción de este tipo de documentos en inglés británico (orientado a

Inglaterra y Gales) y en estadounidense (en particular para el estado de Luisiana). Este novedoso recurso resulta de gran interés no solo para la enseñanza de la traducción jurídica, sino para el mercado inmobiliario en su conjunto, pues podrán utilizarlo no solo traductores profesionales sino abogados, gestores inmobiliarios, notarios e, incluso, particulares que deseen prescindir de intermediarios en el proceso de compra y venta de sus viviendas, por citar solo algunos de sus posibles beneficiarios potenciales. En esta línea, cabe destacar que Arce y Seghiri han sido galardonadas con el Premio George Campbell del Aula María Zambrano de Estudios Transatlánticos de la Universidad de Málaga (2019), en reconocimiento a la herramienta de generación de plantillas de redacción y traducción de contratos de compraventa de viviendas creada.

Por su parte, Carmen Bestué y Patricia Rodríguez-Inés, en «Metodología de corpus en la enseñanza de la traducción de cuentas anuales (inglés-español/catalán): un activo intangible», centran su estudio en un encargo propuesto en la asignatura Traducción jurídica y financiera B-A del Grado de Traducción e Interpretación de la Universitat Autònoma de Barcelona. El texto que los alumnos debían trasladar de inglés a español y catalán correspondía a un fragmento de las cuentas anuales de Sainsbury's, una popular cadena británica de supermercados. La primera parte del estudio consiste en traducir mediante traducción automática (en este caso, a través de Google Translate) el fragmento seleccionado y analizar los problemas terminológicos y sintácticos que presenta la traducción generada por este programa. Posteriormente, los discentes traducen con éxito el mismo texto a través de un corpus *ad hoc* de textos originales en español y catalán y se estudia su utilidad para evitar problemas y solventar errores de traducción.

A continuación, Cristina Toledo Báez, en «Sketch Engine en traducción científico-técnica (francés-español): creación y explotación del corpus *ad hoc* comparable GeneCorp», emplea la herramienta de corpus en línea Sketch Engine para explotar un corpus *ad hoc* comparable bilingüe (francés-español) sobre el ámbito de la genética. Este corpus ha sido compilado de forma semiautomática a través de WebBootCat, integrado en Sketch Engine. En concreto, la autora propone explotar este corpus en la asignatura de Traducción científica y técnica de lengua B francés. Así pues, a lo largo del artículo describe de manera clara y concisa diferentes funcionalidades de Sketch Engine, tales como *Wordlists* (para extraer unidades terminológicas monoléxicas y poliléxicas dentro del corpus); *Word Sketch*, para analizar los patrones gramaticales, colocacionales y sintácticos de los lemas; y *Thesaurus*, que permite crear tesauros distribucionales. En definitiva, consigue demostrar cómo la combinación de los corpus *ad hoc* y herramientas como Sketch Engine para explotarlos resulta muy eficaz para

analizar el comportamiento léxico y gramatical de la terminología especializada y obtener a su vez recursos terminológicos, fraseológicos y temáticos fiables y de forma gratuita y rápida, facilitando así la labor traductológica.

En «Creación de un glosario paralelo bilingüe (inglés-español) sobre discapacidad para la interpretación en el contexto legal», Encarnación Postigo Pinazo se centra en mostrar las ventajas que puede aportar la lingüística de corpus, en concreto en el aula de interpretación. En esta ocasión, son los propios estudiantes quienes, a partir de publicaciones científicas y de documentación oficial de la Comisión Europa, compilan un corpus sobre enfermedades raras. Asimismo, mediante el uso de ParaConc, se extrae un completo glosario bilingüe inglés-español que sirve como apoyo para afrontar una interpretación del ámbito legal relacionada a su vez con el campo médico de la discapacidad física e intelectual. Con ello, la autora demuestra que la compilación y la explotación de corpus lingüísticos no solo puede suponer una herramienta de gran valor para la didáctica o el ejercicio profesional de la traducción, sino también de la interpretación.

Tras ello, se llega al segundo bloque del volumen, compuesto por cinco capítulos, los cuales permiten dar cuenta de las diferentes aplicaciones que presentan los corpus para la didáctica de segundas lenguas, haciendo hincapié en la enseñanza de fraseología y terminología. Abre este bloque «Estrategias heurísticas con corpus para la enseñanza de la fraseología orientada a la traducción», capítulo en el que Carlos Manuel Hidalgo Ternero y Gloria Corpas Pastor muestran con éxito las posibilidades que los corpus en red, tanto monolingües como bilingües, presentan a la hora de identificar e interpretar las unidades fraseológicas. En concreto, basan su estudio en dos corpus monolingües en español (CORPES XXI y esEuTenTen), un corpus monolingüe en inglés (enTenTen) y dos paralelos multilingües (Europarl y Linguee, subcorpus inglés-español). En su estudio, los autores describen las diferentes actividades formativas llevadas a cabo en la asignatura de Lengua y Cultura B aplicadas a la Traducción e Interpretación. A través de esta propuesta didáctica, describen de forma pormenorizada al lector a cómo se puede trabajar en el aula las virtudes de este tipo de corpus para identificar e interpretar correctamente las unidades fraseológicas (en especial aquellas que sean susceptibles de presentar múltiples lecturas).

En «Un estudio de caso basado en corpus sobre el uso de las colocaciones verbales en estudiantes de inglés de nivel avanzado», Miriam Buendía Castro compila corpus de aprendices, es decir, formados por los textos producidos por los propios estudiantes (en este caso, se recogen textos argumentativos redactados en lengua inglesa por alumnos de nivel avanzado), para analizar, a través de TermoStat, las colocaciones

(principalmente verbales) usadas por los discentes y así poder poner de manifiesto la variedad de errores cometidos con este tipo de estructuras. A la luz de los resultados del pormenorizado análisis que lleva a cabo la autora, propone orientaciones para mejorar la enseñanza de las colocaciones y con ello la competencia fraseológica de los estudiantes de nivel avanzado de inglés como segunda lengua.

Por su parte, María Rosario Bautista Zambrana, en su capítulo «El uso de corpus y TextStat para la enseñanza-aprendizaje de la negación y de la fraseología en el aula de alemán como lengua extranjera», compila dos corpus monolingües en alemán que posteriormente explotarán mediante TextStat alumnos tanto del Grado en Estudios Ingleses como del Grado en Turismo.. En concreto, hace uso de dos funciones de TextStat: *Word forms* (para extraer las palabras más frecuentes de cada corpus y que sirva al alumnado para familiarizarse con los textos) y *Concordance* (para que los estudiantes busquen concordancias de los distintos términos y estudien su comportamiento morfosintáctico y fraseológico). A través de distintas propuestas pedagógicas, la autora se sirve de los corpus para trabajar en el aula aspectos gramaticales como la negación en alemán así como contenidos léxicos, fraseológicos y gramaticales relacionados con el tema *Wegbeschreibung* (es decir, a partir de textos basados en la descripción de rutas para llegar a diferentes sitios dentro de una ciudad).

Seguidamente, Cristina Castillo Rodríguez y Alexandra Santamaría Urbieta, con su aportación «Extracción de patrones lingüísticos en inglés y en italiano: un caso práctico con corpus etiquetado en el ámbito del turismo de salud y belleza», ponen de manifiesto las ventajas del empleo de corpus etiquetados para la extracción de patrones en la enseñanza de lenguas extranjeras, en particular, del inglés y del italiano. Las autoras describen en su estudio todo el proceso de compilación del corpus, haciendo especial hincapié en el etiquetado de los textos que lo conforman. En concreto, se han servido de tres programas para el etiquetado y la explotación de este corpus sobre turismo de salud y belleza: EncodeAnt, TagAnt y AntConc. Para mostrar el uso real que los alumnos pueden hacer de las herramientas de análisis de corpus, muestran un ejemplo de búsqueda de patrón lingüístico (adjetivo-nombre en el subcorpus en lengua inglesa y nombre+adjetivo en el subcorpus en lengua italiana). Además, proponen otras pautas no solo para seguir explotando terminológicamente este corpus (en concreto muestra cómo buscar secuencias terminológicas más complejas), sino también para detectar errores en el etiquetado de ciertas palabras del corpus.

Cierra el volumen Isabel Durán Muñoz, quien en el capítulo «Evaluación de recursos-e con corpus desde una perspectiva terminológica» analiza con sumo detalle los recursos en línea disponibles que permiten a los

usuarios realizar consultas a corpus textuales. Asimismo, lleva a cabo una evaluación de 14 de estas herramientas mediante una plantilla propia y puntuable. Así pues, los recursos en línea con corpus que incluye en su evaluación son los siguientes: *Bab.la Dictionary*, *British National Corpus (BNC)*, buscadores (por ejemplo, *Google*), *Corpus of Contemporary American English (COCA)*, *Dictionary.com*, *Glosbe*, *Intelitem*, *Linguee*, *MyMemory*, *Reverso Context*, *Termitur*, *Your Dictionary*, *WebCorp* y *Wordincontext*. A través de la lectura de este capítulo, conoceremos hasta qué punto son aptas para cubrir las necesidades de los traductores profesionales herramientas de los siguientes tipos, a saber, corpus en línea de consulta, buscadores de concordancias basados en corpus, memorias de traducción web y combinaciones de corpus y diccionarios electrónicos.

En suma, este volumen, de carácter eminentemente aplicado, plasma a la perfección la utilidad y relevancia que puede tener tanto para discentes y docentes como para profesionales de la traducción y de la interpretación la explotación de corpus especializados. Gracias a ellos, es posible realizar una labor de documentación precisa, rápida y muy efectiva, pues se basa en datos reales y permite observar cómo se comportan los términos en su contexto. Asimismo, con respecto a sus aplicaciones en la enseñanza-aprendizaje de lenguas extranjeras, la lingüística de corpus pone al alcance de discentes y docentes un gran abanico de posibilidades, entre ellas consultar concordancias o colocaciones léxicas, realizar listados para el estudio del léxico en un campo concreto, analizar los errores más frecuentes en corpus de aprendices o estudiar la fraseología. Finalmente, cabe destacar de la obra aquí reseñada la gran diversidad tipológica y temática de aplicaciones que contiene (desde la terminología de los embutidos, hasta temáticas jurídicas o científico-técnicas); los diferentes tipos de corpus con los que los autores del volumen han trabajado (corpus virtual comparable, corpus de aprendices, corpus paralelos, corpus comparable etiquetado, tanto compilados por ellos mismos como disponibles en red); la diversidad de lenguas en las que se han aplicado los estudios (alemán, catalán, español, francés, inglés e italiano); y, por último, los múltiples programas para la compilación, el etiquetado o la explotación de corpus lingüísticos que se han empleado (*AntConc*, *ParaConc*, *SketchEngine*, *WebBootCat*, *TermoStat*, *tlCorpus*, *Wodsmith 4.0*, *TagAnt* o *ReCor*, entre otros).

En definitiva, este volumen muestra con éxito los últimos avances y enfoques en la lingüística de corpus, y sobre todo, ayuda a observar de un modo eminentemente práctico las enormes ventajas y posibilidades que supone trabajar con corpus lingüísticos tanto en el ámbito (profesional y educativo) de la traducción e interpretación como en la enseñanza-aprendizaje de segundas lenguas. Sin duda, estamos ante una obra

exhaustiva, clara, actual y de gran atractivo para discentes, docentes, investigadores y profesionales de la traducción e interpretación o del campo de la enseñanza de lenguas extranjeras.

[MÍRIAM PÉREZ CARRASCO]