EL USO DE LA LINGÜÍSTICA DE CORPUS PARA LA IDENTIFICACIÓN DE CULTUREMAS Y SUS TÉCNICAS TRADUCTORAS EN UN CONJUNTO DE NOVELAS TRADUCIDAS

VIRGINIA MATTIOLI
Universidad Jaime I, Castellón de la Plana

RESUMEN

Con el objetivo de identificar los culturemas en un conjunto de novelas traducidas y analizar sus técnicas traductoras, se ha aplicado la metodología de la lingüística de corpus a la traducción literaria. Así, se ideó una metodología basada en tres momentos distintos: compilación del corpus, identificación de los culturemas y alineado de traducciones y originales para analizar las técnicas traductoras utilizadas en cada caso. En el artículo se explicará cada una de las tres etapas metodológicas, poniendo particular atención en el proceso de compilación ya que, fusionando métodos diferentes –antologización y método estadístico– se ha conseguido compilar un conjunto de novelas trilingüe representativo, equilibrado, variado y auténtico, que por sus características puede ser útil para ulteriores estudios.

Palabras clave: lingüística de corpus, culturemas, traducción literaria

ABSTRACT

With the aim of identifying cultural elements in a set of translated literary texts and analyzing their translation techniques, corpus linguistics methodology has been applied to literary translation studies. Hence, a methodology based on three different moments has been designed: corpus compilation, cultural elements identification and alignment of translated and original texts in order to identify the translation techniques used in each single case. In this paper, each one of those methodological steps will be explained, focusing particularly on the compilation process. In fact, the fusion of different methods –anthologization and statistical analysis– permitted the creation of a representative, balanced, varied and authentic trilingual corpus that could also be useful in further research.

Keywords: corpus linguistics, cultural elements, literary translation

1. INTRODUCCIÓN

En este artículo se resume una parte de una investigación desarrollada en el seno de un proyecto doctoral en el Departamento de Traducción e Comunicación de la Universitat Jaume I en la cual se compararán los resultado obtenidos de este análisis con los obtenidos del mismo estudio aplicado a diferentes géneros literarios. Específicamente se utilizará la metodología de la lingüística de corpus con el fin de identificar los elementos culturales y sus técnicas traductoras en un conjunto de novelas traducidas del español y del inglés al italiano.

Con este fin, en las páginas a continuación en primer lugar se explicarán los objetivos del estudio y se presentarán las diferentes fases metodológicas utilizadas para alcanzarlos —compilación, análisis y comparación— deteniéndose de manera particular en la primera de estas. Finalmente, se presentarán las principales características del corpus obtenido y se concluirá ofreciendo ulteriores posibilidades de investigación a partir de los resultados alcanzados.

2. OBJETIVOS Y METODOLOGÍA

El objetivo general de este estudio es la determinación de las técnicas de traducción de los elementos culturales identificados en un conjunto de traducciones al italiano de novelas inglesas y españolas publicadas entre 2000 y 2014. Este objetivo general además se puede desglosar en tres objetivos específicos:

- compilación de un corpus representativo y equilibrado que mediante el análisis de algunas novelas traducidas permita generalizar los resultados al entero género literario para el periodo y los idiomas considerados;
- identificación de los elementos culturales presentes en las novelas en análisis mediante la metodología de la lingüística de corpus;
- determinación de las técnicas de traducción utilizadas en cada caso mediante la comparación de los textos traducidos con sus respectivos originales.

Una vez establecidos los objetivos, se escogió el proceso metodológico que está formado por tres etapas, cada una de las cuales se corresponde con uno de los objetivos propuestos anteriormente:

- metodología de compilación;
- metodología de análisis;
- metodología de comparación.

Cada etapa metodológica se divide a su vez en diferentes fases. En los apartados a continuación se presentará cada uno de los tres momentos en detalle, si bien focalizándose mayormente en la sección de compilación, ya que de momento el estudio —todavía en progreso— se ha desarrollado solo hasta ese punto.

2.1 Compilación

La primera etapa metodológica se corresponde con la compilación del corpus que se analizará sucesivamente. Para copilar el conjunto de obras necesario se actuó siguiendo cinco fases:

- identificación del corpus necesario para alcanzar los objetivos propuestos;
- elección de los criterios de compilación;
- elección del método de compilación;
- selección de los textos;
- búsqueda y codificación de los textos.

Considerado el objetivo del estudio, para alcanzar los resultados se optó por analizar un corpus específico (por estar formado por textos de un género determinado), trilingüe (por estar compuesto por obras originales en inglés y español y traducciones en italiano) y compilado *ad hoc* (por no haber encontrado ningún conjunto de textos con estas características después de una minuciosa pesquisa entre la literatura y los estudios precedentes sobre el tema).

Tras establecer los rasgos fundamentales del conjunto de textos que se analizará, en la segunda fase se escogieron los criterios necesarios para su compilación. Después del examen de diversas teorías previas sobre la compilación de corpus, en particular con

referencia a aquellos específicos y *ad hoc*, se consideró seleccionar aquellos cinco parámetros considerados ineludibles por el mayor número de autores estudiados: representatividad tanto cualitativa como cuantitativa, elección de textos entero frente a fragmentos, equilibrio —aunque, habiendo optado por la inclusión de textos enteros, se pudo alcanzar exclusivamente un equilibrio interior o sea, el que consiste en la misma distribución y proporción de los textos entre las diferentes partes y subpartes del corpus (*The University of Manchester* 2011:14)— variedad y autenticidad.

En la fase siguiente, dedicada a la selección del método de compilación, se respetó cada uno de los criterios escogidos con anterioridad. En este caso, el criterio más difícil con el que cumplir fue el de representatividad, vista la falta de claridad sobre los parámetros necesarios para compilar un corpus representativo (Seghiri, 2006:257). Esta dificultad se superó juntando dos métodos diferentes y dividiendo el proceso de compilación en dos momentos distintos: primero, mediante el método de la antologización se escogieron las obras respetando la representatividad cualitativa y sucesivamente se comprobó la representatividad cuantitativa del corpus mediante el método estadístico.

El método de la antologización prevé la selección de las obras entre las incluidas en antologías. Este procedimiento se basa en la idea según la cual la antología es una forma de canon y en consecuencia todas las obras que esta recoge se pueden considerar representativas y en el centro del polisistema literario (Sela-Sheffey 2002 cit. Sánchez Galvis 2012:129). Así que, según la propuesta de Sánchez Galvis (2012), al adoptar este método se remite la elección de las obras a un experto en el tema. Para mantener la posibilidad de atenerse al juicio de críticos competentes a pesar de la escasez de antologías que reúnen textos del género analizado, en este estudio se consideró seleccionar las novelas entre aquellas galardonadas con premios literarios de fama internacional.

Por otro lado, para cumplir con la representatividad cuantitativa se utilizó el programa *ReCor* (Corpas Pastor, Seghiri Domínguez, Maggi 2006) patentado por la Universidad de Málaga que, mediante un cálculo estadístico permite comprobar a posteriori la representatividad cuantitativa de un conjunto de textos (Corpas Pastor,

Seghiri Domínguez, 2007:166). Evidentemente, por su naturaleza a posteriori, este análisis se aplicó al corpus como última fase de la etapa de compilación, después de la elección y de la recopilación de las obras.

Ciñéndose a la metodología escogida, en la cuarta fase se adoptó el método de la antologización para seleccionar los textos. En primer lugar se eligieron cuatro galardones literarios: el premio Nobel[1], de prestigio indiscutible, y tres premios de fama internacional, cada uno otorgado por un país en el cual la lengua oficial es una de las tres analizadas en este estudio[2]. Además, a pesar de la selección equilibrada de los premios entre los idiomas estudiados, en el caso del español tuvo que incorporarse otro galardón con el fin de eliminar las disparidades causadas por la influencia de la diferencia de prestigio entre las dos lenguas de partida en el panorama literario internacional[3].

A continuación, de todas las novelas premiadas en los galardones considerados se seleccionaron exclusivamente aquellas escritas en lengua inglesa o española. De estas se tomaron en consideración solo aquellas publicadas entre 2000 y 2014 —periodo de interés para este análisis— y entre las obras resultantes se escogieron todas aquellas traducidas al italiano.

Finalmente, se procedió a la última fase en la cual se buscaron y se codificaron las obras seleccionadas, mediante los procesos de escaneo —para aquellas obras halladas exclusivamente en papel—, limpieza, conversión en formato *.txt* y alineación de originales y traducciones.

Una vez obtenido el corpus, se comprobó estadísticamente su representatividad cuantitativa a posteriori, a través del programa *ReCor* (Corpas Pastor, Seghiri Domínguez, Maggi, 2006). Como el corpus contiene novelas en diferentes idiomas originales, se analizó cada conjunto lingüístico por separado —antes aquel formado por textos en inglés y luego en español—.

De esta forma, LIT_TRAD, el corpus resultante de este proceso de compilación, está formado por 50 novelas de tema general organizadas en 25 parejas (original y traducción) divididas en dos subcorpus: LIT_TRAD_EN, formado por 13 parejas de novelas inglesas traducidas al italiano, y LIT_TRAD_ES, formado por 12

parejas de novelas españolas traducidas al italiano. En la figura 1 se muestra la distribución de las obras en LIT_TRAD:

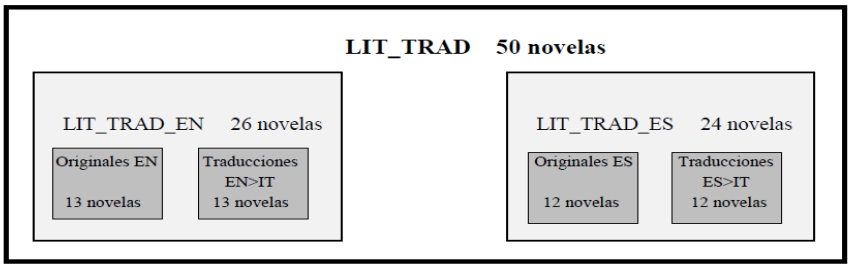

Figura 1. Distribución de las obras en LIT_TRAD

Como se desprende de la Figura 1, los dos subcorpus no son perfectamente equilibrados (LIT_TRAD_EN presenta una pareja de novelas más que LIT_TRAD_ES): esta sutil diferencia —de influencia mínima en el conjunto de obras total— se debe a la decisión de evitar eliminar aleatoriamente una pareja de obras para no afectar la objetividad del proceso de selección.

En definitiva, LIT_TRAD resulta ser un corpus cualitativamente y cuantitativamente representativo —gracias a su compilación por medio de la fusión del método de la antologización con el estadístico—, formado por textos enteros, equilibrado internamente, variado lingüística y cronológicamente —ya que reúne textos de diferentes variedades de inglés y español y publicados en un periodo de 15 años— y auténtico —por estar formado por obras escritas para contextos reales y por escritores nativos—.

2.2 Análisis

Tras haber compilado el corpus LIT_TRAD, se sugirió una metodología de análisis que permite identificar los elementos culturales en el conjunto de obras.

Como la anterior, también esta etapa metodológica se puede desglosar en diferentes fases:

- determinación de la definición y de las clasificaciones de culturemas y de técnicas traductoras adoptadas en este estudio;

- identificación de los culturemas en el corpus LIT_TRAD, mediante las herramientas de listados de palabras, palabras clave, *clusters* y concordancias.

De acuerdo con la primera fase, se adoptaron la denominación y la definición de Vermeer (1983), aceptadas posteriormente por Christiane Nord (1997), según la cual los elementos culturales, denominados culturemas, son

> fenómenos sociales de una cultura X, relevantes para los miembros de esa cultura, que cuando vienen comparados con un fenómeno correspondiente de una cultura Y resultan específicos de la cultura X. (Vermeer, 1983 en Nord, 1997:34)

Entre las diferentes propuestas, se escogió esta definición por la centralidad del factor cultural en la traducción, el énfasis en la relación entre las dos culturas y la importancia del contexto de la traducción.

Una vez establecido qué se entiende por culturemas, por un lado se revisaron las propuestas taxonómicas de los autores previos y se clasificaron los culturemas en 6 categorías culturales (cultura geográfica y natural, cultura material, cultura social, histórica y de folclore, cultura religiosa, cultura conceptual y cultura lingüística). Por otro lado, después de haber considerado las propuestas clasificatorias más aceptadas entre las presentadas hasta hoy por diferentes investigadores, se decidió crear una taxonomía basada en categorías discretas — siguiendo el ejemplo de autores como Nida (1964:226-238), Newmark (1988:103-104), Vinay y Darbelnet (1995:31-40), Santamaría (2001:288) y Molina Martínez (2006:101-104)— organizadas desde la que más crea un efecto de exotismo a la que más propende a la adaptación cultural — enlazándose a la idea de un continuo que representa el grado de fidelidad de la traducción con respecto al original, propuesta entre otros por Mangiron i Hevia (2006:110)—. Dicha taxonomía se compone de 14 clases de técnicas traductoras: transposición, préstamo, naturalización, traducción literal, neutralización, hiperónimo, doblete, traducción estándar aceptada, paráfrasis explicativa, notas a pie de página, supresión de informaciones, total omisión del culturema, equivalente cultural o funcional y creación autónoma.

Como ya se ha introducido al principio de este artículo, el presente trabajo se halla todavía en progreso y de momento se llegó solo hasta este punto. Para seguir con la segunda fase del análisis e identificar los culturemas, la propuesta metodológica —susceptible de modificaciones y mejorías durante su aplicación— prevé cumplir las siguientes tareas:

- creación de un listado de palabras clave dentro del cual identificar los culturemas, o sea aquellos términos que respetan la definición aceptada en este estudio;
- comprobación de la naturaleza cultural de cada culturema por medio de listados de concordancias y *clusters*, que permiten analizar cada elemento en su contexto;
- asignación de una categoría y de una subcategoría cultural entre las establecidas en la taxonomía propuesta a cada uno de los culturemas identificados.

2.3 Comparación

Una vez identificados y clasificados los culturemas del corpus, estos se compararán con los respectivos elementos originales mediante los alineados con el fin de identificar la estrategia de traducción adoptada en cada caso y clasificarla según la propuesta taxonómica de técnicas de traducción presentada anteriormente.

3. CONCLUSIONES

Siguiendo el método presentado en este artículo, se ha conseguido compilar un corpus representativo, formado por textos enteros, equilibrado internamente, variado y auténtico que, además de analizarse en la presente investigación, se podrá utilizar para ulteriores trabajos. Asimismo, se propuso una metodología de análisis que permite estudiar un conjunto de obras literarias mediante la metodología de corpus.

Entre las posibles investigaciones futuras sobre el tema, se pueden sugerir análisis finalizados a ampliar los resultados obtenidos de este estudio —tanto desde un punto de vista comparativo, como

descriptivo—. En el mismo sentido, se puede observar el comportamiento de algunos rasgos o elementos específicos que destaquen de las conclusiones —como profundizar el estudio de la clase de culturema que resulte más productiva o emprender un análisis diacrónico del utilizo de los culturemas en los textos literarios—. De otra forma, la metodología propuesta en esta ocasión, una vez desarrollada y mejorada, se puede volver a aplicar a estudios futuros, bien continuando a investigar sobre las temáticas introducidas en este artículo, o bien persiguiendo objetivos diferentes.

NOTAS

[1] Disponible en http://www.nobelprize.org/nobel_prizes/literature/laureates/ (última visita el día 01/06/2016).

[2] Para la selección de los textos, se consideraron el premio italiano *Grinzane Cavour*, el premio *Princesa de Asturias de las letras* expedido en España y el *America Award in Literature* otorgado por una importante institución estadounidense, disponibles en http://www.epdlp.com/premios.php?premio=Grinzane%20Cavour, http://www.fpa.es/es/premios-princesa-de-asturias/ y http://www.greeninteger.com/america.cfm (última visita el día 01/06/2016).

[3] Para resolver el problema de la disparidad entre el número de novelas galardonadas y traducidas del inglés y del español (19 contra 10), se añadieron las novelas ganadoras del *Premio Rómulo Gallego de Novela*, disponible en http://www.celarg.org.ve/Espanol/Premio%20Romulo%20Gallegos.htm (última visita el día 01/06/2016).

REFERENCIAS BIBLIOGRÁFICAS

Corpas Pastor G., Seghiri Domínguez M. 2007. "Determinación del umbral de representatividad de un corpus mediante el algoritmo N-Cor", *Procesamiento del Lenguaje Natural*, 39 (2007):165-172.

Corpas Pastor G., Seghiri Domínguez M., Maggi R. 2006. ReCor: método para la determinación de la representatividad de un corpus, patente n. ES2320511 de la Universidad de Málaga, http://umapatent.uma.es/es/patent/metodo-para-la-determinacion-de-la-representa4b0/ (última visita el día 01/06/2016)

Mangiron I Hevia C. 2006. *El tractament dels referents culturals a les traduccions de la novel·la Botxan: la interacció entre els elements textuals i extratextuals*, tesis doctoral, Universitat Autònoma de Barcelona, Departamento de Traducción e Interpretación, Barcelona. En http://hdl.handle.net/10803/5270 (última visita el día 01/06/2016).

Molina Martínez L. 2006. *El otoño del pingüino*. Castellón de la Plana: Publicacions de la Universitat Jaume I.

Newmark P.1988. *A textbook of translation*, New York: Prentice Hall.

Nida A. 1964. *Toward a Science of Translating: With Special Reference to Principles and Procedures Involved in Bible Translating.* Leiden: Brill Archive.

Nord C. 1997. *Translating as a purposeful activity.* Manchester: St Jerome.

Sánchez Galvis J. 2012. "Traducción y variedad lingüística: hacia un modelo de reconstrucción dialectal" *Revista electrónica de lingüística aplicada*, 11:125-136.

Santamaría Guinot L. 2001. *Subtitulació i referents culturals. La traducció com a mitjà d'adquisició de representacions mentals*, tesis doctoral, Universitat Autònoma de Barcelona, Departamento de Traducción e Interpretación, Barcelona. En http://hdl.handle.net/10803/5249 (última visita el día 01/06/2016).

Seghiri Domínguez 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, Facultad de Filosofía y Letras. En *atarazanas.sci.uma.es/docs/tesisuma/16754888.pdf* (última visita el día 01/06/2016)

The University of Manchester 2011. "The Translational English Corpus: A practical approach to corpus building", https://artisinitiative.files.wordpress.com/2014/05/the-translational-english-corpus1.pptx (última visita el día 01/06/2016)

Vinay J. P., Darbelnet J. 1995. *Comparative stylistics of French and English: a methodology for translation.* Amsterdam: John Benjamins.