

Trabajo de Final de Máster

Máster en investigación en traducción e interpretación

**IDENTIFICACIÓN Y CLASIFICACIÓN DE CULTUREMAS Y
PROCEDIMIENTOS TRADUCTORES EN EL ARCHIVO DE TEXTOS
LITERARIOS LIT_ENIT_ES: UN ESTUDIO DE CORPUS**

Autor: Virginia Mattioli

Directora: María Calzada Pérez

Fecha de lectura: 17 de noviembre de 2014



Resumen: Este estudio se centra en los elementos culturales, que, de acuerdo con la definición de Christiane Nord (1997:34) y otros autores, se denominan culturemas y se caracterizan por ser elementos abstractos propios de una cultura y útiles por su comparación entre esta y otras culturas diferentes. Debido a su naturaleza cultural, a menudo se encuentran problemas en la traducción de dichos elementos. Estas dificultades han llevado a diversos autores a proponer unas técnicas de trasvase (el préstamo, el calco, la traducción literal, la notas a pie de página, la utilización de equivalentes culturales, las reducciones o ampliaciones de significado, etc.) y a enumerar los factores que se han de considerar en su elección. En este estudio se analizan los culturemas y las técnicas utilizadas para su transposición de una cultura a otra. La metodología elegida para desarrollar este análisis es la de los corpus electrónicos. El objetivo principal del estudio es demostrar que esta metodología facilita la identificación y la clasificación de los culturemas en el corpus de obras traducidas utilizadas en este trabajo. Los objetivos específicos se pueden sintetizar en: obtener un conocimiento más profundo del corpus utilizado que pueda ser útil como base para ulteriores investigaciones, demostrar la operatividad de las clasificaciones propuestas para los culturemas y para las técnicas de traducción utilizadas, y demostrar la efectividad del método utilizado para que se pueda emplear en trabajos futuros, específicamente en una tesis doctoral. Persiguiendo estos fines, el análisis se desarrolla en dos etapas, una descriptiva y otra comparativa, ambas divididas a su vez en dos fases. Con el objetivo de analizar la naturaleza de los culturemas y mejorar el conocimiento del corpus de estudio utilizado, el análisis descriptivo se divide en una primera fase de identificación y clasificación de los culturemas en categorías culturales y una segunda fase de identificación y clasificación de los culturemas más frecuentes en el corpus (con

una frecuencia mayor o igual a 50 ocurrencias). El análisis comparativo permite, en la primera fase, identificar y clasificar las técnicas de traducción, mediante la utilización de un alineado entre obras traducidas y sus respectivos originales, y en la segunda fase analiza y calibra el peso idiosincrásico de los culturemas del corpus de literatura traducida y lo coteja con el de los culturemas localizados en las obras en lengua original.

Palabras clave: culturemas, técnicas de traducción, lingüística de corpus.

ÍNDICE

| | |
|--|----|
| Introducción | 9 |
| I Motivación | |
| personal..... | 9 |
| II Conceptos metodológicos | |
| básicos..... | 10 |
| III Hipótesis y objetivos | |
| | 12 |
| IV Estructura del presente estudio..... | 13 |
| 1. Cultura y culturemas: definición, clasificaciones y propuesta teórica..... | 16 |
| 1.1. El concepto de cultura..... | 16 |
| 1.2. El concepto y las clasificaciones de culturema..... | 20 |
| 1.3 Propuesta de clasificación de culturema..... | 29 |
| 1.4 Marcas culturales y pseudo-marcas culturales..... | 35 |
| 2. Tratamiento de cultura y culturemas..... | 37 |
| 2.1 Diferencia entre técnica, estrategia y método de traducción..... | 37 |
| 2.2 Problemas derivados de la traducción de los culturemas..... | 39 |
| 2.3 Factores que determinan la elección de las soluciones de traducción de culturemas..... | 40 |
| 2.4 Clasificaciones de los tratamientos para traducir los culturemas..... | 42 |
| 2.5 Propuesta de clasificación de tratamiento de los culturemas..... | 51 |
| 3. Metodología: la lingüística de corpus y los métodos del presente trabajo | 55 |
| 3.1 Estudios traductológicos de corpus | 55 |
| 3.1.1 Lingüística de corpus: definición y características..... | 55 |

| | |
|--|----|
| 3.1.1.1 Definición de lingüística de corpus y su concreción en la traductología..... | 55 |
| 3.1.1.2 Objetivos y aplicaciones de investigación basada en corpus..... | 60 |
| 3.1.1.3 Ventajas y límites de los métodos basados en corpus..... | 62 |
| 3.1.2 Breve recorrido histórico de los estudios traductológicos basados en corpus..... | 64 |
| 3.1.3 Clasificación de corpus lingüísticos..... | 66 |
| 3.1.4 Criterios de compilación de corpus lingüísticos..... | 75 |
| 3.1.5 Herramientas de procesamiento y análisis..... | 81 |
| 3.2 Metodología del presente estudio..... | 84 |
| 3.2.1 Hipótesis y objetivos del estudio..... | 85 |
| 3.2.2 Presentación de LIT_ENIT_ES: el archivo de los corpus utilizados para el estudio..... | 87 |
| 3.2.3 Criterios de compilación..... | 94 |
| 3.2.4 Proceso de preparación del corpus..... | 96 |
| 3.2.5 Metodología de análisis..... | 96 |
| 3.2.5.1 Análisis descriptivos: identificación y clasificación de culturemas en el corpus de literatura traducida LIT_ES y en los subcorpus de obras traducidas del inglés LIT_EN_ES y del italiano LIT_IT_ES | 96 |
| 3.2.5.1.1 1ª fase: identificación y clasificación de los culturemas en categorías culturales | 97 |
| 3.2.5.1.2 2ª fase: identificación de culturemas más frecuentes de la literatura traducida | 98 |

| | |
|--|-----|
| 3.2.5.2. Análisis comparativos | 100 |
| 3.2.5.2.1 Análisis comparativo interno | 101 |
| 3.2.5.2.2 Análisis comparativo externo..... | 102 |
| 4 Análisis descriptivos y comparativos del corpus LIT_ENIT_ES..... | 103 |
| 4.1. Análisis descriptivos..... | 103 |
| 4.1.1 1ª fase: identificación y clasificación de culturemas en categorías culturales..... | 105 |
| 4.1.2 2ª fase: identificación de culturemas más frecuentes de la literatura traducida..... | 113 |
| 4.2 Análisis comparativos..... | 158 |
| 4.2.1. Análisis comparativo interno..... | 159 |
| 4.2.2 Análisis comparativo externo | 166 |
| 5 Conclusiones..... | 177 |
| 5.1 Recapitulación del trabajo, de la hipótesis y de los objetivos..... | 177 |
| 5.2 Interpretación de los resultados..... | 179 |
| 5.3 Propuestas para investigaciones futuras..... | 183 |
| Anexos..... | 186 |
| Anexo 1..... | 186 |
| Anexo 2..... | 187 |
| Anexo 3..... | 189 |
| Anexo 4..... | 202 |
| Anexo 5..... | 209 |
| Anexo 6..... | 215 |
| Anexo 7..... | 217 |

| | |
|---------------------------|-----|
| Anexo 8..... | 219 |
| Anexo 9 | 221 |
| Anexo 10..... | 224 |
| Índice de las tablas..... | 226 |
| Índice de los anexos..... | 228 |
| Bibliografía..... | 229 |

AGRADECIMIENTOS

Quisiera expresar mi sincero agradecimiento a la Dra. María Calzada Pérez, directora de esta tesis, por su ayuda, su inmensa disponibilidad y la infinita paciencia.

A Alberto, mi amigo y amante, por haber estado a mi lado durante los largos días de estudio.

A mis padres, por su apoyo y confianza.

INTRODUCCIÓN

I Motivación personal

El presente trabajo fin de máster (TFM) nace de una predilección por la literatura, debida en parte a mis estudios previos en el campo de la filología y en parte a una inclinación personal, unida a un interés por la noción de cultura (y, en consecuencia, de culturema) y por el concepto de traducción, entendido como intercambio cultural y no solo como transposición lingüística. A estas razones se une la fascinación por la tradumática (o tecnología aplicada a la traducción), que ha influido en la metodología elegida: la lingüística de corpus. Por todo ello, este TFM consiste en un estudio basado en corpus acerca de la identificación y clasificación de elementos culturales y de su tratamiento en la literatura traducida. Además, como se explicará en el cuerpo del estudio, este trabajo es el principio de un proyecto de investigación más amplio, que incluye una tesis doctoral que se desarrollará a partir de los resultados obtenidos aquí.

Ya antes de comenzar se consideraba que las conclusiones que podrían derivarse de un estudio de estas características serían interesantes y útiles para la traductología desde diversos puntos de vista. Desde la perspectiva científica, cabía pensar que los resultados abrirían múltiples vías de investigación futura. Así, podrían, por ejemplo, desarrollarse estudios comparativos entre un corpus de literatura traducida, como el utilizado en este estudio (reforzado), con otros corpus (comparables) que reunieran textos de épocas, autores, temas o géneros diferentes para identificar semejanzas y divergencias. También podrían utilizarse los resultados obtenidos del presente estudio como punto de partida para describir aspectos relacionados con la teoría de la traducción, como los universales lingüísticos. Asimismo, podría profundizarse en el conocimiento de un aspecto

específico de los resultados que se obtuvieran en el presente trabajo: una categoría de culturemas determinada, el comportamiento de los elementos culturales en las obras de un autor específico, la influencia del par de lenguas en la cantidad, las categorías o las técnicas traductoras de los culturemas, etc. Por lo demás, desde el punto de vista didáctico, los resultados podrían revelarse útiles, por ejemplo, para la enseñanza de la traducción de los culturemas y la aplicación de las diferentes técnicas según contextos diferentes. En definitiva, el trabajo que nos disponíamos a abordar se iniciaba, desde el principio, con unas perspectivas futuras muy amplias y enriquecedoras.

II Conceptos metodológicos básicos

Analizando más detalladamente el tema presentado en el apartado anterior, el foco principal del análisis de este trabajo se centra en los elementos culturales, a los cuales, de acuerdo con autores como Nord (1997:34), Nadal (2009:97) y Molina Martínez (2001:91), se denomina culturemas y que, con arreglo a la definición adoptada por Christiane Nord (1997:34), son elementos abstractos propios de una cultura y útiles por su comparación entre esta y otras diferentes.

Dados este concepto y esta definición de culturema, durante la actividad de trasvase surgen problemas por la transposición de los textos a entornos culturales distintos. Estas dificultades traductoras han llevado a diversos estudiosos, entre los cuales destacan Nida (1964:226-238), Newmark (1988:103), y Vinay y Darbelnet (1995:31-40), a proponer unas técnicas (como el préstamo, el calco, la traducción literal, la notas a pie de página, la utilización de equivalentes culturales, las reducciones o ampliaciones de significado, etc.) y a enumerar los factores decisivos para su elección. Todos estos

conceptos se estudiarán en el presente TFM en el marco de obras de literatura traducidas en español a partir de originales en inglés y en italiano .

La metodología elegida para desarrollar este análisis es la de los corpus lingüísticos, con ventajas evidentes en cuanto al tamaño del objeto de estudio, su naturaleza electrónica y la precisión ofrecida por las técnicas informáticas, con vistas a alcanzar unos objetivos entre los cuales se mencionan el análisis de los textos, los elementos gramaticales, sintácticos y léxicos, las colocaciones y las combinaciones sintácticas y semánticas (Gandin 2009:141 y 147). La metodología de corpus nació a finales de la década de los 50 del siglo pasado en Estados Unidos y experimentó un desarrollo fulgurante durante los años 70 y 80 del mismo siglo, período al cual se remontan sus primeras aplicaciones en los estudios de traducción. En la década siguiente aparecieron los megacorpus (Sardinha Berber 2000:330-331), fundamentales para el auge exponencial de esta metodología, que ha seguido expandiéndose por todo el mundo, en particular por Europa, hasta llegar en la actualidad a su aplicación tanto en ámbito investigador como en el empresarial.

Para apuntalar esta metodología, muchos autores, como Berber Sardinha (2000:340-341), Bowker (2002), Laviosa (2002:34-38) y Gandin (2009:134-135), han avanzado propuestas sobre la organización y las denominaciones de los corpus, añadiendo numerosos criterios de clasificación que superan la disposición más superficial de corpus comparables (mono/bi/multilingües) y corpus paralelos. Cabe, por último, recordar que los corpus electrónicos se analizan por medio de programas informáticos que disponen de herramientas de procesamiento (lematización, anotación, marcado y alineación) y de herramientas de análisis (listado de palabras, de palabras clave, de frecuencia, de concordancia, etc.).

III Hipótesis y objetivos

La hipótesis principal de la cual se parte en el presente estudio es la siguiente:

una metodología basada en corpus facilita el proceso de identificación de áreas proclives al uso de culturemas y propicia el análisis cuantitativo y cualitativo del tratamiento de los culturemas identificados.

De esta hipótesis se establece el objetivo principal del trabajo: identificar, clasificar y comparar culturemas de literatura traducida frente a culturemas de literatura original con el fin de identificar, clasificar y calibrar tratamientos traductores.

Este objetivo principal puede desglosarse en seis objetivos específicos, que se corresponden con las diferentes fases del análisis de este TFM:

- diseñar un método operativo para identificar los culturemas del corpus analizado en el presente estudio con vistas a posibles investigaciones futuras;
- clasificar los culturemas en categorías para comprobar así la operatividad de la clasificación propuesta en la parte teórica del trabajo;
- comparar los culturemas de la literatura traducida con los de las respectivas obras originales;
- diseñar un método para identificar las técnicas de traducción utilizadas en el corpus analizado que sea operativo y replicable en investigaciones futuras;
- clasificar las técnicas traductoras en categorías y comprobar así la operatividad de la clasificación propuesta en la parte teórica del trabajo; y finalmente

- obtener un conocimiento más detallado del corpus analizado.

Conviene aquí subrayar que, en el presente estudio, no se persiguen los objetivos de exhaustividad ni de significación estadística. En primer lugar, este TFM no se postula como examen irrefutable de la literatura traducida, sino más bien como comprobación de la operatividad de un método de trabajo que se empleará en investigaciones futuras. En segundo lugar, no existen aseveraciones indiscutibles de significación estadística por dos razones. De una parte, porque, en esta fase de la investigación, dicho nivel de precisión se considera inalcanzable, tanto por la identificación de un amplio número de elementos culturales, como por la técnica de resolución de los cálculos, realizados de forma manual, lo cual genera la posibilidad de errores. De otro lado, se han de reconocer ciertas dosis de subjetividad investigadora (limitadas al máximo mediante los postulados teóricos), que influyen el proceso de identificación y de clasificación de los culturemas y de las técnicas de traducción.

IV Estructura del presente estudio

Para corroborar la hipótesis y los objetivos presentados en el apartado precedente, este trabajo se ha estructurado en 5 capítulos, de los cuales los dos primeros se centran en la revisión de postulados teóricos, el tercero pone el foco de atención sobre la metodología del estudio, el cuarto, incluye los análisis propuestos y el quinto se propone a modo de conclusión.

Entrando un poco más en el detalle, en el primer capítulo se realiza una presentación de cultura, tal y como se ha considerado dicho concepto desde el pasado reciente hasta el momento actual. El capítulo sigue con la descripción de planteamientos teóricos acerca de la naturaleza, la definición y las clasificaciones de los culturemas, sobre los que se

basa la clasificación elegida para el análisis. Por último, se concluye con la presentación de la distinción entre marcas culturales y pseudo-marcas culturales.

En el segundo capítulo se aborda el problema que plantea la traducción de los culturemas y, acto seguido, se presentan las técnicas identificadas por los traductólogos para resolver dicha problemática. Se explican, asimismo, los criterios de elección de las técnicas traductoras y las diferentes clasificaciones sugeridas por los investigadores. El capítulo se cierra con una propuesta de categorías de técnicas de traducción de culturemas, realizada a partir los planteamientos teóricos examinados.

El tercer capítulo se divide en dos partes. La primera se ocupará de presentar la lingüística de corpus como metodología de investigación, exponiéndose sus características, objetivos, aplicaciones, límites y ventajas. Se realiza un breve recorrido histórico desde el nacimiento de esta metodología hasta llegar a la época actual y se revisan propuestas de clasificación de los corpus lingüísticos, criterios para su compilación y finalmente herramientas utilizadas para su procesamiento y análisis. En la segunda parte de este capítulo, se explica la metodología del presente estudio. Aquí se plantea la hipótesis de partida y los objetivos perseguidos, se describe el archivo de corpus LIT_ENIT_ES utilizado y se explican los criterios seguidos para su compilación. A continuación se incluye un apartado sobre el proceso de preparación del corpus y se concluye con la explicación de las etapas y las fases del análisis.

El cuarto capítulo, dedicado al análisis, se divide en dos etapas, una descriptiva y otra comparativa, cada una compuesta, a su vez, por dos fases. El análisis descriptivo comprende una primera fase de identificación y clasificación de los culturemas en las categorías culturales precedentemente establecidas y una segunda fase de identificación

de los culturemas más frecuentes del corpus analizado (con una aparición mayor o igual a 50 ocurrencias) y de clasificación según las mismas categorías culturales utilizadas en la fase anterior. Como se ha mencionado, también el análisis comparativo está formado por dos fases. En la primera, se desarrolla un análisis interno para comparar las obras traducidas con sus respectivos originales, por medio de un proceso de alineación, con el que se detecta la técnica utilizada para traducir cada culturema antes identificado. En la segunda fase se concretará un análisis externo, en el cual se confrontarán los corpus analizados con corpus de referencia, para identificar y calibrar el peso idiosincrásico de los culturemas identificados en la literatura traducida y de los localizados en las respectivas obras originales.

Finalmente, en el quinto capítulo, se presentan las conclusiones. Después de haber recapitulado los componentes principales del estudio, se interpretan, a modo de conjeturas que pueden servir como inspiración para trabajos posteriores, los resultados obtenidos en el capítulo de análisis y se cotejan dichos resultados con las hipótesis y los objetivos del estudio para demostrar o desmentir su corroboración y cumplimiento. Finalmente se señalan posibles vías de investigación a partir de los resultados obtenidos.

1. CULTURA Y CULTUREMAS: DEFINICIÓN, CLASIFICACIONES Y PROPUESTA TEÓRICA

En este capítulo se abordarán las nociones de cultura y de culturema fundamentales, para el desarrollo de este estudio. Después de explicar teorías y las definiciones de cultura, se analiza la relación de este concepto y con el de trasvase, donde el traductor muy a menudo desempeña el papel de mediador. Acto seguido, se presentan los términos que han utilizado los estudiosos para referirse a los culturemas, con las respectivas definiciones y las categorías que se utilizan para clasificarlos. A la luz de este examen teórico se propone una clasificación que servirá de base para el presente trabajo. Finalmente, se subraya la diferencia entre marca cultural y pseudo-marca cultural, explicando cómo esta diferencia afecta la traducción de los culturemas y por consecuencia los resultados del presente análisis.

1.1 El concepto de cultura

Cultura es un término al que se han acercado múltiples estudiosos con afán de desentrañarlo. Desde territorios afines a la traductología, ya en 1871, Tylor en su obra *Cultura Primitiva*, de 1871, afirma:

La Cultura o Civilización, tomada en su amplio sentido etnográfico, es ese complejo conjunto que incluye el conocimiento, las creencias, el arte, la moral, las leyes, las costumbres y cualesquiera otras aptitudes y hábitos adquiridos por el hombre como miembro de la sociedad.

(Tylor 1977 [1871]:19)

Después de Tylor, y a lo largo del siglo XX, muchos autores han otorgado diferentes definiciones al concepto de cultura. A finales de los años 80, Newmark (1988:94) la considera un estilo de vida y la vincula a las manifestaciones peculiares de una

comunidad que utiliza lenguaje y expresiones propias. En la misma década y vista la imposibilidad de definir unívocamente el concepto de cultura, Robinson (1988:8-12) considera oportuno examinar ese concepto desde diferentes prismas, por lo que define la cultura desde el punto de vista behaviorista (en función de las acciones de sus miembros), funcionalista, (a partir de las causas de los comportamientos de los individuos), cognitivo (como un modelo de realidad formado por un grupo determinado) y simbólico (como un concepto en evolución constante que depende del cambio social). A finales de los 90, Samovar (2014[1997]:10-12) habla de la cultura como de un depósito de conocimiento, experiencias, ideas, valores, aptitudes, significados, jerarquías sociales, religión, objetos y nociones de espacio, tiempo y universo, adquirido por un grupo de gente a lo largo de las generaciones, añadiendo al concepto de cultura (hasta ahora ligado a ideas, creencias y conocimiento) elementos concretos como las organizaciones sociales y los objetos. Katan (1999: 26) considera la cultura un sistema de valores, creencias, estrategias, etc., en el cual cada aspecto está relacionado con los demás en el seno de un conjunto que representa un modelo del mundo compartido. Por su parte, al comienzo del siglo XXI, Mangiron (2006:51) acepta esta visión de la cultura como sistema pero añade el concepto de regulación establecido por una comunidad y el de aceptación o rechazo de esas normas por parte de los individuos, definiendo la cultura como un

sistema de creences, valors, estratègies i entorns cognitius d'una comunitat que usa un llenguatge determinat per a comunicar-se i que són regulats per una sèrie de normes i convencions compartides per tots els membres d'una comunitat, que poden acceptar-les o bé rebutjar-les (Mangiron 2006:51).

Otros muchos autores han definido el concepto de cultura en sus obras y en todas las definiciones se puede encontrar un punto de contacto entre cultura y lenguaje. Coseriu (1981:272-74) explica esta relación, que considera fundamental ya que el lenguaje, en primer lugar, es una forma primaria de cultura y de creatividad humana; en segundo lugar, refleja los saberes, las ideas y los conocimientos de una comunidad; y finalmente, requiere de una competencia lingüística y otra extralingüística que incluye el conocimiento del mundo sin el cual es imposible comunicar. Una idea compartida por diferentes autores sobre la relación entre cultura y lenguaje es la de Newmark (1988:94), sostenida también por Hurtado Albir (2001:607 cit. Ruiz Pi 2013:25), según el cual no existe una correlación unívoca entre lengua y cultura ya que, dentro de la misma comunidad lingüística, pueden existir varias agrupaciones culturales diferentes. A esta visión de la posibilidad de coexistencia de culturas diferentes se opone, en cambio, la definición de Said (1996 [1994]:14 cit. Molina Martínez 2001:22), que considera la cultura como un archivo de todo lo mejor de lo que cada sociedad ha pensado, inventado, conocido y construido, asociando así este concepto con los de nación e identidad y relacionando, además, la identidad cultural con la noción de conflicto. Delwey (1950:60-61 cit. Molina Martínez 2001:23), a su vez, sostiene que el lenguaje es un producto de la cultura; en cambio, Sapir (1929:214 cit. Molina Martínez 2001:27) defiende que una lengua puede interpretarse solo dentro de su cultura y su discípulo, Whorf, (1971:241) añade que es la estructura de la lengua la que crea las ideas en los individuos de una comunidad. A partir de estas ideas nace la hipótesis de Sapir-Whorf, según la cual el lenguaje determina (según la versión más estricta de esta hipótesis) o influye (según la versión más laxa) en la forma de pensar y percibir la realidad de una comunidad. Consecuencia de esta hipótesis sería la imposibilidad de

traducir ya que, si una lengua determina la visión de la realidad por parte de los individuos que la hablan, sería imposible representar la misma realidad en otra lengua. Esta visión, consecuencia de la hipótesis de Sapir-Whorf, se conoce como determinismo y se opone al universalismo, corriente según la cual todos los idiomas comparten una estructura común formada por elementos comunes llamados universales lingüísticos, que permiten traducir de una lengua a otra sin perder el significado.

Como consecuencia de lo anterior, surgen las teorías sobre la relación entre cultura y traducción. Desde los años 80, se empieza a ver la traducción como un elemento intercultural y al traductor como un mediador entre culturas diferentes. En 1981, Taft (1981:53 y 73 cit. Wu 2013:33) considera al traductor como mediador cultural que, a través de la interpretación de expresiones, intenciones, percepciones y expectativas de los dos grupos culturales, tiene que facilitar la comunicación partiendo de las necesidades lingüísticas y culturales de los interlocutores. Para cumplir esta tarea el traductor ha de conocer la sociedad, la historia y el folclore de las culturas del trasvase y, por cada idioma, tiene que poseer capacidades de comunicación verbal y no verbal, conocimientos técnicos que le permiten la tarea de traducir y conocimientos sociales que le permiten entender cada cultura. Siguiendo esta visión, Hatim y Mason (1995:282) hablan del traductor como comunicador entre culturas y destacan el papel sociocultural de la traducción, considerándola imposible de aislar del marco social en el cual está insertada. En 1997, Nord (1997:34) vuelve a hablar de la centralidad del factor cultural en la traducción, a la que llega a definir como una comparación de culturas. Según esta autora, como las culturas no pueden delimitarse ni tienen confines, ya que países de lenguas diferentes pueden compartir la misma cultura y pueblos que hablan la misma lengua pueden tener culturas diferentes, existen puntos de conflicto respecto de

los cuales se diferencian las culturas, produciéndose incomprensiones. La unión de estos puntos, llamados por el antropólogo Agar (1991:168 cit. Beeby, A. y otros 2000:204) *rich points*, crea una barrera cultural. En este contexto el traductor es el mediador que, a través de la traducción, lleva el texto de un lado a otro de esta barrera lingüística, social y cultural (Nord 1997:25). Katan (1999:20-21) expresa su apoyo a la idea de Hatim y Mason (1995:282 y Katan 1999:20) cuando estos definen la tarea del traductor como una mediación basada en la necesidad de comunicación entre dos partes con diferentes necesidades culturales y lingüísticas. Finalmente, Carbonell i Cortés (en comunicación personal) sostiene que

toda traducción es cultural en tanto que pone en relación al menos dos contextos culturales, con las implicaciones de identidad, diferencias en modos de vida, y hasta de visión del mundo.

1.2 El concepto y las clasificaciones de culturema

Para el concepto de culturema, ligado al de cultura y a los problemas de traducción causados por diferencias culturales, todavía no se ha encontrado una definición unívoca y reconocida, sino que, dependiendo del punto de vista y de los elementos considerados a la hora de su clasificación, cada autor utiliza términos, expresiones y clasificaciones diferentes.

El primer autor en estudiar los problemas de traducción causados por diferencias culturales fue Nida en su obra *Linguistic and Ethnology in Translation Problems* (1945). Aquí el estudioso intenta identificar los elementos que presentan diferencias culturales y que plantean consecuentes problemas de traducción y los divide en cinco ámbitos:

- ecología;
- cultura material;
- cultura social;
- cultura religiosa;
- cultura lingüística.

Desde Nida, son muchos los autores que se han interesado por estos elementos, otorgándole a cada uno de ellos su propio nombre y definición. En 1988 Newmark (1988:95) utiliza la expresión *cultural words* para definir los elementos que representan los objetos o las actividades, connotativos y específicos de una comunidad determinada, y, partiendo de la idea de Nida, propone otra clasificación en cinco categorías:

- ecología;
- cultura material;
- cultura social;
- organizaciones, hábitos, actividades, ideas y conceptos;
- gestos y formas de actuar.

Además de clasificar las palabras culturales, Newmark introduce el concepto de lenguaje cultural para referirse al lenguaje específico de una cultura, en el marco del cual, según el autor, se puede encontrar un foco cultural (o sea una gran cantidad de vocabulario sobre un tema específico propio de esa cultura) que se puede contraponer al lenguaje universal y al idiolecto. (Newmark 1988:94).

Mayoral Asensio (1994:76). en la década siguiente habla de referencias culturales y las define como

elementos del discurso que por hacer referencia a particularidades de la cultura de origen no son entendidos en absoluto o son entendidos en forma parcial o son entendidos de forma diferente por los miembros de la cultura de término.

El autor además de subrayar la dificultad de traducción de dichos elementos pone el acento en su expresividad y semejanza con los idiomatismos.

Vlakhov y Florin (1985) y la escuela eslava hablan de *realia*, y los dividen entre *realia* geográficos y etnográficos, *realia* folklóricos y mitológicos, objetos cotidianos, *realia* sociales e históricos (Chessa 2011:15)

Algunos años más tarde, Vermeer parece haber sido el primero en referirse a los elementos culturales con el término *culturemas*, que define como

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 (cfr. Nord 1997:34).

Nord (1997:34) adopta el término y la definición propuestos por Vermeer y explica que, con el término *comparados*, se apunta tanto a las situaciones en las cuales los dos elementos de la comparación tienen formas diferentes con la misma función como a aquellas otras situaciones en las que diferentes elementos tienen la misma forma con funciones diferentes.

La idea de los culturemas como elementos abstractos se puede encontrar también en el artículo de Lucía Luque Nadal «Los culturemas: ¿unidades lingüísticas, ideológicas o culturales?», de 1999, en el cual la autora los describe como elementos ilimitados y diferentes según cada sociedad y los define como

cualquier elemento simbólico específico cultural, simple o complejo, que corresponda a un objeto, idea, actividad o hecho que sea suficientemente conocido entre los miembros de una sociedad, que tenga valor simbólico y sirva de guía, referencia, o modelo de interpretación o acción para los miembros de dicha sociedad. (Nadal 2009:97).

Aunque no proponga una clasificación de los culturemas, que según la autora son imposibles de catalogar por ser ilimitados y diferentes en cada sociedad, Lucía Luque Nadal (2009:98-101) distingue entre:

- aquellos ligados a la cultura clásica, a la religión, a la historia y a los cuentos tradicionales;
- aquellos nacionales y supranacionales;
- aquellos que se refieren a personajes o situaciones arquetípicos.

Además, esta autora establece cuatro criterios que han de cumplir los culturemas para ser considerados como tales (Nadal 2009:104-107):

- vitalidad, figuratividad y motivación: el reconocimiento y la utilización del culturema en cuestión por los hablantes;
- productividad fraseológica: el número de veces en las que es utilizado en un idioma;

- frecuencia de aparición: frecuencia de aparición del culturema en los diferentes medios de comunicación (TV, radio, textos narrativos, cine, chistes, etc.);

- complejidad y estructura simbólica.

En el siglo XXI, a fin de otorgar al elemento cultural un significado más amplio, Santamaria utiliza el término *referente cultural* para definir

los objetos y eventos creados dentro de una cultura determinada con un capital cultural distintivo, intrínseco en el conjunto de la sociedad, capaz de modificar el valor expresivo que se otorga a los individuos que están relacionados con el mismo (Santamaria 2001:13).

Esta autora (Santamaria 2001:288) divide los referentes culturales en 6 categorías temáticas, cada una con ulteriores divisiones en subcategorías según las áreas:

Ecología

Geografía y topografía

Meteorología

Biología

Ser humano

Historia

Edificios

Eventos

Personalidad

Estructura social

Trabajo

Organización social

Política

Instituciones culturales

Bellas artes

Arte

Religión

Educación

Medios de comunicación

Universo social

Condiciones sociales

Geografía cultural

Transporte

Cultura material

Alimentación

Indumentaria

Cosmética y peluquería

Ocio

Objetos materiales

Tecnología

Ese mismo año, Molina Martínez (2001:91) utiliza el término *culturema* y adopta la definición de Nord, añadiéndole el concepto de dinamicidad, según el cual los culturemas se desarrollan y aumentan ilimitada y continuamente. Además, esta misma autora subraya la importancia del contexto de los culturemas, afirmando que existen solo en una situación de transferencia cultural, es decir en el momento de traducir, y que tienen una estricta dependencia de esta. Así la autora divide los culturemas en cuatro ámbitos

- medio natural;
- patrimonio cultural;
- cultura social;
- cultura lingüística;

a los cuales añade otra categoría, la de la interferencia cultural, de la cual forman parte los falsos amigos y la injerencia culturales (Molina Martínez 2001:89-98).

En la misma década, Carme Mangiron (2006:63). opta por la expresión *referencia cultural* y define los culturemas como «elements discursius presents en un text que al·ludeixen a una cultura determinada i aporten significat, expressivitat o color local (o una combinació de tots ells)».

Esta autora (Mangiron 2006:138-140) propone una división de los elementos culturales aún más detallada, clasificándolos en 7 categorías:

- medio natural;
- historia;

- cultura social;
- instituciones culturales;
- cultura material;
- cultura lingüística;
- injerencias culturales.

Cada una además tiene ulteriores clasificaciones y subcategorías muy detalladas que se presentan a continuación:

- | | | | |
|------------------|---------------------------------------|-----------------------|-----------------------------|
| - Medio natural | Geología | | |
| | Biología | Flora | |
| | | Fauna | |
| - Historia | Edificios | | |
| | Acontecimientos históricos | | |
| | Instituciones y personajes históricos | | |
| | Símbolos nacionales | | |
| - Cultura social | Trabajo | Unidades de medida | |
| | | Unidades monetarias | |
| | Condiciones sociales | Antropónimos | Antropónimos convencionales |
| | | | Antropónimos simbólicos |
| | | Relaciones familiares | |

| | | |
|----------------------------|---------------------------------------|-------------------------------|
| | | Relaciones sociales |
| | | Costumbres |
| | | Geografía cultural |
| | | Transporte |
| - Instituciones culturales | Bellas artes | Pintura, cerámica y escultura |
| | | Artes floreales |
| | | Música y danza |
| | Arte | Teatro |
| | | Literatura |
| | Religión | |
| | Educación | |
| - Cultura material | Hogar | |
| | Alimentación | Comidas |
| | | Bebidas |
| | Indumentos | |
| | Tiempo libre | Juegos |
| | | Deportes y artes marciales |
| | | Hoteles y restaurantes |
| | Objetos materiales | |
| - Cultura lingüística | Sistema de escritura | |
| | Dialectos | |
| | Expresiones, refranes y frases hechas | |

| | | |
|--------------------------|--|-------------------------------|
| | Juegos de palabras | |
| | Insultos | |
| | Onomatopeyas | |
| - Injerencias culturales | Referencias a otros idiomas | |
| | Referencias a instituciones culturales | Pintura, cerámica y escultura |
| | | Literatura |
| | Referencias históricas | |

De todas formas, Mangiron (2006:137) sostiene que las categorías tendrían que decidirse según el contenido y los referentes culturales que aparecen en cada texto analizado.

Después de una breve presentación de las expresiones y de las definiciones utilizadas por los autores previos para referirse a los culturemas, Carbonell i Cortés propone su teoría recientemente (como parte del curso de metodología de la investigación en traducción literaria del máster cursado). En su opinión, un *realia* o elemento cultural específico es un «elemento de un texto que plantea una dificultad de traducción debido a su contraste con el referente habitual en la cultura de destino. Es, por lo tanto, un problema de equivalencia» (Carbonell i Cortés: en comunicación personal).

1.3 Propuesta de clasificación de culturemas

Partiendo de las definiciones y de las clasificaciones de los autores revisados arriba, para este estudio se utilizará el término culturema por varios motivos. De hecho esta palabra, así como la han definido los estudiosos, destaca con especial vehemencia la centralidad del factor cultural en la traducción, enfatiza la relación existente entre dos

culturas y la comparación de los elementos entre culturas diferentes y finalmente, como subraya Molina Martínez (2001:91), hace un especial hincapié en el contexto de la traducción ya que los culturemas existen solo en relación con su transferencia de una cultura a otra.

Por estas mismas razones, la definición que se adopta en el presente estudio será la de Christiane Nord (1997:34), que la propia autora atribuye a Vermeer: el culturema es un elemento abstracto propio de una cultura y útil por su comparación entre esta y otras culturas diferentes. En consecuencia, el término culturema puede apuntar tanto a elementos de lengua como de comportamiento.

La clasificación que se utilizará en el marco del presente trabajo adoptará, como punto de partida, una propuesta resultante de las vistas hasta ahora. Ello permitirá iniciar el análisis a partir de una perspectiva múltiple con la que deseamos cubrir la más amplia casuística posible. Por lo demás, dicha clasificación será susceptible de enmienda a partir de los datos textuales que se manejen, más adelante, en los apartados analíticos de nuestro estudio.

Así, en la clasificación que proponemos aparecen 7 categorías básicas:

- cultura geográfica y natural: incluye todas las referencias a sitios geográficos, topónimos, flora y fauna y resumiría las categorías “ecología” de Nida, Newmark y Santamaria; la de “medio ambiental” de Lucía Molina Martínez, y la de “medio natural” de Mangiron;
- cultura material: en la cual se incluyen todos los objetos, los productos, las obras de arte específicas de una cultura y que recoge la “cultura material” de Nida, Newmark,

Mangiron y Molina Martínez; “los *realias* de los objetos cotidianos” de Vlachov y Florin; y la categoría “patrimonio cultural” de Lucía Molina Martínez;

- cultura social, histórica y de folclore: en la cual se insertan todos los hábitos sociales, los nombres propios de personajes famosos, históricos y políticos, las acciones cotidianas típicas de un pueblo determinado, todo lo que se refiere a los trabajos, al tiempo libre, al deporte, a los orígenes de un pueblo, de una etnia, de un país, etc. Esta categoría incluiría las categorías “cultura social” de Nida y Newmark; los “*realia* sociales e históricos” de Vlachov y Florin; y la de la cultura social de Lucía Molina Martínez;

- cultura religiosa: que comprende todas las formas de práctica religiosa, las ideas de sacralidad, los cultos de una cultura. Forman parte de esta categoría la cultura religiosa de Nida; los *realia* mitológicos de Vlachov y Florin, etc.;

- cultura conceptual: categoría de la cual forman parte los conceptos y las ideas políticas, filosóficas de una cultura determinada, de las categorías de ideas y conceptos de Newmark;

- cultura lingüística: que incluye las formas idiomáticas, los refranes y los elementos lingüísticos de una cultura determinada e incluiría las categorías de la cultura lingüística de Nida y Newmark;

- interferencias culturales: que comprenden las inequivalencias, los falsos amigos culturales y la injerencia cultural y deriva directamente de las categorías añadidas por Newmark a las de Nida y de la de Lucía Molina Martínez.

Algunas de las categorías enumeradas aquí arriba además se dividen en subcategorías que resumen más detalladamente las propuestas de los autores analizados hasta ahora.

De modo que la categoría de la cultura geográfica i natural se divide en:

- flora y fauna (que resume las categorías “ecología” propuesta por Nida y Newmark y los *realias* geográficos de Vlahov y Florin);

- topónimos y referencias a lugares específicos (con referencia a las propuestas de la subcategoría geografía y topografía de Santamaria);

La categoría de la cultura material reúne todas las referencias a los objetos específicos de una cultura.

La cultura social, histórica y de folclore se divide en las subcategorías de:

- referencias a personajes históricos, políticos y culturales, a obras de arte y a hechos de importancia histórica (procedente de la categoría de la cultura social de Nida, de los *realia* sociales, históricos y folclorísticos de Florin, de la categoría de las instituciones culturales propuesta por Santamaria y de la categoría de los elementos culturales que se refieren a instituciones culturales teorizada por Lucia Luque Nadal);

- familia (de la subcategoría “relaciones familiares” propuesta por Mangiron);

- amistad (procedente de la subcategoría “relaciones sociales” de Mangiron);

- organizaciones y papeles sociales (de la categoría de organizaciones y ámbitos descrita por Newmark y las propuestas de las subcategorías de trabajo, organización social, condiciones sociales y política de Santamaria);

- acciones de vida cotidiana (de la unión entre la categoría “gestos y formas de actuar” de Newmark y la subcategoría “tiempo libre” de Mangiron).

La categoría de la religión y la de la cultura conceptual se mantienen intactas mientras que la categoría de la cultura lingüística se divide en:

- juegos de palabras, refranes, etc. (subdivisión de Nida, también propuesta por Mangiron como subcategoría de refranes, expresiones y frases hechas);
- nivel de formalidad (de las subcategorías “Insultos” y “dialectos” de Mangiron);
- tabús (subcategoría de la categoría de la cultura religiosa de Nida).

Finalmente la categoría de la interferencia cultural se subdivide en:

- inequivalencias (desde la propuesta de subcategoría de referencias a otros idiomas por Mangiron);
- falsos amigos culturales (de la clasificación de Lucía Molina Martínez);
- injerencia cultural (categoría propuesta por Lucía Molina Martínez).

Así, el sistema de categorías definitivo que se utilizará en esta investigación se resume en la ficha que se incluye a continuación, en cuya primera columna se presentan las categorías de culturemas que se utilizarán en el presente estudio; en la segunda columna se registran las subcategorías y en la tercera se recoge un ejemplo de elemento cultural que representa cada categoría:

| Categoría culturema | Subcategorías | Ejemplo |
|--------------------------------|----------------------|----------------|
| | | |

| | | |
|---|--|--------------------|
| Cultura geográfica y natural | Flora y fauna | Zebra |
| | Topónimos y referencias a lugares específicos | Harvard University |
| | Referencias a un grupo étnico o a una tribu | Los Sioux |
| Cultura material | Referencias a objetos específicos de una cultura | Dumplings |
| Cultura social, histórica y de folclore | Referencias a personajes históricos y políticos | Botticelli |
| | Familia | Hermanito |
| | Amistad | Colega |
| | Organización y papeles sociales | Sir |
| | Acciones de vida cotidiana | Merendar |
| Cultura religiosa | Religión | Padrenuestro |

| | | |
|------------------------|---|--|
| Cultura conceptual | Creencias, tradiciones y supersticiones | Gato negro |
| Cultura lingüística | Juegos de palabras, refranes, etc. | Hasta el diablo tiene que bajar la vista para verla |
| | Nivel de formalidad | Usted |
| | Tabús | Tabús religiosos, políticos, etc. |
| Interferencia cultural | Inequivalencias | Bagel |
| | Falsos amigos culturales | Búho o lechuza (en la cultura occidental indica sabiduría y en la cultura oriental mal agüero) (Molina Martínez 2001:98) |
| | Injerencia cultural | Flamenco (utilizado en inglés para representar algo exótico y típico de otra cultura que en español se traduce con otro término típico de una cultura diferente) |

Tabla 1: clasificación de culturemas

1.4 Marcas culturales y pseudo-marcas culturales

Helena Tanqueiro (2001) se refiere a los culturemas como a marcas culturales y los define como

denominaciones, expresiones, referencias presentes en un texto (literario) cuyo valor connotativo y/o carga afectiva son entendidas y

compartidas, de forma implícita, por el autor solo con los integrantes del área cultural a la que se dirige la obra y que están vetadas o provocan asociaciones falsas a los integrantes del área cultural a la que va dirigida la traducción (Soria Orti 2012:19 cit. Tanqueiro 2001).

Según esta estudiosa, el autor y el lector comparten unos conocimientos previos que les permiten entender aquellas referencias de un texto que tienen un valor connotativo o una carga afectiva importante para la comprensión del texto mismo. En este sentido, las marcas culturales se pueden considerar portadoras de cultura y el traductor que no las detecta y, por consiguiente no las transfiere al idioma meta, comete un error de infratraducción, o sea la omisión de elementos con una carga semántica importante a la hora de interpretar un texto. (Tanqueiro 2002: 118-119).

Según la misma autora, en un texto literario existen también elementos que parecen marcas culturales, pero en realidad la referencia a la cultura de partida es utilizada solo como convención textual sin ningún valor connotativo. En este caso, la autora habla de pseudo-marcas culturales (Soria Orti 2012:19 cit. Tanqueiro 2001) y añade que, si el traductor no distingue la pseudo-marca cultural de la marca cultural y al traducir le otorga valor semántico ausente en el original, cumple un error de sobretraducción, es decir otorga demasiada importancia a un elemento que originariamente no tenía ningún valor semántico (Tanqueiro 2002:119).

2 TRATAMIENTO DE CULTURA Y CULTUREMAS: DEFINICIÓN, CLASIFICACIONES Y PROPUESTA TEÓRICA

Después de haber presentado los culturemas y haber explicado cómo y por qué pueden ser causas de problemas de traducción, en este capítulo se describen los tratamientos que diferentes autores han propuesto para trasvasar los culturemas manteniendo la misma carga semántica y cultural que tienen estos en el texto original. Con este propósito, en primer lugar se establece la diferencia entre técnica, estrategia y método de traducción, necesaria para motivar la elección del término *técnica de* utilizado en este estudio. En segundo lugar, se explicitan los problemas que provocan los culturemas durante la traducción, los factores que determinan la elección de las técnicas y las clasificaciones de estas con arreglo a los postulados anteriores. Finalmente se presenta aquí una propuesta propia de técnicas que se utiliza a lo largo del presente estudio y se explican las razones de su elección.

2.1 Diferencia entre técnica, estrategia y método de traducción

Antes de explicar los tratamientos de los culturemas, es imprescindible distinguir y definir los conceptos de método, estrategia y técnica de traducción. Según Hurtado Albir (2001:633-645 cit. Ruiz Pi 2013:31), estos tres conceptos, se sitúan en una escala que se extiende desde lo más general hasta lo más particular. En primer lugar, partiendo de una visión general, el traductor escoge el método que determinará el proceso de traducción. En segundo lugar, toma decisiones sobre dicho proceso en su totalidad a través de la elección de estrategias de traducción. Finalmente escoge procedimientos concretos, es decir técnicas de traducción, utilizadas para cada elemento del texto con el

objetivo de alcanzar la equivalencia entre texto de partida y de llegada. La misma autora (Hurtado Albir 2001:633-645 cit. Ruiz Pi 2013:31), define cada uno de estos conceptos:

Método traductor: desarrollo de un proceso traductor determinado regulado por unos principios en función del objetivo del traductor, respondiendo a una opción global que recorre todo el texto. Los métodos cambian en función del contexto y de la finalidad de la traducción.

Estrategia traductora: procedimientos, conscientes e inconscientes, verbales y no verbales, internos y externos, utilizados por el traductor para resolver los problemas encontrados en el desarrollo del proceso traductor y mejorar su eficacia en función de sus necesidades específicas.

Técnica de traducción: procedimiento, visible en el resultado de la traducción, que se utiliza para conseguir la equivalencia traductora a microunidades textuales; las técnicas se catalogan en comparación con el original. La pertinencia del uso de una técnica u otra es siempre funcional, según el *tipo textual*, la *modalidad de traducción*, la finalidad de la traducción y el *método* elegido.

Aceptando la misma clasificación, Lucía Molina Martínez (2006:100-101) añade algunas características de las técnicas de traducción, que en su opinión afectan al resultado de la traducción misma. Estas técnicas se catalogan con respecto al original, se refieren a microunidades textuales, tienen características discursivas y contextuales y son funcionales.

A pesar de las definiciones de Hurtado Albir (2001: 633-645 cit. Ruiz Pi 2013:31), los teóricos de la traducción mezclan la terminología y la utilizan de un modo personal, refiriéndose a las soluciones de traducción con los términos métodos, procedimientos, procesos y técnicas indiscriminadamente. Sin embargo, siendo el presente estudio contrastivo y basado en el análisis de extractos concretos, se considera que el concepto más útil para desarrollar esta investigación es el de la técnica, que permite proceder con un análisis desde lo más concreto a conclusiones más generales. Así que, para referirse al tratamiento de los culturemas utilizado a lo largo del proceso de transferencia de una lengua a otra o de una cultura a otra, se utilizará la expresión *técnica de traducción*. En resumen y aceptando la definición de Hurtado Albir (2001:633-645 cit. Ruiz Pi), por técnica de traducción se entiende ese procedimiento que se aplica a microunidades textuales, que en este caso específico corresponden a cada uno de los culturemas identificados.

2.2 Problemas derivados de la traducción de los culturemas

Como ya se ha mencionado en el capítulo anterior, la presencia de culturemas crea problemas en la traducción por la naturaleza de estos elementos. Siendo los culturemas parte integrante de una cultura, es difícil transferirlos de una lengua a otra manteniendo la misma carga semántica, porque esto significaría trasplantar un elemento propio de una cultura X a otra cultura Y, en la cual ese elemento no tiene la misma importancia o el mismo reconocimiento que tiene en la cultura de partida o simplemente no existe. De hecho, consideradas estas dificultades, Carbonell i Cortés (en comunicación personal) define el elemento cultural como «cualquier elemento de un texto que plantea una dificultad de traducción debido a su contraste con el referente habitual en la cultura de destino. Es, por lo tanto, un problema de equivalencia».

Después de haber clasificado los elementos culturales, Stolze (1992 cit. Tanqueiro2002:105) intenta agrupar las soluciones de traducción útiles para hacer frente a las incongruencias culturales que generan estos elementos. Según este autor, existen tres tipos de incongruencias:

- incongruencias reales: cuando en el texto original se hace referencia a algo que no existe en el texto de llegada;
- incongruencias formales: cuando en el texto original hace referencia a algo que en la cultura de llegada existe pero se expresa a través de una forma diferente;
- incongruencias semánticas: en el caso de referencias con connotaciones falsas o que provocan reacciones negativas.

2.3 Factores que determinan la elección de las soluciones de traducción de culturemas

Algunos autores han intentado localizar los elementos de los que depende la elección de las soluciones traductoras de los culturemas. Entre estos, Newmark (1988:103), sugiere que para transferir los elementos culturales de una lengua a otra hay que considerar seis factores:

- finalidad del texto;
- motivación y nivel cultural de los lectores;
- importancia del referente cultural en el texto original;
- ámbito;
- novedad del término;

- futuro del término.

En 1993, Florin (cit. Trujillo 2012:5) sostiene que la elección de la técnica de traducción depende de cinco factores:

- el tipo de texto;
- la importancia del elemento cultural que ha de traducirse en el contexto y en el texto;
- la naturaleza de dicho elemento cultural (o sea el registro, el grado de familiaridad, etc.);
- las parejas de lenguas entre las cuales se traduce;
- las características del lector.

También Hurtado Albir (2001 cit. Chessa 2011: 28-29), más tarde, presenta los que en su opinión son los factores de partida que hay que considerar para traducir los culturemas:

- la relación entre cultura de partida y cultura meta;
- el género textual de la obra original;
- la función de culturema original en el texto;
- la naturaleza del culturema en cuestión (es decir sus características como la universalidad, el grado de novedad, el registro, etc.);
- las características del destinatario;
- la finalidad de la traducción, que determina el método y las técnicas de traducción.

Para todos los autores analizados entonces, los factores esenciales que hay que considerar para escoger las soluciones de traducción apropiadas a cada situación, son las características propias del culturema, su contexto, las finalidades de traducción y las características del destinatario.

2.4 Clasificaciones de los tratamientos para traducir los culturemas

A lo largo de la historia de la traducción, muchos son los autores que han intentado proponer y clasificar las soluciones de traducción para los elementos culturales.

Ya en 1991, Hewson y Martin (cit. Chessa 2011:25), considerando la relación entre la lengua de partida y la lengua de llegada, proponen cuatro soluciones de traducción:

- reducción, si el sistema de la lengua de partida domina el sistema de la lengua de llegada;
- marginalización, si el sistema de la lengua de llegada domina el sistema de la lengua de partida;
- inserción, si existe la posibilidad de integrar algunos valores culturales;
- conversión, si en los dos sistemas lingüísticos se utilizan sistemas parecidos;

Vinay y Darbelnet (1995:31-40), en cambio, se interesan por la traducción específica de los culturemas y, en su obra *Comparative stylistics of French and English* (1958), denominan los métodos de traducción procesos o procedimientos y los dividen entre traducción directa, o sea literal, y traducción oblicua. Además proponen siete «*methods of translation*» que los mismos autores definen «*processes*» o «*procedures*»:

- préstamo (utilización de una palabra original de la lengua de partida);

- calco (traducción del contenido semántico de una palabra o de una expresión extranjera, mediante unidades lingüísticas propias de la lengua de llegada);
- traducción literal (mantenimiento del contenido, de la forma y del estilo del término original);
- transposición (producción de un cambio de categoría gramatical);
- modulación (producción de un cambio en la forma del mensaje debido a una variación del punto de vista);
- equivalencia (utilización de una expresión semánticamente equivalente);
- adaptación (adaptación de la expresión en lengua original a una expresión propia de la lengua de llegada, con un referente distinto del original pero con una connotación cultural similar).

En 1964, Nida (1964:226-238) se refiere a las soluciones de traducción definiéndolas como técnicas de ajuste y, de acuerdo con Taber y Margot (cit. Chessa 2011:25), propone tres soluciones para traducir los culturemas cuando en la lengua meta no existe un equivalente:

- adición (adición de elementos al texto original);
- omisión (omisión de elementos del texto original);
- conversión (utilización de valores parecidos en la lengua meta).

Sucesivamente el autor añade otras soluciones a las ya presentadas: la utilización de préstamos, de términos clasificadores o de equivalentes descriptivos y la sustitución cultural (la sustitución de un elemento problemático con otro que en la lengua de

llegada tenga la misma función). Además, el mismo estudioso, incorpora a las técnicas de ajuste también las notas a pie de página para añadir informaciones sobre el contexto sociocultural original (Nida 1964:238-239).

En 1970 Vlahov y Florin (cit. Chessa 2011:27) en cambio sugieren seis métodos de traducción:

- transcripción (transposición de los sonidos de una palabra original de la lengua de partida por medio de la fonética de la lengua meta);
- formación de un término nuevo (creación de un término inexistente en la lengua de llegada);
- asimilación cultural (adopción de elementos culturales propios de la cultura de partida);
- traducción aproximada (traducción del significado general del término de la lengua de origen sin equivalencia formal);
- traducción descriptiva (traducción de un término mediante su paráfrasis).

En los años siguientes, Vázquez Ayora (1977:251-384) habla de procedimientos técnicos de ejecución identificando entre estos:

- transposición (producción de un cambio de categoría gramatical);
- modulación (producción de un cambio en la forma del mensaje debido a una variación del punto de vista);
- equivalencia (utilización de una expresión semánticamente equivalente);

- adaptación (adaptación de la expresión en lengua original a una expresión propia de la lengua de llegada, con un referente distinto del original pero con una connotación cultural similar);
- amplificación (adición de elementos lingüísticos);
- explicitación (explicitación de elementos mantenidos implícitos en el texto original);
- omisión (eliminación de elementos lingüísticos);
- compensación (pérdida de un elemento compensada por la adición de otro).

También Newmark (1988:103-104) utiliza el término procedimientos y propone su propia clasificación:

- traducción literal (mantenimiento del contenido, de la forma y del estilo del término original);
- naturalización (adaptación de la traducción a la pronunciación en la lengua meta);
- equivalente cultural (utilización de un referente distinto de la lengua original pero con una connotación cultural similar);
- equivalente funcional (adición de elementos para explicar y especificar la función que desempeñan un término o una expresión en la lengua original);
- equivalente descriptivo (descripción del término original);
- sinónimo (utilización de palabras diferentes pero con el mismo significado);
- calco (traducción del contenido semántico de una palabra o de una expresión extranjera, mediante unidades lingüísticas propias de la lengua de llegada);

- transposición (producción de un cambio de categoría gramatical);
- modulación (producción de un cambio en la forma del mensaje debido a una variación del punto de vista);
- traducción estándar aceptada (utilización de una traducción reconocida en la lengua meta);
- etiquetas de traducción (utilización de traducciones provisionales);
- compensación (pérdida de un elemento compensada por la adición de otro);
- análisis componencial (explicitación de las características el culturema);
- reducción y expansión (omisión o adición de informaciones);
- paráfrasis (explicación del término original);
- expresiones equivalentes (utilización de expresiones idiomáticas diferentes que mantienen equivalencia semántica);
- adaptación (adaptación de la expresión de la lengua original a una expresión propia de la lengua de llegada, con un referente distinto del original pero con una connotación cultural similar);
- glosas y notas (adición de explicaciones en el texto mediante glosas y notas).

En 1992 Hervey y Higgins (cit. Chessa 2011:28) argumentan que en cualquier texto se encuentran filtros culturales, formales, semánticos, lingüísticos y de género. Los filtros culturales, como también los culturemas, necesitan de una transposición en el pasaje de una lengua a otra. Los autores consideran que las diferentes soluciones de transposición

se pueden ordenar en una escala, desde la que más se acerca al exotismo (o al texto original) hasta la que más se acerca a la adaptación (es decir, al texto meta) (cit. Chessa 2011:28). Estos dos autores proponen un continuo, como una línea que une el texto de partida con el texto de llegada, a lo largo de la cual se desarrollan las diferentes soluciones de traducción de los elementos culturales.

Lucía Molina Martínez (2006:101-104) también expone su propia clasificación, dividiendo las técnicas traductorales de los elementos culturales en dieciocho categorías:

- adaptación (sustitución de un elemento cultural con otro propio de la cultura meta);
- ampliación lingüística (añadir elementos lingüísticos);
- amplificación (añadir informaciones y explicaciones);
- calco (traducción literal de una palabra o de una sintagma extranjero);
- compensación (pérdida de un elemento compensado por otro añadido en otra parte del texto);
- compresión lingüística (síntesis de elementos lingüísticos);
- creación discursiva (establecimiento de una equivalencia efímera e imprevisible);
- descripción (sustitución de término con su descripción);
- equivalente acuñado (traducción estándar aceptada);
- generalización (utilización de un término más general o neutro);
- modulación (cambio punto de vista);
- particularización (utilización de un término más preciso o concreto);

- préstamo (utilización de una palabra original de la lengua de partida);
- reducción (omitir informaciones y explicaciones presentes en el texto original);
- sustitución (sustituir elementos lingüísticos con elementos paralingüísticos o viceversa);
- traducción literal (mantenimiento del contenido, de la forma y del estilo del término original);
- transposición (cambio de categoría gramatical);
- variación (cambio de elementos lingüísticos o paralingüísticos que afectan aspectos de la variación lingüística).

En los mismos años, Mangiron (2006:110) vuelve a adoptar la idea de un continuo entre la obra original y la obra meta y propone cuatro estrategias de traducción, ordenadas de la que más se acerca a la cultura de origen a la que más se acerca a la cultura del lector meta:

- conservación (que incluye transferencia, ampliación, reducción y sustitución);
- neutralización (explicación de un elemento mediante palabras que describen su función o características);
- omisión (eliminación de un elemento del texto original);
- adaptación (adaptación de la expresión de la lengua original a una expresión propia de la lengua de llegada, con un referente distinto del original pero con una connotación cultural similar)

Según la misma autora, cada una de estas estrategias corresponde un comportamiento del traductor, así que al enfrentarse con un elemento cultural en el texto original, el traductor tiene cuatro opciones (Mangiron 2006:110-114):

- mantener la referencia a la cultura original;
- neutralizar la referencia a la cultura original;
- omitir la referencia a la cultura original;
- adaptar la referencia a la cultura meta.

Carbonell i Cortés sigue la misma línea de Mangiron. Partiendo de la teoría según la cual en el lenguaje se encuentran unos modelos a través de los cuales se representa la realidad, el autor considera que el lector construye su propia imagen e identidad, a las cuales sigue, por consecuencia, la construcción de la imagen de los demás como “otros”, diferentes de la autopercepción y realidad propias. Prosiguiendo con esta idea, el autor presenta el dilema de Schleiermacher, o sea la tendencia hacia el texto y la cultura original bien hacia el texto y la cultura meta (Carbonell i Cortés: en comunicación personal). Esta oposición refleja la dicotomía entre equivalencia formal y dinámica teorizada por Nida (1964) y es de aquí de donde Carbonell i Cortés (en comunicación personal) deduce que todos los textos representan bien la alteridad o extranjería (es decir a través de referentes exóticos presentan una realidad ajena), bien la identidad o la ideología (es decir que a través de referentes familiares presentan una realidad cercana o propia, en la cual los lectores se pueden identificar). En consecuencia, el traductor para cumplir con su función de intermediario intercultural entre autor y lector tiene dos posibilidades: acercar el lector a la realidad del autor o acercar el autor a la realidad del lector. En el primer caso utilizará la técnica del

exotismo, mientras en el segundo la de la identidad. A partir de estos dos extremos, Carbonell i Cortés teoriza la existencia de un continuo que se extiende desde el exotismo hacia la identidad, a lo largo del cual se pueden disponer las diferentes estrategias de traducción de los elementos culturales (Carbonell i Cortés en comunicación personal) desde el préstamo (el más cercano a la cultura de origen y entonces al exotismo), hasta la creación autónoma (más cercana a la cultura meta y entonces a la adaptación cultural). Las estrategias de traducción que el autor incluye en ese continuum, ordenadas de la más cercana al exotismo hasta la más cercana a la adaptación cultural, son:

- préstamo / transferencia (utilización de una palabra original de la lengua de partida);
- naturalización (préstamo adaptado a la pronunciación, ortografía y morfología de la lengua de llegada);
- traducción literal (mantenimiento del contenido, de la forma y del estilo del término original);
- neutralización / explicitación (explicación de un elemento mediante palabras que describen su función o características);
- hiperónimo o universalización (generalización);
- pareja o doblote (combinación de dos técnicas de traducción);
- traducción estándar aceptada (utilización de una traducción reconocida en la lengua meta);
- paráfrasis o glosa (adición de explicaciones en el texto);

- supresión (omisión de elementos);
- adaptación intracultural (sustitución por otro referente de la cultura original, pero más transparente);
- equivalente cultural o adaptación intercultural (utilización de un referente distinto de la lengua original pero con una connotación cultural similar);
- creación autónoma (adición de un elemento ausente en el texto original).

2.5 Propuesta de clasificación de tratamiento de los culturemas

A partir de las propuestas de los autores previos, se ha determinado una clasificación de técnicas de traducción de los culturemas que pueda ser utilizada en el presente estudio. Esta propuesta se basa sobre todo en la clasificación de Carbonell i Cortés presentada en el párrafo anterior, menos específica que la de Molina Martínez (2006:110-114) y más compleja que la de Mangiron, que se limita al listado de las posibilidades de traducción, con respecto al mantenimiento o la pérdida de los elementos culturales del texto original. Sin embargo, se han añadido a la clasificación propuesta por Carbonell i Cortés algunas técnicas más, avanzadas por autores arriba citados, que parecen interesantes y útiles para los fines de esta investigación. Además se ha mantenido la idea de una línea recta a lo largo de la cual se encuentran las diferentes técnicas de traducción, de la que más fiel permanece a la cultura del texto original a la que más se acerca a la cultura del lector del texto meta, idea propuesta inicialmente por Hervey y Higgins (1992 cit. Chessa 2011:28) y seguida por Mangiron (2006) y Carbonell i Cortés (en comunicación personal).

El resultado por el que aquí se aboga se compone de 13 técnicas, obtenidas añadiendo, omitiendo o fusionando algunas de las presentadas arriba. En la tabla que sigue, se ordenan desde la que más se acerca a la cultura original y mantiene sus rasgos hasta la que más se acerca a la cultura del lector de la lengua meta, según el continuo anteriormente explicado y representado por la flecha que aparece junto a la tabla. En la primera columna se presenta la técnica de traducción, en la segunda la explicación de dicha técnica y en la tercera columna un ejemplo de cada técnica, para que sea más fácil entender su uso. Por la imposibilidad de presentar ejemplos de notas a pie de página, de omisión y de creación autónoma fuera del texto, las casillas de la tercera columna correspondiente a estas técnicas se quedarán vacías. Además se utilizará el símbolo > para indicar la traducción de una lengua a otra (por ejemplo con la fórmula A > B se entiende que A ha se ha traducido a B).

+ exotismo

| Técnica de traducción | Explicación | Ejemplo |
|------------------------------|---|---------------------------------|
| Transposición | Mantenimiento de la palabra en lengua original | Fish and Chips > Fish and Chips |
| Préstamo | Mantenimiento de la palabra en lengua original en el caso de palabras ya reconocidas por los diccionarios | Web > Web |
| Naturalización | Adaptación a la fonética de la lengua meta | Check > Chequeo |

| | | |
|------------------------------|---|---|
| Traducción literal | Traducción literal del culturema | Email > Correo electrónico |
| Neutralización | Explicación con palabras que describen su función o características | turf > cuidado césped del patio (Carbonell i Cortés:en comunicación personal) |
| Hiperónimo | Generalización | Machete > knife |
| Doblete | Uso de dos o más estrategias de traducción para traducir el mismo culturema | Combinación de transferencia + neutralización: Tortilla > Tortilla, una frittata di patate. |
| Traducción estándar aceptada | Traducción aceptada por los diccionarios y las gramáticas de la lengua meta | Federal Court > Corte Federal (Carbonell i Cortés: en comunicación personal) |
| Paráfrasis explicativa | Adición de informaciones en el interior del texto | Gondola > Góndola, esas barcas típicas de los canales venecianos |
| Notas a pie de página | Adición de informaciones en notas a pie de página | Pregare santa Lucia > Rogar a santa Lucía ¹ : santa Lucía en Italia es famosa por ser la santa protectora de la vista. |
| Supresión de | Omisión de informaciones sobre el culturema presentes en el texto | Góndola, esas barcas típicas de los canales venecianos > |

| informaciones | original | Gondola |
|----------------------------------|--|--|
| Omisión | Omisión total del culturema | Guardare Studio Aperto in televisione > mirar la televisión |
| Equivalente cultural o funcional | Palabras diferentes de las originales pero que mantienen la misma función cultural | <i>BA degree</i> > licenciatura (Carbonell i Cortés: en comunicación personal) |
| Creación autónoma | Introducción de un elemento cultural ausente en el texto original | Terminare gli studi > Sacarse la licenciatura |

Tabla 2: clasificación de técnicas de traducción de culturemas

+ adaptación cultural



3. METODOLOGÍA: LA LINGÜÍSTICA DE CORPUS Y LOS MÉTODOS DEL PRESENTE TRABAJO

Este capítulo se divide en una parte teórica general y otra más concreta y aplicada a este estudio. En la primera se explica la lingüística de corpus como método de investigación, sus objetivos, ventajas y límites. Además se recurre, brevemente, a la historia de los estudios traductológicos basados en corpus, y se presentan las clasificaciones, los criterios según los cuales se compilan los corpus y las herramientas que se utilizan para analizarlos. En la segunda parte se aplican dichos conceptos al presente estudio y se presentan los corpus utilizados junto con su metodología de compilación y las herramientas de análisis que se utiliza para alcanzar los objetivos propuestos en el capítulo que sigue a este.

3.1 Estudios traductológicos de corpus

3.1.1 Lingüística de corpus: definición y características

3.1.1.1 Definición de lingüística de corpus y su concreción en la traductología

Un corpus, del latín *corpus* que significa cuerpo, es un conjunto de documentos utilizado desde la antigüedad para estudiar rasgos comunes a todos ellos. En los años 90, Sinclair (1991:171) intenta definir el corpus como una colección de textos naturales (*naturally occurring*), escogidos para caracterizar un estado o una variedad de lenguaje. Esta definición aclara que los textos que componen un corpus son naturales, es decir producidos por seres humanos, y auténticos, o sea utilizados en contextos reales y ya existentes en la lengua. Esto permite determinar que aunque el corpus sea artificial, es decir construido con fines específicos, los textos que lo componen son naturales. De

aquí la aportación de Mercado Percia (2008:4), que subraya la diferencia entre lengua y corpus, ya que el corpus representa una lengua y esa representación puede tener diferentes grados de fiabilidad. Sin embargo la definición de Sinclair no aporta ninguna información sobre el tamaño y la selección del corpus que, según Berber Sardinha (2000:337-338), para ser útil a los fines de una investigación, ha de ser amplio (con el fin de reflejar la variedad de la lengua elegida de la forma más fiel posible) y tiene que incorporar solo el material útil para el objeto de investigación. Sánchez (1995:8-9) aporta una definición más completa, que reúne todas estas características al tiempo que menciona el origen, el propósito, la composición, la formación la representatividad y la extensión de todo corpus:

Un conjunto de datos lingüísticos (pertenecientes al uso oral o escrito de la lengua, o a ambos) sistematizado según determinados criterios, suficientemente extensos en amplitud y profundidad de manera que sean representativos del total del uso lingüístico o de algunos de sus ámbitos, y dispuestos de tal modo que puedan ser procesados mediante ordenador con el fin de obtener resultados varios y útiles para la descripción y el análisis.

En 2001, Tognini-Bonelli (2001:55) definen el corpus como:

a computerised collection of authentic texts, amenable to automatic or semi-automatic processing or analysis. The texts are selected according to explicit criteria in order to capture the regularities of a language, a language variety or a sub-language.

Además de estas definiciones generales, Mercado Percia (2008:6) propone otras, complementarias, desde puntos de vista más concretos:

- desde el punto de vista morfológico, el corpus es un conjunto de palabras derivadas de una lengua;
- desde el punto de vista sintáctico, el corpus es un conjunto variado de frases de una lengua;
- desde el punto de vista de la programación neurolingüística, el corpus es un recurso lingüístico que permite construir herramientas para el tratamiento del lenguaje natural.

Entre los estudiosos de lingüística y traducción se ha discutido ampliamente sobre la naturaleza de la lingüística de corpus: algunos la consideran una metodología, otros una disciplina. Por un lado, hay que considerar que la lingüística de corpus no tiene un objetivo delimitado y se ocupa de varios fenómenos, por lo que resulta ser una metodología. Algunos de los autores que se han mostrado de acuerdo con esta postura son McEnery y Wilson (1996:2) y Leech (Leech 1992:105), según el cual la lingüística de corpus es una plataforma sobre la que puede desarrollarse la investigación lingüística. Sin embargo hay que considerar, por otro lado, que la lingüística de corpus produce conocimiento nuevo y, por ende, podría considerarse disciplina.

Berber Sardinha (2000:355-356) participa en el debate sobre la definición de la lingüística de corpus y, para establecer su naturaleza, considera necesario definir lo que se entiende por metodología; sobre esta base, avanza la posibilidad de considerar la lingüística de corpus bien una metodología bien una herramienta. En opinión de este autor, si al término metodología se le otorga un sentido instrumental, entonces la lingüística de corpus puede ser considerada una herramienta; en cambio, si con el

término metodología se hace referencia a un modo típico de aplicar presupuestos teóricos, entonces sí que la lingüística de corpus puede ser considerada una metodología. Por lo demás, dado el carácter transdisciplinar de la lingüística de corpus, Hoey (1997 cit. Berber Sardinha 2000:357), sostiene que esta solo puede definirse como una perspectiva (ni como una metodología ni como una disciplina). Consideradas sus múltiples aplicaciones en diferentes ámbitos y la aceptación como método por parte de numerosos autores, en el presente estudio se considerará la lingüística de corpus como una metodología, en torno a la cual desarrollaremos nuestro análisis (véase capítulo 4).

La lingüística de corpus es, por tanto, una metodología empírica que se opone a la corriente racionalista sostenida por Chomsky y que se basa en el convencimiento de que la lengua es un sistema probabilístico en el cual, aunque sean posibles muchos rasgos, no todos ocurren con la misma frecuencia. Berber Sardinha (2000:350-352) explica esta aproximación, según la cual existen frecuencias diferentes, que no son aleatorias sino que dependen de patrones regulares. Por tanto, de acuerdo con los defensores de la lingüística de corpus, es posible conocer dicha frecuencia de aparición y ese conocimiento se puede alcanzar a través de la observación empírica. El mismo autor hace referencia a Leech (1992:107) para resumir las diferencias fundamentales que oponen la lingüística de corpus a la lingüística racionalista teorizada por Chomsky:

| Lingüística de corpus | Lingüística racionalista |
|-------------------------------|---|
| Enfoque: práctica lingüística | Enfoque: competencia |
| Descripción de la lingüística | Descripción de los universales lingüísticos |

| | |
|----------------------------------|---------------------------------|
| Modelo cuantitativo del lenguaje | Modelo cualitativo del lenguaje |
| Investigación empírica | Investigación racionalista |

Tabla 3: comparación lingüística de corpus y lingüística racionalista

En cuanto al encaje de la lingüística de corpus, en tanto que método, en el ámbito de la traductología, Brigitte Lépinette (2004:2-3) la relaciona con el tercero de los modelos investigadores que siguen:

- modelo sociológico-cultural, que considera el contexto social y cultural de un fenómeno en el momento de su producción y recepción;
- modelo histórico descriptivo, que incluye la categoría del modelo descriptivo-comparativo basado en el estudio de las diferentes teorías de traducción y en la evolución de los conceptos de dichas teorías en el tiempo;
- modelo descriptivo-contrastivo, basado en las elecciones tomadas por los traductores.

En concreto, la autora argumenta que la lingüística de corpus tiene un valor muy señalado para estudiar la relación entre el texto fuente y el texto meta y para describir las estrategias de traducción elegidas por los traductores.

Según Lépinette (2004:7), esta descripción puede tener un enfoque global, si se analiza todo el texto, o selectivo, si se analizan unos aspectos relevantes del texto. En el primer caso, el del enfoque global, se examina el texto en el plano macrotraductológico, es decir se estudian sus aspectos generales (la división en diferentes partes, la presencia o ausencia de grandes segmentos de textos, por ejemplo de párrafos o capítulos, etc.). En el segundo caso, en cambio, se examina el texto en el plano microtraductológico, es

decir se estudian sus aspectos más específicos (la organización discursiva, las estrategias de traducción utilizadas, etc.). En este segundo caso, el elemento objeto de análisis se escoge *a priori*, por ser relevante desde diferentes perspectivas:

- por ser la manifestación de un determinado fenómeno lingüístico;
- por ser la manifestación de un fenómeno ligado a una determinada época histórica;
- por ser un elemento ligado a una determinada historia literaria.

3.1.1.2 Objetivos y aplicaciones de investigación basada en corpus

Como hemos visto arriba, de acuerdo con Sinclair (1991:171), la lingüística de corpus ha nacido como respuesta a la necesidad de estudiar el lenguaje natural a través de ejemplos reales. No obstante, para establecer más específicamente los objetivos y las aplicaciones de la lingüística de corpus hay que considerar las disciplinas en las cuales se utilizan los corpus. Petra Procházková (2006:7-8) indica algunos de los objetivos que se intentan alcanzar a través de la metodología de corpus aplicada a diferentes materias lingüísticas:

- en dialectología/sociolingüística: distinción entre dialectos y sociolectos y entre estos y la lengua estándar;
- en lingüística histórica: estudio de la evolución de un palabra, de una construcción gramatical o de su utilización a lo largo de un período histórico;
- en psicolingüística: estudio de la influencia de los fenómenos lingüísticos en el procesamiento del lenguaje;

- en lexicografía: estudio de la frecuencia, de la presencia, de la utilización y de las colocaciones de las palabras, estudio de la presencia de más de un género para un mismo sustantivo, estudio de los términos utilizados en los diferentes registros;
- en sintaxis: estudio de las condiciones de existencia de determinadas construcciones sintácticas;
- en semántica: estudio del sentido y de la utilización metafórica de las palabras;
- en fonología: estudio, análisis y clasificación de la pronunciación y de los acentos de hablantes nativos y no nativos;
- en morfología: estudio de los afijos y de su productividad;
- en lingüística computacional: elaboración automática de un diccionarios, memorias de traducción y diccionarios de colocaciones;
- en didáctica: estudio de los errores más frecuentes en hablantes no nativos y análisis del proceso de adquisición de la lengua.

En traductología, Xiao y Yue (2009:237) utilizan las palabras de Toury (1995:265 cit. Xiao y Yue 2009:237) para presentar la investigación de corpus como una aproximación destinada a indagar en las regularidades y los comportamientos actuales de la lengua (traducida) y subrayan su utilidad para estudiar los procedimientos, los productos y las funciones de la traducción. A estos objetivos, Gandin (2009:134) añade otro fin general, el de identificar semejanzas y diferencias entre textos. Además según Gandin (2009:147) las aplicaciones traductológicas más útiles de la lingüística de corpus son:

- profundizar en el tema de los universales lingüísticos;

- crear corpus sobre el lenguaje especializado de grandes dimensiones;
- estudiar las lenguas no europeas sobre las cuales todavía no hay gran disponibilidad de corpus.

Finalmente la estudiosa (Gandin 2009:141 y 147) añade que, para las investigaciones concretamente desarrolladas con corpus paralelos, los objetivos más comunes suelen ser:

- identificar la terminología específica del lenguaje analizado;
- identificar las colocaciones;
- analizar estructuras sintácticas, gramaticales y estilísticas;
- remontar al significado conceptual de un término
- realizar glosarios bilingües;
- estudiar combinaciones sintácticas y semánticas del lenguaje analizado.

3.1.1.3 Ventajas y límites de los métodos basados en corpus

Como subraya Gandin (2009:137), la mayor ventaja de la utilización de los corpus lingüísticos es su carácter electrónico, característica que permite analizar muchos datos en poco tiempo, trabajar con material siempre actualizado y buscar la información de forma más rápida y sencilla. Esto permite estudiar una gran cantidad de textos auténticos, hecho que aumenta la posibilidad de alcanzar una mayor representatividad.

Entre los límites de esta metodología, Rojo (2002:3) presenta el problema de la recogida, selección y procesamiento del material. Estos procedimientos, realizados por

personas diferentes, no son seguidos por un único experto que puede controlar todo el proceso en su totalidad, al revés de lo que ocurre en la lingüística empírica tradicional y, por consiguiente, se pierde la visión de una perspectiva completa de la investigación. Otro límite de la lingüística de corpus, según el mismo autor, es la gran cantidad de datos que resultan del proceso de análisis, entre los cuales se encuentran los resultados de interés para la investigación que tienen que ser identificados y aislados. Malmkjaer (1998:534) añade dos problemas adicionales relativos a la utilización de corpus paralelos. En primer lugar, la falta de contexto lingüístico ofrecido por las herramientas de concordancia no facilita el análisis de rasgos del texto en su globalidad ni de fenómenos semánticos. En segundo lugar, los corpus paralelos suelen incluir solo una traducción por cada texto original, hecho que no permite estudiar las características generales de la traducción por falta de datos. La misma autora propone solucionar este último problema complementando el corpus principal y más amplio, compilado para estudiar las normas de traducción, con otro corpus de dimensiones más pequeñas formado por más traducciones del mismo texto original hechas por traductores diferentes. A esta propuesta de solución, Laviosa (2002:28) contesta con un problema añadido: la dificultad en cuanto al acceso de múltiples traducciones, que a menudo están disponibles solo para los textos literarios y también en ese caso pertenecen a épocas diferentes, con lo cual son difícilmente comparables. Finalmente, entre los límites de la lingüística de corpus se pueden indicar la restricción del número de elementos que se pueden analizar y el grado de representatividad que se atribuyen los corpus y que determina el alcance de los resultados.

3.1.2 Breve recorrido histórico de los estudios traductológicos basados en corpus

A finales de los años 50, se asiste a un cambio de paradigma lingüístico que lleva a la presencia de dos corrientes opuestas: por un lado, el empirismo de la lingüística de corpus recién nacida y, por otro, las teorías racionalistas del lenguaje propugnadas por Chomsky. De acuerdo con Rojo (2002:2), al comienzo, la lingüística de corpus se encuentra en una situación marginal con respecto al resto de enfoques de la época en boga en los Estados Unidos: el distribucionalismo ya en decadencia y la gramática generativa transformacional, en alza. Los corpus compilados en este período son manuales y con objetivos pedagógicos, al contrario de los de hoy en día, electrónicos y descriptivos. Entre los pioneros de esa nueva metodología López-Mezquita Molina (2007:125-126) recuerda a Thorndike que, en 1921, compila *The Teacher's Wordbook*, el primer corpus manual que incluye las palabras más frecuentes de la lengua inglesa; diez años más tarde este estudioso amplía su trabajo alcanzando las 20.000 palabras y, en 1944 junto a Lorge, llega a las 30.000. El segundo autor nombrado por la estudiosa es Michael West que, en 1953, compila el *General Service List (GSL)*, un corpus manual compuesto por la descripción de 20.000 palabras de lengua inglesa con fines didácticos. A este recorrido histórico, Berber Sardinha (2000:326) añade el corpus *Survey of English Usage (SEU)* de Randolph Quirk, compilado a partir de 1953 en Londres. Este corpus manual reúne un número fijo de textos (200 textos) cada uno con la misma cantidad de palabras (500 palabras por texto) analizadas gramaticalmente, hecho que lo convierte en la base de los etiquetadores gramaticales actuales.

Considerando las diferencias de los corpus manuales con los electrónicos, Bowker (2002:43) subraya la mayor limitación de aquellos: la gran cantidad de tiempo necesario

para compilarlos y, sobre todo, para analizarlos y consultarlos. De esa restricción proviene la crítica más frecuente que, por aquel entonces, hacen Chomsky y sus seguidores a esta metodología: la falta de confiabilidad del recuento manual en presencia de grandes cantidades de datos. Una respuesta y una solución a esta acusación llegan en los años 60, con los primeros ordenadores que permiten contar los datos de forma electrónica, con mayor sencillez, fiabilidad y rapidez. Así que, en 1964, nace el primer corpus electrónico, el *Brown Corpus (Brown University Standard Corpus of Present Day American English)*. Paralelamente, en cuanto a los estudios de traducción, Laviosa (2002:18-19) añade a los avances informáticos dos fenómenos que impulsaron el desarrollo y la utilización de los corpus electrónicos: el cambio de perspectiva desde los estudios prescriptivos a los descriptivos y la teoría de los polisistemas. El primer fenómeno fue propugnado por Mona Baker, a la que Laviosa (2002:18) considera la primera estudiosa en proponer la investigación de corpus para los estudios de traducción, con el objetivo de describir los universales de traducción. El segundo fenómeno es la creciente influencia de la teoría del polisistema en los estudios literarios y de traducción, teoría que nace al final de los años 70 con Even Zohar y sigue en las décadas sucesivas con Toury. Por lo demás, Xiao y Yue (2009:237-238) consideran que la metodología de corpus ha sufrido la influencia de otro cambio de perspectiva ocurrido en las décadas precedentes: el paso de la micro a la macroperspectiva, que empieza en los años 50 y sigue desarrollándose hasta los 80.

En los años 80, la lingüística de corpus sigue creciendo gracias a la invención de los micro-ordenadores personales. Nace en esta década el *Birmingham University International Language Database* de 1987, que consigue alcanzar los 20 millones de palabras. Una decena de años más tarde se compila el *British National Corpus (BNC)*,

que une la lengua inglesa escrita y hablada en un archivo de 100 millones de palabras. El BNC es el primer *megacorpus* de estas dimensiones y el único actualmente disponible a la venta (para más datos sobre los corpus, ver por ejemplo Sardinha Berber 2000:330-331). Los corpus fundamentales mencionados arriba, como aclara Bowker (2002:44), establecen factores ineludibles aun hoy día para los diversos ámbitos en los que se utilizan (ocupando en este sentido la traductología un lugar prominente): su amplitud (que los hace más representativos), su naturaleza electrónica (que facilita su manipulación y análisis) y, finalmente, su compilación (basada en criterios explícitos).

En la actualidad, proliferan los centros de investigación basada en corpus, sobre todo en Europa y en concreto en Inglaterra y en los países escandinavos. Fuera de Europa la presencia de estos centros de investigación sigue siendo modesta, como en los Estados Unidos, quizás por el fuerte protagonismo de la corriente de la lingüística generativa, que se apoya en la introspección y de fundamentos opuestos al empirismo de la lingüística de corpus. Además, a día de hoy, esta metodología se desarrolla también en ámbito empresarial, fenómeno gracias al cual muchos de los centros de investigación se encuentran financiados por empresas, sobre todo de las áreas informáticas y de las telecomunicaciones. Entre los principales estudios basados en corpus de nuestros días Berber Sardinha (2000:358) subraya los que indagan en la investigación de corpus, la descripción de las herramientas de investigación de corpus, la aplicación de los estudios de corpus y el objetivo más común: la descripción del lenguaje (original y traducido).

3.1.3 Clasificación de corpus lingüísticos

En este apartado se hará un repaso de algunas de las clasificaciones de corpus que se han propuesto en la disciplina con el fin de que sirvan de trasfondo para las tareas de

diseño y compilación de este estudio. Dado el elevado número de clasificaciones que existen y la multiplicidad de parámetros que estas contemplan, el apartado es algo prolijo. Sin embargo, se ha intentado contener al mínimo con criterios ilustrativos (es decir, hemos elegido clasificaciones que sirvan de ejemplo del tipo de estructuración propuesta) y se han organizado los enfoques de una forma lógica: desde aquellos procedentes del ámbito de la lingüística de corpus en general hasta los que se defienden desde la traductología basada en corpus. El apartado se concluye con dos clasificaciones que tienen en cuenta la codificación de los textos del corpus y se cierra con un trabajo completo y equilibrado, propuesto desde la Universitat Jaume I, marco en el que se realiza este estudio.

Desde la lingüística de corpus en general, los estudiosos se han esforzado por organizar la desbordante proliferación de compilaciones ocurrida en el ámbito. Ello ha dado como resultado un crecimiento paralelo de clasificaciones poliédricas (con niveles variados y hasta complejos). El ejemplo más típico de esto lo constituye Berber Sardinha (2000:340-341) cuyo ordenamiento se basa en numerosos parámetros diferentes, algunos de los cuales dan lugar a las clasificaciones también propuestas por otros autores y otros a clasificaciones completamente originales basadas en:

- modalidad: oral
- escrito
- tiempo: sincrónico, con textos de un período determinado
- diacrónico, que incluye textos de varios períodos

- contemporáneo, con textos escritos en el presente
- histórico, con textos escritos en épocas pasadas
- selección:
 - de muestra
 - monitor, que representa el estado actual de la lengua o de la variedad analizada
 - dinámico/abierto
 - equilibrado, con cantidades semejantes distribuidas equilibradamente
- contenido:
 - especializado
 - regional/dialectal
 - múltiple, con textos de más idiomas
- autoría:
 - de aprendiz, con textos de autores o hablantes no nativos
 - de lengua nativa
- finalidad:
 - de estudio, descriptivo
 - de referencia, contrastivo
 - de entrenamiento o test, para desarrollar aplicaciones
- disposición interna:
 - paralelo
 - alineado

El mismo autor añade otros parámetros clasificatorios, entre los que destacan:

- pluralidad de autoría;
- origen de autoría;
- medio;
- integralidad;
- especificidad;
- dialecto;
- equilibrio;
- fecha;
- renovación;
- temporalidad;
- plurilingüismo.

Algunos estudiosos de la lingüística de corpus intentan contener el afán subdivisorio y proponen clasificaciones algo más sencillas. Gandin (2009:134-135) es una de ellos. La autora propone una división basada principalmente en la disposición interna de los textos, y distingue entre:

- corpus paralelo, es decir un corpus formado por textos originales y sus respectivas traducciones;
- corpus comparable, es decir un corpus formados por textos solo originales o solo traducidos.

En una segunda capa de estructuración, la misma autora propone subclasificaciones de estos dos grupos. Así, de un lado, los corpus paralelos pueden seguir diferentes modelos:

- unidireccional, si el corpus incluye solo textos traducidos de la misma lengua de partida a la misma lengua de llegada (por ejemplo solo textos traducidos del inglés al español);

- bidireccional, si el corpus incluye textos traducidos de la lengua de partida a la lengua de llegada y viceversa (por ejemplo con textos traducidos del inglés al español y del español al inglés);

- modelo estrella, si el corpus incluye textos traducidos de una misma lengua de partida a diferentes lenguas de llegada (por ejemplo con textos traducidos del inglés al español, del inglés al francés y del inglés al italiano);

- modelo diamante, si el corpus incluye todas las combinaciones posibles de traducción entre tres o más lenguas (por ejemplo con textos traducidos del inglés al español, del inglés al italiano, del español al inglés, del español al italiano, del italiano al inglés y del italiano al español).

De otro lado, los corpus comparables, según Gandin (2009:135), serían:

- corpus monolingüe;

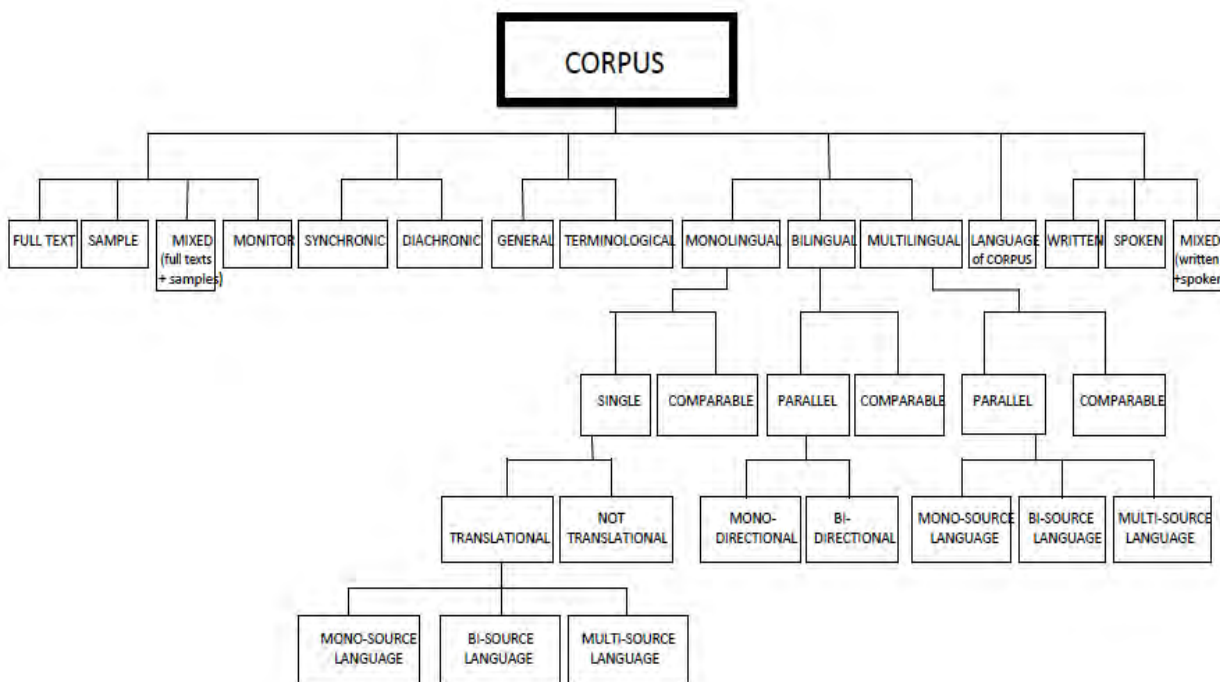
- corpus bilingüe.

Una tercera forma de estructurar los corpus (desde la lingüística), lo propone Procházková (2006:5), quien los organiza a partir de una lógica de contraste; es decir,

aquí no se presentan clasificaciones según (múltiples) parámetros sino que se los agrupa, principalmente (aunque no solo), por características opuestas. Así la autora clasifica los corpus en:

- bases de datos de árboles;
- orales ≠ de textos;
- multimodales;
- sincrónico ≠ diacrónico;
- monolingüe ≠ multilingüe;
- de referencia ≠ monitor;
- histórico;
- dialectal.

La lógica del contraste nos sirve de transición para revisar propuestas desde el ámbito de la traductología de corpus. Aquí dos de las clasificaciones más canónicas son las de Laviosa (2002) y Bowker (2002), publicadas curiosamente, en el mismo año. Laviosa (2002:34-38) sigue un método equiparable, por su profusión y detalle, al de Berber Sardinha (2000). En su obra *Corpus-Based Translation Studies: Theory, Findings, Applications* (2002) organiza los corpus jerárquicamente en cuatro niveles, cada uno con divisiones y subgrupos. Esta organización jerárquica se presenta en el organigrama a continuación:



Además la misma autora distingue, según el objetivo de la investigación de corpus, entre los corpus utilizados en estudios que:

- consideran la traducción como producto;
- consideran la traducción como proceso;
- sirven de entrenamiento para los traductores;
- se basan en la lingüística contrastiva.

Bowker (2002:45-46), sin embargo, parece favorecer un método estructurador algo más sencillo, afín, por ello, a las formulaciones que, desde la lingüística de corpus, se ha visto que hace Gandin (2009) y clasifica los corpus según el número de lenguas (monolingüe, bilingüe y multilingüe), el nivel de especialización (generales y especializados) y la relación que se crea entre los textos (paralelos, si se unen textos

originales con sus respectivas traducciones; comparables, si se unen textos con características comunes).

Esta clasificación de Bowker (2002) ha inspirado, recientemente, para sus planteamientos de recapitulación a Xiao y Yue (2009:240-242) quienes construyen sobre el trabajo de Bowker (2002), añadiendo una clasificación de los corpus paralelos entre:

- corpus unidireccional (de una lengua de partida a una lengua de llegada);
- corpus bidireccional (de una lengua de partida a una lengua de llegada y viceversa);
- corpus multidireccional (de una lengua de partida a diferentes lenguas de llegada).

Asimismo, completan el trabajo de Bowker (2002) con una novedosa distinción entre corpus:

- comparables, es decir formados por textos en lenguas diferentes pero con características comunes y comparables;
- comparativos, es decir con el objetivo de comparar las variantes lingüísticas (sincrónica o diacrónicamente).

Finalmente, existen estudiosos que, tienen en cuenta, para sus clasificaciones, factores de anotación. Entre ellos Alcina Caudet y Estellés Palanca (2006:1415), que diferencian los corpus dependiendo del procesamiento en el momento de su compilación entre:

- corpus crudo, sin elementos etiquetados;
- corpus etiquetado, si cada elemento contiene informaciones comunicativas.

Mercado Percia (2008:13-17) sigue la línea exhaustiva de Berber Sandinha (2000) y Laviosa (2002) hasta tal punto que, junto con la consideración de numerosos parámetros, pone en un lugar destacable el criterio de anotación y etiquetado. En ese sentido, los corpus pueden ser:

- simple;

-codificado y anotado (ASCII).

Esta subdivisión complementa a otras basadas en el porcentaje de distribución de textos (corpus grande; equilibrado; piramidal; comparable; multilingüe; oportunista), la especificidad de los textos (general; especializado; genérico; canónico; periódico/cronológico; diacrónico) la cantidad de los textos incluidos en el corpus (textuales; de referencia; léxicos) y contenido (orales; de texto; multimodales).

Como broche final, se recoge aquí la categorización de Isabel García Izquierdo (2005:4) que acepta la división precedentemente propuesta por G. Corpas (2001:157-159) y Baker (1995) y junto con Esther Monzó (2003) presenta la división siguiente:

- Según el porcentaje y la distribución de los diferentes tipos de texto:

corpus grande o extenso, corpus equilibrado, corpus piramidal, corpus monitor, corpus paralelo y corpus comparable.

- Según la especificidad de los documentos contenidos en el corpus:

corpus general, corpus

especializado, corpus canónico, corpus periódico o cronológico y

corpus diacrónico.

- Según la cantidad de texto que se reúne en cada uno de los documentos que componen el corpus: corpus textual, corpus de referencia y corpus léxico.
- Según la codificación y anotación: corpus no anotado, corpus anotado.
- Según la documentación que acompaña al corpus: Corpus documentado y corpus no documentado.

3.1.4 Criterios de compilación de corpus lingüísticos

Para la compilación de un corpus es necesario seguir unos prerequisites y escoger los textos a partir de parámetros establecidos previamente. De nuevo la bibliografía en este sentido es prolija y procede de terrenos distintos (desde la lingüística de corpus hasta la traductología, entre otros). Aquí, en un esfuerzo de condensación ilustrativo se recogen aquellas propuestas que parecen más útiles para el trabajo posterior.

Cabe en primer lugar destacar que numerosos autores (como Gandin 2009:143-146) no entienden el establecimiento de criterios de forma aislada sino como parte de un proceso más amplio, que atañe tres fases:

- el diseño del corpus;
- la codificación de los textos;
- los objetivos del proyecto de investigación.

Gandin (2009: 143-146) entra de lleno en cada una de las tres fases y su explicación de los prerequisites es exhaustiva, quizá demasiado extensa y algo dispersa. En la primera fase, según la autora, hay que determinar:

- área temática;
- elección del corpus: cerrado (si se entiende estudiar la evolución de la lengua o de un fenómeno) / abierto (si se entiende analizar la lengua o un determinado fenómeno en un período específico);
- selección de los textos: enteros (si el objeto de la investigación es el lenguaje específico) / muestras de textos (si el objetivo de investigación es el lenguaje general);
- cantidad de textos;
- calidad de los textos;
- autoría de los textos
- fecha de publicación de los textos;
- disponibilidad de los textos y de su permiso de utilización.

La segunda fase es la de codificación, por la que la autora (Gandin 2009:145) entiende el proceso de preparación de los textos para su inserción en el corpus. Para compilar un corpus, la característica fundamental es que todos los archivos estén en formato electrónico. Para alcanzar esta prerrogativa, o bien los textos se recogerán ya en formato digital o bien habrán de ser escaneados con el método OCR. Después de tener todos los textos en formato electrónico, se pasa a la fase de *mark-up*, es decir de etiquetado y de codificación en lenguaje informático, que para los corpus suele ser el *Corpus Encoding*

Standard (CES). Finalmente se llega a la fase de la alineación para enlazar los textos originales con sus respectivas traducciones.

La tercera fase es la determinación de los objetivos del proyecto investigador basado en corpus. Aquí Gandin (2009:146) sostiene que, tanto para investigaciones con fines didácticos como para aquellas con fines de investigación, el objetivo principal suele ser el estudio de fenómenos típicos de un género lingüístico, de un lenguaje especializado, de un autor o de un personaje determinados a través del análisis electrónico del corpus en cuestión que, como se presentará más detalladamente en el apartado siguiente, permite obtener datos sobre:

- el número de *tokens*, el número de las palabras totales presentes en el corpus;
- el número de *types*, el número efectivo de las palabras diferentes utilizadas en el corpus, que permite establecer la variabilidad lingüística;
- la lista de palabras;
- la lista de frecuencia;
- la lista de colocaciones.

Como Gandin (2009), también Berber Sardinha (2000) es meticuloso en la descripción de los criterios que han de cuidar los investigadores que trabajen con corpus. Sin embargo, en vez de cubrir todo el proceso, Berber Sardinha (2000:338-339) focaliza su atención exclusivamente sobre cuatro prerequisites:

- autenticidad, ya que los textos tienen que estar escritos en lengua natural;

- naturalidad, ya que quienes escriben los textos tienen que ser hablantes nativos (si se utilizan textos de hablantes no nativos hay que indicarlo explícitamente hablando de corpus de aprendiz);
- criterios de elección de los textos según sus características;
- representatividad de un idioma o de una variedad de lengua.

Partiendo de estos cuatro prerrequisitos, el mismo autor (Berber Sardinha 2000:339-349) explica cada uno de los elementos que es necesario considerar en la compilación de un corpus:

- tipología, que ya se ha explicado en el apartado anterior del presente estudio, dedicado a la clasificación de los corpus;
- representatividad, ya que el corpus posee una función representativa de la lengua, de la variante o del fenómeno que se pretende estudiar. La representatividad de un corpus depende de sus dimensiones. Cuanto más amplio sea el corpus, más posibilidades tendrá este de representar todos los casos de una lengua (o parte de esta). La dimensión de un corpus puede calcularse con respecto al número de palabras, de textos o de géneros que lo componen. Sin embargo, la representatividad nunca podrá ser absoluta por tres cuestiones: en primer lugar la lengua es un sistema probabilístico difícil de prever; en segundo lugar el corpus es una muestra de una población de tamaño desconocido porque no se sabe realmente cuántos son los hablantes de una lengua; y finalmente porque la representatividad depende también de los usuarios ya que un corpus puede ser representativo para algunos individuos pero no para otros;

- extensión, de la cual, como se acaba de explicar, depende la representatividad del corpus. Para elegir la extensión de un corpus hay que considerar tres aproximaciones: impresionista, histórica y estadística. La aproximación impresionista se basa en las teorías de autores reputados en el área que indican la cantidad de palabras necesarias para que un corpus se pueda considerar representativo. La aproximación histórica se basa en la observación de los corpus utilizados en la comunidad y cuya extensión se tomará de ejemplo. La aproximación estadística se basa en la aplicación de teorías estadísticas que intentan identificar el tamaño del corpus necesario para determinadas investigaciones a partir de vertientes internas, externas o relativas;

- especificidad del corpus, es decir delimitar al máximo las variedades incluidas en el corpus;

- adecuación del corpus a los objetivos de investigación, es decir que el corpus tiene que compilarse para alcanzar objetivos determinados.

Si Berber Sardinha (2000) concede un lugar destacado a la representatividad, Procházková (2006:6) hace de este su único foco de interés y se esmera en identificar los parámetros que aseguren este criterio:

- modalidad: lengua oral o escrita;

- diferentes registros especiales (medicina, finanza, economía, etc.);

- diversos parámetros demográficos (edad, grupo social, género y religión);

- clasificación por época o por siglo;

- medios de comunicación (libros, periódicos, correo, etc.);

- niveles lingüísticos (coloquial, formal, familiar, etc.);
- tipos de textos (novela, poema, formulario, etc.).

La traductología también ha dado enumeraciones claras y pertinentes de los criterios para la selección y compilación de los corpus. Siguiendo la estela de los estudiosos de la lingüística de corpus, Tognini-Bonelli (2001: 54-64) postula que han de tenerse en cuenta estos tres factores:

- autenticidad;
- representatividad (que se alcanza según la elección de un corpus variado y complejo formado por textos de diferentes tipos, frecuencia de aparición, factores de aparición y variación);
- selección de textos.

En una línea algo más exhaustiva, encontramos el trabajo de Bowker y Pearson (2002:45), quienes enumeran los parámetros necesarios para compilar un corpus:

- el tamaño del corpus su naturaleza finita o abierta, su dinamicidad;
- elección entre textos enteros o extractos;
- número de textos;
- medio (textos escritos, orales o de ambos tipos);
- tema de los textos;
- tipología de textos;
- autoría de los textos;

- lengua de los textos;
- fecha de publicación de los textos.

3.1.5 Herramientas de procesamiento y análisis

Las herramientas de procesamiento son esos instrumentos que preparan el corpus para el análisis, según los objetivos de investigación y el enfoque adoptado durante el estudio.

Forman parte de las herramientas de procesamiento los procesos de:

- lematización: la unión de las diferentes formas de la misma palabra bajo el mismo lema;
- anotación o etiquetado: el etiquetado de determinados elementos de los textos según los intereses de investigación (elementos pragmáticos, semánticos, léxicos, sintácticos, etc.);
- marcación: la introducción en el corpus de informaciones lingüísticas;
- alineación: la relación de segmentos del texto en lengua original con los correspondientes segmentos de los textos en lengua meta.

Una vez preparado el texto (es decir una vez compilado y procesado) se puede pasar al análisis, o sea a la identificación de los elementos útiles para alcanzar los objetivos de investigación. Bowker (2002:46-68) y Bowker y Pearson (2002:109-130) presentan las herramientas de análisis más comunes.

La más utilizada es la lista de todas las palabras que aparecen en el corpus. Esta herramienta puede presentar los resultados en tres diferentes formatos:

- análisis estadístico (con las informaciones sobre el número de *tokens*, *types*, la relación entre ellos, la extensión del corpus total y de cada texto);
- listado de frecuencias (con todas las palabras que aparecen en el corpus organizadas en orden de frecuencia; es esta una herramienta útil para saber, por ejemplo, el tema principal del texto analizado);
- listado de palabras en orden alfabético (útil sobre todo para identificar las diferentes formas del mismo lema).

Como se acaba de ver, la lista de palabras divide el número de *tokens* (el número total de las palabras de un corpus) del número de *types* (el número de palabras diferentes utilizadas en el corpus), permitiendo así calcular la variedad léxica. Además, gracias a la proporción entre palabras léxicas y gramaticales, se puede identificar la densidad léxica y la carga informativa del texto en análisis. Otras informaciones estadísticas proporcionadas por esta herramienta son la longitud media de las palabras, de las oraciones y de los párrafos.

La segunda herramienta que presentan Bowker (2002:46-68) y Bowker y Pearson (2002:109-130) es la *stop list*, un listado de palabras que se desea ignorar durante el análisis. Puede constar de palabras o clases gramaticales.

La tercera herramienta de la que hablan Bowker (2002:46-68) y Bowker y Pearson (2002:109-130) es el listado de concordancias. Este listado se utiliza para buscar todas las ocurrencias de un patrón en su cotexto y contexto y puede aplicarse a los corpus monolingües como a los bilingües. En el caso de los corpus monolingües, el listado de concordancias puede presentarse de dos formas: como una búsqueda en el texto completo, evidenciando las palabras en el interior del texto o como una búsqueda de

índice, es decir como un índice de todas las palabras que muestra el número de líneas en las cuales se encuentran estas en el interior del texto, para que sea posible analizarlas en su contexto.

Estos listados permiten también encontrar las concordancias de una forma lematizada y realizar una *context search*, es decir buscar dos términos separados entre ellos por un número de palabras determinado especificado previamente con un número fijo o un intervalo. Los resultados de la búsqueda de concordancias en corpus monolingües se presentan en formato KWIC, es decir la frase en la cual se encuentra la concordancia no aparece de forma completa y la palabra buscada, que Sinclair (2003:xiii) define como el nodo de la búsqueda, aparece en el medio de la secuencia de palabras. El listado de concordancias aplicado a los corpus bilingües funciona de la misma forma que el listado de concordancias aplicado a corpus monolingüe, pero muestra los resultados de forma parecida a la de la alineación, organizándose, a menudo, las colocaciones en lengua original a la izquierda (o arriba) y las en lengua meta a la derecha (o debajo). Este listado permite además observar las colocaciones en lengua original de forma independiente de las de lengua meta (por ejemplo disponiendo las primeras en orden alfabético y las segundas en orden de frecuencia) y se puede escoger si buscar solo las concordancias de palabras con una determinada traducción (por ejemplo la herramienta puede identificar solo las concordancias de las palabras *drive* traducidas con *conducir*). Asimismo, se pueden buscar las concordancias de palabras con traducciones diferentes a una determinada opción (por ejemplo la herramienta puede identificar concordancias de la palabra *drive* que no se hayan traducido como *conducir*). Esto resulta muy útil para calibrar el uso de una traducción en concreto.

Bowker (2002:64-68) destaca, asimismo, los buscadores de grupos y colocaciones, es decir las coocurrencias, o sea palabras que se encuentran juntas con una probabilidad superior a la casualidad. Para encontrar las colocaciones, suele preferirse la fórmula de información mutua (MI), que confronta la probabilidad de aparición conjunta de los componentes de la posible colocación frente a la aparición de estos, por separado, en otros contextos. Una vez identificadas las colocaciones estas pueden ordenarse en orden alfabético, de frecuencia o de posición.

Finalmente, a todas estas herramientas, Bowker y Pearson (2002 109-119) añaden la lista de palabras claves, que permite encontrar las palabras más frecuentes de un texto en comparación con otro (por ejemplo comparando un texto determinado con otros del mismo corpus; o comparando las palabras frecuentes de un corpus frente a las de otro corpus más amplio, general y representativo).

3.2 Metodología del presente estudio

Se explican a continuación aspectos reseñables de la metodología que se utilizará en el presente trabajo. En concreto, examinaremos debajo hipótesis y objetivos (3.2.1.), el archivo de corpus empleado (3.2.2.), los criterios de compilación (3.2.3.), el proceso de preparación de los corpus (3.2.4) y el método de análisis (3.2.5).

Conviene dejar claro, desde el primer momento, que este estudio es un primer paso para un trabajo más complejo, que se presentará en el marco de la tesis doctoral. Por ello, esta metodología es provisional y se propone a modo de prueba para examinar, de un lado, la operatividad de marco teórico sobre culturemas y técnicas traductoras y, de otro, la fortaleza y debilidad de las decisiones investigadoras que se han tomado en este

trabajo. En definitiva, esta metodología se plantea como un primer acercamiento que permita reforzar la calidad de trabajos futuros (entre ellos, la tesis doctoral).

3.2.1 Hipótesis y objetivos del estudio

La hipótesis principal del presente estudio es la siguiente:

Una metodología basada en corpus facilita el proceso de identificación de áreas proclives al uso de culturemas y propicia el análisis cuantitativo y cualitativo del tratamiento de los culturemas identificados.

Con el fin de corroborar o refutar esta hipótesis, el objetivo principal de este estudio es identificar, clasificar y comparar culturemas de literatura traducida frente a culturemas de literatura original para así identificar, clasificar y calibrar tratamientos traductores. Como parece claro, este objetivo puede desglosarse en los propósitos parciales que aparecen a continuación:

- diseñar un método para identificar los culturemas presentes en el corpus objeto de estudio;
- clasificar los culturemas presentes en el corpus objeto de estudio y probar, así, la operatividad de las clasificaciones que se han examinado en el capítulo 1;
- analizar/comparar culturemas de literatura traducida con culturemas de literatura original;
- diseñar un método para identificar las técnicas de traducción utilizadas en el corpus objeto de estudio;

- clasificar las técnicas de traducción utilizadas en el corpus objeto de estudio este género, probar, así, la operatividad de las clasificaciones que se han examinado en el capítulo 2 y calibrar las más productivas en nuestro corpus;
- obtener un conocimiento más detallado del corpus de literatura traducida objeto de estudio (para más datos del corpus, véase apartado 3.2.2), en el que podrá profundizarse en trabajos futuros.

Confrontando el objetivo principal de este trabajo (y su desglose en propósitos parciales) con los objetivos planteados por los estudiosos del área (véase apartado 3.1.1.2.) cabe subrayar que nuestras metas:

- refuerzan el afán de Sinclair (1991) por estudiar la lengua natural, ya que auténticos son los textos literarios (originales y traducidos) que se han elegido para su estudio;
- coinciden con los postulados traductológicos de Xiao y Yue (2009) en cuanto a la indagación de regularidades y comportamientos traductores. Siguiendo con Xiao y Yue (2009), el presente trabajo se marca objetivos vinculados con el estudio de procedimientos traductores a partir del examen del producto de trasvase;
- se encuentran en la línea de Gandin (2009) cuando ponen el foco en las similitudes y diferencias de originales y traducciones. Si bien este trabajo, así planteado, no aspira a participar en el debate de universales de traducción, puede servir como un primer paso muy incipiente en este sentido. Más bien, y de acuerdo con Gandin (2009), la propuesta planteada se centra en el análisis de estructuras lingüísticas que, en nuestro caso, encapsulan culturemas.

Por último, resulta necesario destacar en un lugar prominente que los objetivos plateados arriba vienen acompañados por el propósito investigador de aplicar la metodología provisional aquí propuesta de forma que compruebe su operatividad con vistas a la tesis doctoral.

3.2.2 Presentación de LIT_ENIT_ES: el archivo de los corpus utilizados para el estudio

El archivo de corpus objeto de estudio, LIT_ENIT_ES, está compuesto por 20 obras en total (10 textos fuente y 10 textos meta), de originales escritos entre 1930 y 2000.

Nótese que el total de *tokens* del archivo es 1.670.436 y el total de *types* es 105.322:

| Obra original / Traducción | <i>Tokens</i> | <i>Types</i> |
|--|----------------------|---------------------|
| <i>Mice and Men</i> | 30.297 | 2.810 |
| <i>The Catcher in the Rye</i> | 77.518 | 3.986 |
| <i>Lolita</i> | 113.360 | 13.844 |
| <i>The Joy Luck Club</i> | 94.526 | 8.738 |
| <i>Trainspotting</i> | 107.157 | 10.509 |
| <i>Il Gattopardo</i> | 74.576 | 13.927 |
| <i>Todo Modo</i> | 32.847 | 6.245 |
| <i>Se una notte d'inverno un viaggiatore</i> | 76.375 | 12.089 |
| <i>Il nome della rosa</i> | 180.769 | 21.754 |
| <i>Sostiene Pereira</i> | 42.757 | 5.454 |

| | | |
|--|------------------|----------------|
| <i>De ratones y ombre</i> | 30.229 | 4.659 |
| <i>El guardián entre el centeno</i> | 66.202 | 6.791 |
| <i>Lolita</i> | 111.163 | 17.242 |
| <i>El club de la buena estrella</i> | 96.483 | 12.207 |
| <i>Trainspotting</i> | 109.146 | 13.818 |
| <i>El gatopardo</i> | 78.142 | 13.034 |
| <i>Todo Modo</i> | 35.443 | 6.340 |
| <i>Si una noche de invierno un viajero</i> | 79.261 | 11.834 |
| <i>El nombre de la rosa</i> | 188.735 | 20.171 |
| <i>Sostiene Pereira</i> | 45.450 | 5.851 |
| Total | 1.670.436 | 21.2089 |

Tabla 4: presentación corpus LIT_ENIT_ES

Según las clasificaciones presentadas en el apartado 3.1.3 del presente capítulo, en general, este archivo está formado por textos enteros; se compone de texto escrito — Laviosa (2002:34-38), Mercado Percia (2008:13-17), Procházková (2006:5) —; es equilibrado por tener el mismo número de textos con las mismas características (en este caso el mismo número de textos traducidos de cada uno de los idiomas analizados); presenta una pluralidad de autoría y se ha diseñado con una finalidad descriptiva o de estudio (Berber Sardinha, 2000:340-341). En cuanto a su formato, y por la ausencia de etiquetas, es un corpus crudo (Alcina Caudet y Estellés Palanca, 2006:14-15), o simple

(Mercado Percia, 2008:13-17) y, por consiguiente, no está anotado (García Izquierdo, 2005:4).

Como vemos, el archivo global puede estructurarse a modo de complejo de corpus paralelos, uno formado por las 5 obras inglesas con su relativas traducciones en español (836.081 *tokens*); y otro formado por las 5 obras en italiano con las relativas traducción en español (834.355 *tokens*).

| Obra original | <i>Tokens</i> | <i>Types</i> | | Obra traducida | <i>Tokens</i> | <i>Types</i> |
|-------------------------------|----------------------|---------------------|---|-------------------------------------|----------------------|---------------------|
| <i>Mice and Men</i> | 30.297 | 2.810 | > | <i>De ratones y hombres</i> | 30.229 | 4.659 |
| <i>The catcher in the rye</i> | 77.518 | 3.986 | > | <i>El guardián entre el centeno</i> | 66.202 | 6.791 |
| <i>Lolita</i> | 113.360 | 13.844 | > | <i>Lolita</i> | 111.163 | 17.242 |
| <i>The Joy Luck Club</i> | 94.526 | 8.738 | > | <i>El club de la buena estrella</i> | 96.483 | 12.207 |
| <i>Trainspotting</i> | 107.157 | 1.009 | > | <i>Trainspotting</i> | 109.146 | 13.818 |
| Total | 422.858 | 39.887 | | | 413.223 | 54.717 |

Tabla 5: correspondencias corpus paralelo inglés-español

| Obra original | <i>Tokens</i> | <i>Types</i> | | Obra traducida | <i>Tokens</i> | <i>Types</i> |
|----------------------|----------------------|---------------------|---|-----------------------|----------------------|---------------------|
| <i>Il Gattopardo</i> | 74.576 | 13.927 | > | <i>El gatopardo</i> | 78.142 | 13.034 |

| | | | | | | |
|--|----------------|---------------|---|--|----------------|---------------|
| <i>Todo Modo</i> | 32.847 | 6.245 | > | <i>Todo Modo</i> | 35.443 | 6.340 |
| <i>Se una notte d'inverno un viaggiatore</i> | 76.375 | 12.089 | > | <i>Si una noche de invierno un viajero</i> | 79.261 | 11.834 |
| <i>Il nome della rosa</i> | 180.769 | 21.754 | > | <i>El nombre de la rosa</i> | 188.735 | 20.171 |
| <i>Sostiene Pereira</i> | 42.757 | 5.454 | > | <i>Sostiene Pereira</i> | 45.450 | 5.851 |
| Total | 407.324 | 59.469 | | | 427.031 | 57.230 |

Tabla 6: correspondencias corpus paralelo italiano-español

En este sentido y aludiendo a las clasificaciones del apartado 3.1.3 del presente capítulo, el corpus LIT_ENIT_ES es un corpus unidireccional, ya que todos los textos se vierten al español (Gandin, 2009 y Xiao y Yue, 2009). El complejo paralelo es multilingüe, por la presencia de más de dos idiomas, y de proveniencia de dos lenguas originales (*bi-source language*) (Laviosa, 2002).

Al mismo tiempo, el archivo global también puede estructurarse a modo de corpus monolingüe en lengua traducida, LIT_ES, compuesto por 10 obras vertidas al español, procedentes tanto del inglés como del italiano. Este corpus contiene 840.254 *tokens* y 52.664 *types* en total, de los cuales 413.223 *tokens* y 33.002 *types* son de textos traducidos del inglés y 427.031 *tokens* y 35.667 *types* son de textos traducidos del italiano. En la tabla que sigue a continuación complementaremos la información aportada hasta el momento, centrándonos ahora en el corpus LIT_ES.

| Título de la obra traducida | Autor | Año de publicación del original | Lengua original de la obra | Traductor | Tokens | Types |
|--|------------------------------|--|-----------------------------------|--------------------|---------------|--------------|
| <i>De ratones y hombres</i> | John Steinbeck | 1937 | Inglés | Román A. Jimenez | 30.229 | 4.659 |
| <i>El guardián entre el centeno</i> | J.D. Salinger | 1951 | Inglés | Carmen Criado | 66.202 | 6.791 |
| <i>El gatopardo</i> | Giuseppe Tomasi de Lampedusa | 1955 | Italiano | Fernando Gutiérrez | 78.142 | 13.034 |
| <i>Lolita</i> | Vladimir Nabokov | 1955 | Inglés | Enrique Tejedor | 111.163 | 17.242 |
| <i>Todo Modo</i> | Leonardo Sciascia | 1974 | Italiano | Joaquín Jordá | 35.443 | 6.340 |
| <i>Si una noche de invierno un viajero</i> | Italo Calvino | 1980 | Italiano | Esther Benítez | 79.261 | 11.834 |
| <i>El nombre de la rosa</i> | Umberto Eco | 1980 | Italiano | Ricardo Potchar | 188.735 | 20.171 |
| <i>El club de la buena estrella</i> | Amy Tan | 1989 | Inglés | Jordi Fibla | 96.483 | 12.207 |

| | | | | | | |
|--------------------------------|---------------------|------|----------|---|----------------|----------------|
| <i>Trainspotting</i> (1993) | Irvine Welsh | 1993 | Inglés | Federico Corriente | 109.146 | 13.818 |
| <i>Sostiene Pereira</i> | Antonio Tabucchi | 1994 | Italiano | Carlos Gumpert y Xavier González Rovira | 45.450 | 5.851 |
| Total | | | | | 840.254 | 111.947 |

Tabla 7: presentación subcorpus LIT_ES

Para los fines de este trabajo, el corpus comparable de literatura traducida LIT_ES se puede dividir en dos subcorpus: uno compuesto por las obras traducidas del inglés, denominado LIT_EN_ES y formado por 413.223 *tokens* y 33.002 *types*, y otro compuesto por las obras traducidas del italiano, denominado LIT_IT_ES que incluye 427.031 *tokens* y 35.667 *types*. Los dos subcorpus se presentan a continuación:

LIT_EN_ES

| Título de la obra traducida | Autor | Año de publicación del original | Traductor | <i>Tokens</i> | <i>Types</i> |
|-------------------------------------|-------------------|--|---------------------|---------------|--------------|
| <i>De ratones y ombre</i> | John Steinbeck | 1937 | Román A. Jimenez | 30.229 | 4.659 |
| <i>El guardián entre el centeno</i> | J.D. Salinger | 1951 | Carmen Criado | 66.202 | 6.791 |
| <i>Lolita</i> | Vladimir | 1955 | Enrique | 111.163 | 17.242 |

| | | | | | |
|-------------------------------------|-----------------|------|-----------------------|----------------|---------------|
| | Nabokov | | Tejedor | | |
| <i>El club de la buena estrella</i> | Amy Tan | 1989 | Jordi Fibla | 96.483 | 12.207 |
| <i>Trainspotting</i> (1993) | Irvine Welsh | 1993 | Federico Corriente | 109.146 | 13.818 |
| Total | | | | 413.223 | 54.717 |

Tabla 8: presentación corpus LIT_EN_ES

LIT_IT_ES

| Título de la obra traducida | Autor | Año de publicación del original | Traductor | Tokens | Types |
|--|------------------------------|--|--------------------|---------------|--------------|
| <i>El gatopardo</i> | Giuseppe Tomasi de Lampedusa | 1955 | Fernando Gutiérrez | 78.142 | 13.034 |
| <i>Todo Modo</i> | Leonardo Sciascia | 1974 | Joaquín Jordá | 35.443 | 6.340 |
| <i>Si una noche de invierno un viajero</i> | Italo Calvino | 1980 | Esther Benítez | 79.261 | 11.834 |
| <i>El nombre de la rosa</i> | Umberto Eco | 1980 | Ricardo Potchar | 188.735 | 20.171 |
| <i>Sostiene Pereira</i> | Antonio Tabucchi | 1994 | Carlos Gumpert y | 45.450 | 5.851 |

| | | | | | |
|--------------|--|--|------------------------------|----------------|---------------|
| | | | Xavier González Rovira | | |
| Total | | | | 427.031 | 57.230 |

Tabla 9: presentación subcorpus LIT_IT_ES

3.2.3 Criterios de compilación

Se presentan a continuación los criterios más destacables que se han seguido para compilar el archivo de corpus (LIT_ENIT_ES) que será objeto de nuestro estudio. De entre los diversos criterios que barajan los estudiosos del ámbito para el diseño de corpus (véase apartado 3.1.4. del presente capítulo), nos ha parecido oportuno dedicar especial atención al comentario de los siguientes:

- Representatividad: el archivo LIT_ENIT_ES no aspira, por el momento, a cumplir el criterio de representatividad. Compilar un corpus plenamente representativo de novela del siglo XX traducida al español desde el italiano y el inglés entraña problemas que van más allá de los compromisos que se asumen en un TFM (ya habíamos reseñado en el apartado 3.1.4. que para muchos autores la representatividad es imposible). No obstante, para este estudio, no hemos elegido cualquier novela sino que nos hemos preocupado por seleccionar textos que se han considerado pilares del género. En cualquier caso, y dadas las limitaciones de representatividad del corpus, conviene destacar que los resultados que se obtengan al final del estudio solo serán aplicables al corpus elegido y constituirán un primer paso para probar la teoría y metodología elegidas con vistas al futuro. En trabajos posteriores se aumentará el tamaño del corpus y se perseguirá dar una imagen más poliédrica del género estudiado.

- Exhaustividad: a pesar de que el archivo de corpus estudiado solo se representa a sí mismo, se ha intentado superar el alcance de los exámenes manuales (a la antigua usanza: con un único par de textos de original-traducción) para que los resultados obtenidos sean lo suficientemente variados como para cubrir una casuista amplia. Por consiguiente, hemos compilado 20 obras y todas ellas completas (frente a estudios con extractos que también existen en la disciplina).
- Medio: se eligen textos escritos, que en principio han sido objeto de estudio de la mayor parte de trabajos basados en corpus.
- Tipología textual y tema: se escogen novelas traducidas al español de temática variada, producidas originariamente en inglés e italiano con un elevado potencial cultural.
- Lengua(s): es importante destacar que la lengua central del estudio es el español traducido y que se han elegido dos lenguas originales distintas, inglés e italiano, para comprobar el comportamiento de productos de trasvase que parten de dos focos lingüísticos diferentes. En este momento, el método de trabajo es solo provisional y, precisamente por ello hemos incluido una variante lingüística en cuanto al punto de partida. En estudios futuros se mantendrá esta variante que puede ser muy fructífera en cuanto a la homogeneidad / variabilidad de los resultados.
- Fecha de publicación: se opta por obras cuyos originales se publicaron entre 1930 y 2000. Por un lado, los textos que constituyen el archivo de corpus LIT_ENIT_ES pertenecen a décadas distintas del siglo pasado (caracterizadas por acontecimientos sociales, políticos y culturales diversos) con el fin de tener acceso a un amplio abanico de referencias culturales. Por otro, las décadas permiten cierto paralelismo y comparabilidad entre los originales ingleses e italianos.

3.2.4 Proceso de preparación del corpus

Como ya explicaba Gandin (2009:145) en el apartado 3.1.4 del presente capítulo, antes de empezar el análisis, el corpus pasa por fases de preparación (que incluyen o no procesos de codificación). En nuestro caso, la preparación del archivo LIT_ENIT_ES se inicia con la descarga de los textos del archivo en formato digital. A continuación los documentos se procesan con editor *JEdit*, con el que se limpian, ajustan manualmente y, en último término, se convierten en formato UTF-8. Llegado a este punto, el corpus paralelo se alinea con *Intertext Editor* de modo que se pueda comparar los elementos culturales encontrados en los textos traducidos con los de los textos originales.

3.2.5 Metodología de análisis

Para lograr los objetivos prefijados y corroborar la hipótesis de partida, los corpus objeto de estudio se analizan con la herramienta AntConc en dos etapas diferenciadas:

- análisis descriptivos;
- análisis comparativos.

3.2.5.1 Análisis descriptivos: identificación y clasificación de culturemas en el corpus de literatura traducida LIT_ES y en los subcorpus de obras traducidas del inglés LIT_EN_ES y del italiano LIT_IT_ES

El objetivo de esta fase es conocer más a fondo el complejo de corpus de literatura traducida LIT_ES para confirmar la aplicabilidad de las clasificaciones teóricas (capítulo 1) y obtener más información acerca del mismo con vista a posibles estudios futuros. Con este fin se procederá con arreglo a dos fases:

- 1ª fase: identificación y clasificación de los cultuemas en categorías culturales

- 2ª fase: : identificación de cultuemas más frecuentes de la literatura traducida

3.2.5.1.1 1ª fase: identificación y clasificación de los cultuemas en categorías culturales

En esta primera fase se intentará identificar los elementos culturales del corpus de literatura traducida LIT_ES sin considerar su frecuencia, mediante el análisis de la lista entera de palabras, para luego clasificarlos según las categorías culturales establecidas en el apartado 1.3 del primer capítulo del presente estudio. Para alcanzar este objetivo se desarrollarán tres tareas:

- se utilizará el buscador de concordancias *AntConc* para crear un listado de palabras en orden alfabético; se utilizará una *stop-list* para excluir del corpus las palabras gramaticales, que no pueden considerarse elementos culturales y, por lo tanto, no resultan útiles para este estudio;

- se identificarán, entre las palabras resultantes, los elementos que podrían tener una connotación cultural (por ejemplo el término *California* podría tener una connotación cultural, mientras el término *mesa* no);

- por cada palabra se analizará el listado de concordancias y colocaciones para identificar su contexto y forma de uso en eventuales expresiones coloquiales, idiomáticas y exclamativas con el fin de comprobar (o descartar) su uso como elemento cultural (por ejemplo el término *nazca* fue inicialmente considerado un referente cultural, por su posible pertenencia a la categoría de los topónimos y referencias a lugares específicos, pero una vez consultadas sus concordancias y su contexto —“Y,

una vez que nazca, será ciudadano...” — ha resultado ser una forma verbal del verbo nacer, sin ningún matiz cultural y, por esto, se borró de la lista de culturemas);

- una vez establecida su naturaleza como elemento cultural, las palabras se clasificarán según la propuesta de categorías del apartado 1.3 del capítulo 1, nacida de la unión de los planteamientos de diferentes estudiosos.

Concluida esta fase se procederá a contabilizar resultados totales en el complejo de corpus LIT_ES y en los corpus parciales que componen el complejo LIT_EN_ES y LIT_IT_ES:

| Categoría cultural | | Número de culturemas por categoría | | |
|--------------------|--------------|---|--|--|
| Categoría | Subcategoría | En el corpus de literatura traducida (LIT_ES) | En el subcorpus inglés-español (LIT_EN_ES) | En el subcorpus italiano-español (LIT_IT_ES) |
| | | | | |

3.2.5.1.2 2ª fase: identificación de culturemas más frecuentes de la literatura traducida

Para identificar los culturemas más frecuentes del complejo de corpus traducidos al español (LIT_ES) y de cada uno de los corpus que lo constituyen (LIT_EN_ES Y LIT_IT_ES) se hará uso de la herramienta *AntConc*, con la que:

- se generará un listado de palabras en orden de frecuencia, del que se examinarán los elementos léxicos con una aparición superior a 50. Para la identificación de elementos léxicos (frente a palabras funcionales) se utilizarán una *stop list*. La elección de 50 como umbral para la selección de los términos (es decir, términos que constituyan un 0,01% del total) corresponde a una decisión intuitiva con la que se pretende descartar términos con poca incidencia en el corpus pero incluir un abanico de resultados lo suficientemente amplio como para acceder a culturemas variados;

- se identificarán, entre las palabras resultantes, los elementos culturales ya identificados en la primera fase del presente estudio descriptivo (apartado 3.2.5.1.1);

- se dividirán los culturemas encontrados según su categoría cultural (recuérdese que estas categorías culturales se corresponden a las del marco teórico propuesto en el capítulo 1);

Al concluir esta segunda fase, por cada uno de los subcorpus (LIT_EN_ES y LIT_IT_ES) se cumplimentará una ficha de los culturemas encontrados divididos por categoría cultural, con la frecuencia, la posición en el listado de frecuencia, y las utilidades más o menos frecuentes, como la presentada a continuación:

| Culturema | Categoría cultural | Ocurrencias | Posición en el listado de frecuencia | Comentarios en cuanto al uso |
|------------------|---------------------------|--------------------|---|-------------------------------------|
| | | | | |

En resumen, las dos fases presentadas en este apartado representan los dos métodos seguidos en este análisis. En la primera fase, de identificación y clasificación de culturemas en categorías culturales (3.2.5.1.1), se ha seguido un método cualitativo que permite identificar un mayor número de culturemas recurriendo a la clasificación establecida en el marco teórico (1.3). En la segunda fase, de identificación de culturemas más frecuentes de la literatura traducida (3.2.5.1.2), se ha adoptado un método más cuantitativo que permite destacar los culturemas más usados en el corpus analizado.

3.2.5.2. Análisis comparativos

Tras la descripción de los corpus, procederemos a comparar los textos traducidos con sus originales (3.2.5.2.1.) con el fin de identificar las técnicas traductoras, clasificarlas (a partir del marco teórico propuesto en el capítulo 2), y calibrar su peso en el archivo. Asimismo proponemos cruzar nuestras listas de frecuencias con listas de corpus de referencia para extraer conclusiones vinculadas al peso idiosincrásico de los culturemas.

3.2.5.2.1 Análisis comparativo interno

En el primer estudio comparativo se hará uso de los corpus paralelos LIT_EN_ES y LIT_IT_ES (apartado 3.2.2) a fin de comparar los culturemas identificados en la fase anterior con los de las respectivas obras originales, para identificar, clasificar y calibrar las técnicas de traducción utilizadas en dichos corpus. Para la clasificación de las técnicas se recurrirá al marco teórico del apartado 2.5 del capítulo 2 del presente estudio.

El análisis se realizará con la ayuda del alineador *Intertext Editor* que no solo servirá para enlazar originales y traducciones sino que se utilizará también para buscará

elementos culturales en originales y traducciones. Este programa resulta útil para el objetivo del trabajo ya que consigue dividir el texto en segmentos y alinear cada segmento de la obra original con el segmento correspondiente en la obra traducida, permitiendo así comparar más rápida y fácilmente los elementos culturales previamente identificados en el corpus de literatura traducida con los elementos correspondientes en los textos originales.

Para realizar los análisis comparativos internos se acometerán las siguientes tareas:

- identificación de los elementos culturales en los textos originales;
- anotación de los mismos junto a las respectivas traducciones en español;
- determinación de técnicas traductoras para cada culturema (por medio de la comparación entre los culturemas utilizados en originales y traducciones) según la clasificación de técnicas establecidas y explicadas en el segundo capítulo del presente estudio (véase apartado 2.5).

Los datos resultantes se insertarán en una ficha en la cual se recoge la frecuencia de cada técnica de traducción utilizada en el complejo de corpus traducido (LIT_ES) así como LIT_EN_ES y LIT_IT_ES. Dicha ficha será como la presentada continuación:

| Técnica de traducción | Aparición por lengua | | Aparición total |
|------------------------------|-------------------------------------|-----------------------------------|------------------------|
| | del italiano (LIT_IT_ES) | del inglés (LIT_EN_ES) | |
| | | | |

3.2.5.2.2 Análisis comparativo externo

El objetivo de este análisis comparativo es identificar las palabras más idiosincrásicas del corpus de literatura traducida (LIT_ENIT_ES) para ver en qué porcentaje estas se corresponden con los culturemas encontrados en las fases precedentes de este estudio y verificar en qué medida los culturemas contribuyen a la especificidad de los corpus. Además, por medio de este análisis comparativo externo, se podrá establecer cuáles son los culturemas que adquieren mayor importancia en el momento de establecer el perfil idiosincrásico del corpus analizado. Asimismo el análisis externo también nos permitirá identificar las palabras más idiosincrásicas de los corpus originales en inglés y en italiano para confrontar la idiosincrasia de originales y traducciones.

Para desarrollar este análisis se necesita un generador de palabras clave (como el que ofrece el *AntConc*) y listados de las palabras más frecuentes de corpus de referencia en español, inglés e italiano. Para este estudio se han utilizado:

- el listado de las 10.000 palabras más frecuentes del Corpus del Español Actual (CREA);
- el listado de las 10.000 palabras más frecuentes del Bank of English (BoE);
- el listado *Leeds Italian Internet Frequency List*, recopilado por Serge Sheroff con las 10.000 palabras más frecuentes de un vastísimo corpus de referencia de italiano.¹

El análisis se compone de las siguientes tareas.

- Creación de listado de palabras clave del complejo de corpus de literatura traducida al español (LIT_ES) (como es bien sabido para crear listados de palabras clave se cruza un corpus de referencia con el corpus de análisis, ambos en la misma lengua, mediante

¹ <http://www.comp.leeds.ac.uk/ssharoff/>

programas informáticos como *AntConc*);

- Análisis del peso idiosincrásico de los culturemas en LIT_ES, considerándose los primeros 100 términos del listado de palabras clave;
- Creación de listado de palabras clave del corpus de literatura traducida al español desde el inglés (LIT_EN_ES);
- Análisis del peso idiosincrásico de los culturemas en LIT_EN_ES, considerándose los primeros 100 términos del listado de palabras clave;
- Creación de listado de palabras clave del corpus de literatura traducida al español desde el italiano (LIT_IT_ES);
- Análisis del peso idiosincrásico de los culturemas en LIT_IT_ES, considerándose los primeros 100 términos del listado de palabras clave;
- Creación de listado de palabras clave del corpus original de literatura en inglés;
- Análisis del peso idiosincrásico de los culturemas en el corpus de literatura en inglés, considerándose los primeros 100 términos del listado de palabras clave;
- Creación de listado de palabras clave del corpus original de literatura en italiano;
- Análisis del peso idiosincrásico de los culturemas en el corpus de literatura en italiano, considerándose los primeros 100 términos del listado de palabras clave;
- Comparación de los resultados de los análisis anteriores.

En resumen, en este capítulo se han explicado conceptos teóricos de la metodología de corpus y conceptos concretos de nuestro trabajo. En el capítulo sucesivo se procederá a aplicar todos estos conceptos al análisis de nuestro archivo de corpus.

4 ANÁLISIS DESCRIPTIVOS Y COMPARATIVOS DEL CORPUS LIT_ENIT_ES

Siguiendo la metodología adoptada y explicada en el capítulo anterior, el análisis de este estudio se desarrollará en dos etapas:

- análisis descriptivos;
- análisis comparativos.

En la primera etapa, se identificarán y se clasificarán los elementos culturales del corpus LIT_ES según las categorías presentadas en el marco teórico del apartado 1.3. En la segunda etapa, se realizará primero un comparación interna consistente en el cotejo de los elementos culturales de LIT_ES con los de las obras originales con el fin de establecer y clasificar su tratamiento en torno a las técnicas de traducción propuestas en el apartado 2.5. Asimismo, se desarrollará un análisis comparativo externo para estudiar el peso idiosincrásico de los culturemas en la literatura traducida y compararlo con el peso que estos tienen en las obras en lengua original.

4.1. Análisis descriptivos

Como ya se ha explicado en el capítulo de metodología (véase el apartado 3.2.5.1 del capítulo precedente), los análisis descriptivos corresponderán a dos fases:

- identificación y clasificación de los culturemas en categorías culturales;
- identificación de culturemas de la literatura traducida más frecuentes.

4.1.1 1ª fase: identificación y clasificación de los culturemas en categorías culturales

Como ya se ha explicado detalladamente en el apartado relativo a esta fase del análisis descriptivo en el capítulo de metodología (véase apartado 3.2.5.1.1), para clasificar los culturemas del corpus de literatura traducida LIT_ES según las categorías culturales establecidas en el apartado 1.3 del presente estudio, se desarrollan tres tareas:

- se utiliza el buscador de concordancias *AntConc* para crear un listado de palabras ordenadas en orden alfabético (se utiliza una *stop-list*, que puede consultarse en el anexo 1, para excluir del corpus las palabras gramaticales que no pueden ser consideradas elementos culturales y por lo tanto no resultan útiles para este estudio). Así resulta una lista de palabras formada por 515.727 *tokens* y 52.587 *types*).

- Entre las palabras resultantes, se identifican los elementos que podrían tener una connotación cultural. De este procedimiento resultan 11.012 elementos que, por pertenecer a las categorías de clasificación de culturemas establecidas, podrían ser considerados culturemas.

- Por cada palabra se analiza el listado de concordancias y colocaciones para identificar su contexto y forma de uso en eventuales expresiones coloquiales, exclamaciones, expresiones idiomáticas con el fin de comprobar (o descartar) su uso como elemento cultural. Con esta operación se eliminan 61 elementos (presentados en el anexo 2) por no tener naturaleza cultural, así que el listado de culturemas definitivo está formado por 10.951 elementos, de los cuales 6.810 aparecen en el subcorpus de obras traducidas del inglés LIT_EN_ES y 4.141 en el subcorpus de obras traducidas del italiano LIT_IT_ES.

- Una vez comprobada su naturaleza como elemento cultural, las palabras se clasificarán según las categorías precedentemente elegidas en la propuesta definitiva de clasificación (presentada en el apartado 1.3 del capítulo 1), nacida de la unión de las diferentes propuestas de los autores de estudios de traducción precedentes a este análisis. Los resultados de esta clasificación se reúnen en la ficha ya presentada en el apartado 3.2.5.1.1, en la cual se indica la categoría cultural, la subcategoría, el número de cultreemas otorgados a dicha categoría en el complejo de corpus LIT_ES y en los corpus parciales que componen el complejo LIT_EN_ES (inglés-español) y LIT_IT_ES (italiano-español). La ficha total obtenida se presenta a continuación:

| Categoría cultural | | Número de cultreemas por categoría | | |
|------------------------------|---|---|--|--|
| Categoría | Subcategoría | En el corpus de literatura traducida (LIT_ES) | En el subcorpus inglés-español (LIT_EN_ES) | En el subcorpus italiano-español (LIT_IT_ES) |
| Cultura geográfica y natural | | 2.274 | 1.258 | 1.016 |
| | Flora y fauna | 637 | 271 | 366 |
| | Topónimos y referencias a lugares específicos | 1.637 | 987 | 650 |
| Cultura material | Referencias a | 1.033 | 725 | 308 |

| | | | | |
|---|--|--------------|--------------|--------------|
| | objetos específicos de una cultura | | | |
| Cultura social, histórica y de folclore | | 2.304 | 1.271 | 1.033 |
| | Referencias a personajes históricos y políticos | 396 | 139 | 257 |
| | Familia | 371 | 270 | 101 |
| | Amistad | 3 | 3 | 0 |
| | Organización y papeles sociales | 1.278 | 649 | 629 |
| | Acciones de vida cotidiana | 256 | 210 | 46 |
| Cultura religiosa | Religión | 1.546 | 232 | 1.314 |
| Cultura conceptual | Creencias, tradiciones y supersticiones | 59 | 51 | 8 |
| Cultura lingüística | | 1.468 | 1.436 | 32 |
| | Juegos de palabras, refranes, etc. | 41 | 27 | 14 |
| | Nivel de | 1.383 | 1.366 | 17 |

| | | | | |
|------------------------|--------------------------|--------------|--------------|--------------|
| | formalidad | | | |
| | Tabús | 44 | 43 | 1 |
| Interferencia cultural | | 0 | 0 | 0 |
| | Inequivalencias | 0 | 0 | 0 |
| | Falsos amigos culturales | 0 | 0 | 0 |
| | Injerencia cultural | 0 | 0 | 0 |
| Total | | 8.684 | 4.973 | 3.711 |

Tabla 10: clasificación culturemas LIT_ES

Como se puede observar de la fila de los resultados, en el corpus LIT_ES 8.684 elementos culturales (del total de 10.951) se han clasificado según las categorías establecidas en el marco teórico presentado en el capítulo 1. Del análisis del corpus resulta que los 2.267 culturemas restantes se pueden agrupar en una categoría que no se había considerado en la propuesta de categorización establecida previamente y presentada en el apartado 1.3. Esta categoría, que reúne los nombres de personajes (excluyendo los nombres de los protagonistas de las novelas, que no se han considerado por el alto número de ocurrencias, lo cual habría falseado los resultados) y los nombres propios de personas sin ningún interés histórico, político o social, se ha denominado categoría de los nombres propios y se ha añadido a los resultados obtenidos por el presente estudio.

| Categoría cultural | | Número de cultuemas por categoría | | |
|--------------------|--------------|---|--|--|
| Categoría | Subcategoría | En el corpus de literatura traducida (LIT_ES) | En el subcorpus inglés-español (LIT EN_ES) | En el subcorpus italiano-español (LIT IT_ES) |
| Nombres propios | | 2.267 | 1.837 | 430 |

Tabla 11: categoría nombres propios

En resumen, resulta que los cultuemas más utilizados en el corpus de literatura traducida LIT_ES son los que pertenecen a las categorías presentadas a continuación en orden de frecuencia:

| Categoría cultural | Ocurrencias | Porcentaje del número total de cultuemas |
|---|-------------|--|
| Cultura social, histórica y de folclore | 2.304 | 21% |
| Cultura geográfica y natural | 2.274 | 20,7% |
| Nombres propios | 2.267 | 20,7% |

Tabla 12: cultuemas más utilizados en LIT_ES

Por lo que se refiere a las subcategorías, las más fructíferas en el corpus de literatura traducida LIT_ES, presentadas por orden de frecuencia, son:

| Categoría cultural > subcategoría | Ocurrencias | Porcentaje del número total de culturemas |
|--|--------------------|--|
| Cultura geográfica y natural > Topónimos y referencias a lugares específicos | 1.637 | 14,9% |
| Cultura lingüística > Nivel de formalidad | 1.383 | 12,6% |
| Cultura social, histórica y de folclore > Organización y papeles sociales | 1.278 | 11,6% |

Tabla 13: subcategorías más utilizadas en LIT_ES

La categoría menos fructífera es la de la cultura conceptual, con 59 culturemas, que representan el 0,5% (8 culturemas en las obras traducidas del italiano, correspondientes al 0,1% y 51 en las obras traducidas del inglés, o sea un 0,7%). Otro punto de interés es la categoría de la cultura lingüística, compuesta por 1.468 culturemas (el 13,4% del total), la mayoría de los cuales pertenecen a la subcategoría del nivel de formalidad, con 1.383 culturemas, mientras las demás subcategorías, la de los juegos de palabras o refranes y la de los tabús, se quedan respectivamente con solo 41 y 44 culturemas (el 0,3% y el 0,4% del total). Finalmente, es imprescindible destacar que la categoría de la interferencia cultural se ha quedado completamente vacía en cada una de sus subcategorías. Una explicación a este fenómeno podría ser la dificultad en identificar dichos elementos en este nivel de mera clasificación, quizás por la falta de una comparación con los textos originales. Esta hipótesis se podrá comprobar en la siguiente

etapa del presente estudio, el análisis comparativo, a través del alineado y de la comparación de las obras con los respectivos textos originales.

Si se consideran los subcorpus analizados por separado (LIT_EN_ES, formado por obras traducidas del inglés y LIT_IT_ES, formado por obras traducidas del italiano), el orden de frecuencia de las categorías culturales es diferente de lo que resulta del análisis del corpus global de literatura traducida LIT_ES. Por esta razón y para presentar los resultados del estudio en su totalidad, a continuación se presentan las categorías y las subcategorías más utilizadas en los dos subcorpus (LIT_EN_ES inglés-español y LIT_IT_ES italiano-español):

Categorías culturales más utilizadas en el subcorpus de obras traducidas del inglés (LIT_EN_ES)

| Categoría cultural | Ocurrencias | Porcentaje del número total de culturemas |
|---|--------------------|--|
| Nombres propios | 1.837 | 26,9% |
| Cultura lingüística | 1.436 | 21% |
| Cultura social, histórica y de folclore | 1.271 | 18,6% |

Tabla 14: culturemas más utilizados en LIT_EN_ES

Categorías culturales más utilizadas en el subcorpus de obras traducidas del italiano (LIT_IT_ES)

| Categoría cultural | Ocurrencias | Porcentaje del número total de culturemas |
|---------------------------|--------------------|--|
|---------------------------|--------------------|--|

| | | |
|---|-------|-------|
| Cultura religiosa | 1.314 | 31,7% |
| Cultura social, histórica y de folclore | 1.033 | 24,9% |
| Cultura geográfica y natural | 1.016 | 24,5% |

Tabla 15: culturemas más utilizados en LIT_IT_ES

Si entre los resultados de la división en categorías, como se acaba de presentar, hay bastantes diferencias entre el corpus global LIT_ES y cada uno de los subcorpus que lo compone LIT_EN_ES y LIT_IT_ES, los resultados relativos a la división en subcategorías para el corpus global LIT_ES y los dos subcorpus analizados LIT_EN_ES y LIT_IT_ES son más parecidos:

Subcategorías más utilizadas en el subcorpus de obras traducidas del inglés (LIT_EN_ES)

| Categoría cultural > subcategoría | Ocurrencias | Porcentaje del número total de culturemas |
|--|--------------------|--|
| Cultura lingüística > Nivel de formalidad | 1.366 | 20% |
| Cultura geográfica y natural > Topónimos y referencias a lugares específicos | 987 | 14,4% |
| Cultura social, histórica y de folclore > Organización y papeles sociales | 649 | 9,5% |

Tabla 16: subcategorías más utilizadas en LIT_EN_ES

Subcategorías más utilizadas en el corpus de obras traducidas del italiano (LIT_IT_ES)

| Categoría cultural > subcategoría | Ocurrencias | Porcentaje del número total de culturemas |
|--|--------------------|--|
| Cultura geográfica y natural > Topónimos y referencias a lugares específicos | 650 | 15,6% |
| Cultura social, histórica y de folclore > Organización y papeles sociales | 629 | 15,1% |
| Cultura geográfica y natural > Flora y fauna | 366 | 8,8% |

Tabla 17: subcategorías más utilizadas en LIT_IT_ES

Por lo que se refiere a las subcategorías, es interesante notar que los resultados del corpus general LIT_ES coinciden con los resultados de los dos subcorpus analizados (LIT_EN_ES y LIT_IT_ES), aunque con perfiles de uso diferentes.

4.1.2 2ª fase: identificación de culturemas más frecuentes de la literatura traducida

Siguiendo el apartado dedicado al análisis descriptivo del capítulo de metodología (véase 3.2.5.1.2), en esta fase se analizan el corpus de literatura traducida LIT_ES y los dos subcorpus LIT_EN_ES y LIT_IT_ES que lo componen para identificar los culturemas más frecuentes de las obras traducidas escogidas.

Siguiendo los pasos explicados en el apartado de metodología correspondiente (3.2.5.1.2), se actúa en las fases siguientes:

- utilizando la herramienta *AntConc*, se genera un listado de palabras por orden de frecuencia, del que se examinarán los elementos léxicos con una aparición superior a 50. Para la identificación de elementos léxicos (frente a palabras funcionales) se utilizarán la misma *stop list* utilizada en la primera fase del análisis (apartado 4.1.1) y presentada en el anexo 1. De este procedimiento se obtienen tres listas de palabras:

- una relativa al corpus de literatura traducida LIT_ES formada por 556.881 *tokens* y 5.2587 *types* . De esta lista los términos con una frecuencia mayor o igual a 50 corresponden a las primeras 1.439 palabras y se puede consultar en el anexo 3;
- una relativa al subcorpus inglés-español LIT_EN_ES formada por 267.393 *tokens* y 32.931 *types*. De esta lista los términos con una frecuencia mayor o igual a 50 corresponden a las primeras 760 palabras y se puede consultar en el anexo 4;
- una relativa al subcorpus italiano español LIT_IT_ES, formada por 257.897 *tokens* y 35.595 *types*. De esta lista se consideran los términos con una frecuencia igual o mayor a 50, es decir las primeras 714 palabras que se pueden consultar en el anexo 5 al final de este estudio.

Obsérvese que el número de *tokens* y de *types* no coincide aquí con el número total de *tokens* y de *types* de cada uno de los corpus mencionados (apartado 3.2.2 del presente

capítulo) porque para el análisis se ha utilizado la *stop list* presentada en el anexo 1 al final de este estudio. En efecto, comparando el número de *types* de cada lista de palabras aquí presentadas con el número total de *types* de los respectivos corpus descritos en el apartado 3.2.2, resulta que en cada lista de palabras faltan un número de *types* menor o igual a 77, cifra que corresponde al número de palabras que componen la *stop list* utilizada.

- Se consideran, entre las palabras resultantes en cada lista, los elementos culturales ya identificados en la primera fase del presente estudio descriptivo (véase el apartado precedente 4.1.1). De este modo entre los términos con frecuencia igual o superior a 50 de cada lista de palabras resultan:

- 11.021 elementos culturales para el corpus de literatura traducida LIT_ES
- 6.175 elementos culturales para el subcorpus inglés-español LIT_EN_ES;
- 3.219 elementos culturales para el subcorpus italiano-español LIT_IT_ES.

- Se dividen los culturemas encontrados según las categorías culturales establecidas en el apartado 1.3 y utilizadas en la primera fase del análisis descriptivo (apartado 4.1.1) recogiendo los resultados en fichas como la presentada en el apartado 3.2.5.1.2 del capítulo de metodología.

En las ficha total cumplimentada al final de esta segunda fase, que se presentan a continuación, se muestran en la primera columna los elementos culturales identificados

entre las palabras con frecuencia igual o mayor de 50 de cada subcorpus, en la segunda columna la categoría cultural a la cual pertenecen, en la tercera el número de ocurrencias (en caso de haber encontrado el culturema tanto con letra inicial mayúscula como con letra inicial minúscula, se pondrán en la casilla en cuestión los dos números de apariciones unidos con el símbolo +), en la cuarta columna la posición en el listado de frecuencia y, en la quinta columna, aparecen comentarios en torno al uso que se le da al término (en expresiones, colocaciones, refranes, etc., que se indican junto con el número de frecuencia de cada locución señalado por una x seguida por el número en cuestión al lado de cada expresión). Si no existe nada especial que reseñar, la casilla se deja vacía.

Ficha de los culturemas (de más de 50 ocurrencias) más frecuentes en el corpus de literatura traducidas (LIT_ES)

| Culturema | Categoría cultural | Ocurrencias | Posición en el listado de frecuencia | Comentarios en cuanto al uso |
|------------------|--|--------------------|---|---|
| Madre | Cultura social, histórica y de folclore > Familia | 988 | 34 | Expresión: De puta madre x 8 (Welsh) Sentido religioso: Reina Madre de los Cielos x 4 Madre superiora x10 (Welsh) Madre de Dios x2/de Cristo x1/ de la Divina Gracia x5 Santa Madre Iglesia x6 Madre naturaleza x1 |

| | | | | |
|-------|---|-----|-----|--|
| | | | | Utilización en sentido familiar: mi madre x641, tu madre x74, su madre x109 |
| Padre | Cultura social, histórica y de folclore > Familia | 811 | 53 | En sentido religioso x446: Padre todopoderoso x1, padre Santo x4, padre herbolario x2, Padre nuestro x2... Utilización en sentido familiar x365: mi padre x199, tu padre x49, su padre x57, nuestro padre x5, padre de familia x2 |
| Usted | Cultura lingüística > Nivel de formalidad | 751 | 64 | Utilizado solo como pronombre de cortesía |
| Dios | Cultura religiosa | 162 | 214 | Exclamación: !Dios mío! x42 Gracias a Dios x3 Exclamaciones: ¡Dios! x2 /¡Dios Santo! x4/ ¡Ay Dios! x4 (solo en Welsh) Dios la bendiga x5 Jurar por Dios x1 Dios del Cielo x3 |

| | | | | |
|-------|---|----------|---------|--|
| | | | | Dios lo quiera x4 Voluntad de Dios x5 Santo Dios x3 Dios santo x4 Como Dios manda x8 Que Dios guarde x2 que Dios perdone x8 que Dios proteja x4 |
| Tío | Cultura social, histórica y de folclore > Familia | 453 | 132 | Utilizado en sentido de amistad |
| Señor | Cultura lingüística > Organización y papeles sociales | 196 +199 | 170/339 | Nuestro Señor x27 Señor Jesucristo x5 ¡Señor!; Señores (exclamaciones)x 6 Querido señor x6 El señor+ nombre propio x45 ¡Oh Señor! X6 |
| Abad | Cultura religiosa | 318 | 203 | |
| Boy | Nombres propios | 265 | 238 | |
| Hija | Cultura social, histórica y de folclore | 257 | 248 | |

| | | | | |
|---------|---|-----|-----|---|
| | > Familia | | | |
| Capullo | Cultura lingüística > Nivel de formalidad | 232 | 279 | |
| Tía | Cultura social, histórica y de folclore > Familia | 226 | 285 | |
| Cabrón | Cultura lingüística > Nivel de formalidad | 217 | 295 | Cacho Cabrón x10 |
| Putá | Cultura lingüística > Nivel de formalidad | 211 | 310 | Hostia puta x6 Ni puta idea x3 De una puta vez x8 Hijo/a/os/as de puta x37 De puta madre x4 |
| Hijo | Cultura social, histórica y de folclore > Familia | 210 | 312 | En sentido religioso: Hijo de Dios x4 Hijo de María x2 Hijo de Eva x1 En sentido familiar: Expresiones: hijo de esa época x1 Hijo de la perdición x1 Hijo de perra x7 Hijo de puta x23 |

| | | | | |
|---------|---|-----|-----|--|
| | | | | |
| Begbie | Nombres propios | 209 | 151 | |
| Familia | Cultura social, histórica y de folclore > Familia | 198 | 340 | |
| Señora | Cultura lingüística > Organización y papeles sociales | 193 | 356 | Amable señora x2 Honorada señora x1 |
| Hermano | Cultura social, histórica y de folclore > Familia | 143 | 249 | Expresiones: Como hermano y hermana x2 (Tan) Como una hermana que protege a su hermano menor (Tan) Con sentido familiar y de amistad |
| Mierda | Cultura lingüística > Nivel de formalidad | 173 | 400 | Le importa una mierda x2 Una puta mierda x5 Un montón de puta mierda x2 Exclamaciones: ¡Vete a la mierda! x11 |

| | | | | |
|----------|---|-----|-----|---|
| Puto | Cultura lingüística > Nivel de formalidad | 158 | 439 | |
| Padres | Cultura social, histórica y de folclore > Familia | 153 | 452 | |
| Perro | Cultura geográfica y natural > Flora y fauna | 143 | 486 | |
| Caballo | Cultura geográfica y natural > Flora y fauna | 142 | 487 | Utilizado como sinónimo de heroína en el lenguaje coloquial x32 |
| Cristo | Cultura religiosa | 140 | 494 | Pobre Cristo x3 (Sciascia) |
| Señorita | Cultura lingüística > Organización y papeles sociales | 125 | 561 | Utilizado solo como pronombre de cortesía |
| Marido | Cultura social, | 123 | 577 | |

| | | | | |
|-----------|--|-----|-----|-----------------------------------|
| | histórica y de folclore > Familia | | | |
| Mamá | Cultura social, histórica y de folclore > Familia | 121 | 586 | |
| Charlotte | Nombres propios | 116 | 611 | |
| Dolly | Nombres propios | 106 | 666 | |
| Hijos | Cultura social, histórica y de folclore > Familia | 106 | 667 | También con sentido de amistad |
| Concetta | Nombres propios | 104 | 676 | |
| Candy | Nombres propios | 103 | 682 | |
| Billy | Nombres propios | 100 | 703 | |
| Adelmo | Nombres propios | 99 | 714 | |
| Tíos | Cultura social, | 99 | 722 | Utilizado como sinónimo de amigos |

| | | | | |
|----------|---|----|-----|--|
| | histórica y de folclore > Familia | | | |
| Diablo | Cultura religiosa | 93 | 762 | |
| Rey | Cultura social, histórica y de folclore > Organización y papeles sociales | 93 | 766 | |
| Dólares | Cultura material | 91 | 775 | |
| Dolores | Nombres propios | 89 | 797 | |
| Lisboa | Cultura geográfica y natural > Topónimos y referencias a lugares específicos | 89 | 799 | |
| Hermanos | Cultura social, histórica y | 88 | 806 | En sentido religioso: hermanos laicos x2 Venerables hermanos |

| | | | | |
|------------|--|----|-----|----------------------------------|
| | de folclore > Familia | | | x3 |
| Excelencia | Cultura social, histórica y de folclore > Organización y papeles sociales | 83 | 851 | |
| Jodido | Cultura lingüística > Nivel de formalidad | 81 | 880 | |
| Café | Cultura material | 80 | 886 | |
| Colega | Cultura social, histórica y de folclore > Amistad | 80 | 887 | Gran colega x2 Buen colega x4 |
| Taxi | Cultura material | 77 | 932 | |
| Hermanas | Cultura social, histórica y de folclore > Familia | 76 | 934 | |

| | | | | |
|-----------|--|----|-------|--|
| Pub | Cultura geográfica y natural > Topónimos y referencias a lugares específicos | 74 | 959 | |
| Cardenal | Cultura religiosa | 69 | 1.022 | |
| Emperador | Cultura social, histórica y de folclore > Organización y papeles sociales | 69 | 1.024 | |
| Capullos | Cultura lingüística > Nivel de formalidad | 68 | 1.034 | |
| York | Cultura geográfica y natural > Topónimos y referencias a lugares | 67 | 1.069 | |

| | | | | |
|-------------|---|----|-------|--------------|
| | específicos | | | |
| Papá | Cultura social, histórica y de folclore > Familia | 59 | 639 | Papá Noel x2 |
| Ministro | Cultura social, histórica y de folclore > Organización y papeles sociales | 64 | 1.121 | |
| Gato | Cultura geográfica y natural > Flora y fauna | 63 | 1.139 | |
| Joder | Cultura lingüística > Nivel de formalidad | 63 | 1.142 | |
| Universidad | Cultura social, histórica y de folclore > Organizació | 63 | 1.150 | |

| | | | | |
|---------|---|----|-------|--|
| | n y papeles sociales | | | |
| Conejos | Cultura geográfica y natural > Flora y fauna | 62 | 1.153 | |
| Italia | Cultura geográfica y natural > Topónimos y referencias a lugares específicos | 60 | 1.201 | |
| Londres | Cultura geográfica y natural > Topónimos y referencias a lugares específicos | 59 | 1.218 | |
| Señores | Cultura social, histórica y de folclore > Organizació | 58 | 1.238 | |

| | | | | |
|----------|--|----|-------|--|
| | n y papeles sociales | | | |
| Aviñón | Cultura geográfica y natural > Topónimos y referencias a lugares específicos | 56 | 1.260 | |
| Hijas | Cultura social, histórica y de folclore > Familia | 56 | 1.266 | |
| Princesa | Cultura social, histórica y de folclore > Organización y papeles sociales | 56 | 1.274 | |
| Club | Cultura material | 55 | 1.289 | |
| Ay | Cultura lingüística > Juegos de palabras, | 54 | 1.313 | |

| | | | | |
|-----------|---|----|-------|---|
| | refranes, etc. | | | |
| Cabrones | Cultura lingüística > Nivel de formalidad | 54 | 1.315 | |
| Crooks | Nombres propios | 54 | 699 | |
| Francia | Cultura geográfica y natural > Topónimos y referencias a lugares específicos | 54 | 1.322 | |
| Rosa | Cultura geográfica y natural > Flora y fauna | 54 | 1.335 | Expresiones: Hay una rosa entre las espinas x1 (Welsh) Fresca como una rosa x1 (Tabucchi) |
| Jesús | Cultura religiosa | 53 | 1.352 | Exclamaciones: ¡Jesús! X6 ¡Jesús María! X3 Jesús crucificado x3 |
| Beardsley | Cultura geográfica y natural > | 52 | 723 | |

| | | | | |
|--------------|---|--------------|-------|--|
| | Topónimos y referencias a lugares específicos | | | |
| Caballos | Cultura geográfica y natural > Flora y fauna | 52 | 1.368 | |
| Cojones | Cultura lingüística > Nivel de formalidad | 52 | 1.372 | |
| Drogas | Cultura material | 51 | 1.395 | |
| Total | | 11021 | | |

Tabla 18: culturema más frecuentes en LIT_ES

De la ficha presentada resulta que las categorías de culturemas (de más de 50 ocurrencias) más frecuentes en el corpus de literatura traducida LIT_ES son:

| Categoría cultural > subcategoría | Frecuencia | Types de culturemas por categoría | Porcentaje del total de los culturemas |
|---|-------------------|--|---|
| Cultura social, histórica y de folclore > Familia | 4.167 | 18 | 37,8% |

| | | | |
|---|---------------|-----------|-------------|
| Cultura lingüística> Nivel de formalidad | 2.008 | 10 | 18,2% |
| Cultura social, histórica y de folclore > Organización y papeles sociales | 1.251 | 11 | 11,3% |
| Nombres propios | 1.245 | 10 | 11,2% |
| Cultura religiosa | 835 | 6 | 7,5% |
| Cultura geográfica y natural > Flora y fauna | 516 | 6 | 4,6% |
| Cultura geográfica y natural > topónimos y referencias a lugares específicos | 511 | 8 | 4,6% |
| Cultura material | 354 | 5 | 3,1% |
| Cultura social, histórica y de folclore > Amistad | 80 | 1 | 0,7% |
| Cultura lingüística > Juegos de palabras, refranes, etc. | 54 | 1 | 0,4% |
| Total | 11.021 | 75 | 100% |

Tabla 19: categorías culturemas más frecuentes LIT_ES

Para que la distribución de las categorías culturales de los elementos culturales identificados resulte más clara, a continuación se presentan los datos recogidos en un gráfico visual:

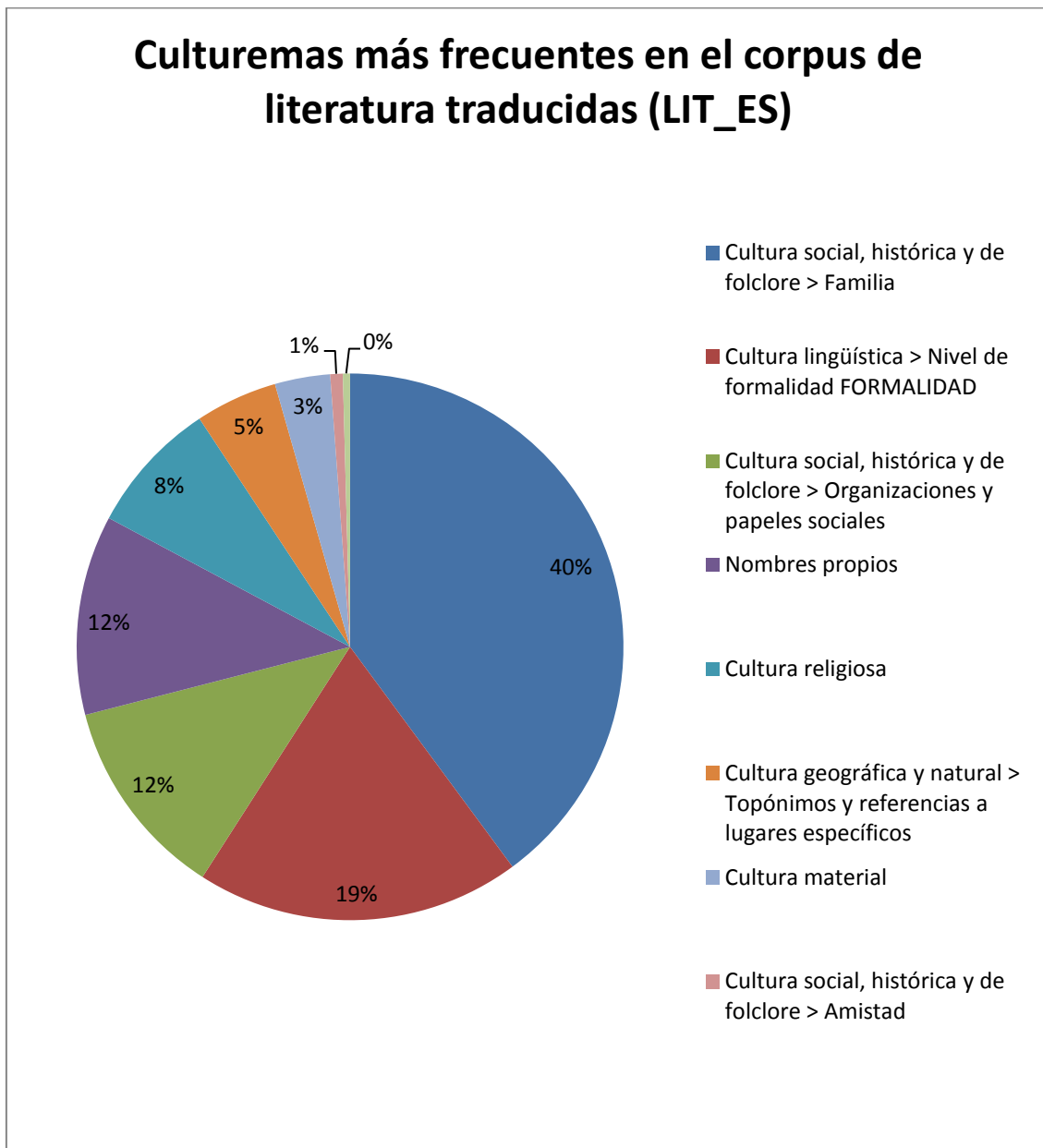


Tabla 20: Gráfico visual culturemas más frecuentes en el corpus d literatura traducida LIT_ES

De la categorización efectuada resultan necesarias algunas observaciones. En primer lugar es interesante notar que, en el corpus de literatura traducida LIT_ES, solo la subcategoría de la familia, (la más utilizada entre los culturemas de más de 50 ocurrencias) incluye más de un tercio del total de los culturemas identificados (37,8%). Otro fenómeno interesante es que 3 de las 10 categorías identificadas en este análisis de culturemas frecuentes son subcategorías de la categoría de la cultura social, histórica y de folclore que, con el 49,8 % de los culturemas totales encontrados (un 37,8% en la subcategoría de la familia, un 11,3% en la de la organización y papeles sociales y un 0,7% en la de la amistad), se puede considerar la categoría más fructífera del corpus analizado.

Además, para algunos de los culturemas identificados, es interesante comparar la categoría cultural (que se pueden consultar en la segunda columna de la ficha) con las expresiones y colocaciones (en la quinta columna). En la tabla que se presenta a continuación se compara la categoría establecida inicialmente y con el uso del elemento cultural en su contexto. En la primera columna se encuentra el elemento cultural, en la segunda la clasificación inicial (establecida mediante el listado de frecuencia), en la tercera las expresiones y colocaciones identificadas para cada culturema y en la última columna la nueva categoría semántica, resultante del análisis del elemento cultural en su contexto.

| Elemento cultural | Categoría cultural | Comentarios en cuanto al uso | Nueva categoría cultural (en algunos contextos) |
|--------------------------|---------------------------|-------------------------------------|--|
| | | | |

| | | | | |
|--------|---|--|---|--|
| Madre | Cultura social, histórica y de folclore > Familia | Expresión: De puta madre x 8 Sobre todo utilización en sentido familiar Sentido religioso: Reina Madre de los Cielos x 4 Madre superiora x10 | > | Cultura lingüística > Nivel de formalidad Cultura religiosa |
| Padre | Cultura social, histórica y de folclore > Familia | En sentido religioso x446: Padre todopoderoso x1, padre Santo x4, padre herbolario x2, Padre nuestro x2... Utilización en sentido familiar x365: mi padre x199, tu padre x49, su padre x57, nuestro padre x5, padre de familia x2 | > | Cultura religiosa |
| Tío | Cultura social, histórica y de folclore > Familia | Utilizado en sentido de amistad | > | Cultura social, histórica y de folclore > Amistad |
| Señor | Cultura social, histórica y de folclore > Organización y papeles sociales | Nuestro Señor x27 Señor Jesucristo x5 ¡Señor!/¡Señores (exclamaciones)x 6 Querido señor x6 El señor+ nombre propio x45 ¡Oh Señor! X6 | > | Cultura religiosa |
| Señora | Cultura social, | Amable señora x2 | > | Cultura religiosa |

| | | | | |
|----------|--|---|---|---|
| | histórica y de folclore > Familia | Honorada señora x1 | | |
| Hermano | Cultura social, histórica y de folclore > Familia | Expresiones: Como hermano y hermana x2 (Tan) Como una hermana que protege a su hermano menor (Tan) Con sentido familiar y de amistad | > | Cultura social, histórica y de folclore > Amistad Cultura religiosa |
| Caballo | | Utilizado como sinónimo de eroina en el lenguaje coloquial x32 | > | Cultura material |
| Hijos | Cultura social, histórica y de folclore > Familia | También con sentido de amistad | > | Cultura social, histórica y de folclore > Amistad |
| Tíos | Cultura social, histórica y de folclore > Familia | Utilizado como sinónimo de amigos | > | Cultura social, histórica y de folclore > Amistad |
| Hermanos | Cultura social, histórica y de folclore > Familia | En sentido religioso: hermanos laicos x2 Venerables hermanos x3 | > | Cultura religiosa |

Tabla 21: cambio de categoría culturemas en LIT_ES

De la tabla de comparación resulta que algunos elementos de la subcategoría de la familia, una vez analizados en su contexto, resultan pertenecer a la categoría de la

cultura religiosa. Así el término *madre* se utiliza también en sentido religioso con referencia a la Virgen, en expresiones como *Reina Madre de los Cielos* (4 veces), a la institución eclesiástica *Santa Madre Iglesia* (6 veces) o a los cargos religiosos como en *Madre superiora* (en 10 casos); *padre*, se utiliza con referencia a Dios como en *Padre Santo* (4 veces) o *Padre todopoderoso* (1 vez) y como apelativo de frailes y sacerdotes, como en *padre herbolario* (2 veces); y *hermanos* presenta las expresiones *hermanos laicos* (2 veces) y *venerables hermanos* (3 veces).

También el título de cortesía *señor*, inicialmente incluido a la subcategoría de la organización y papeles sociales, después de la comparación con algunas de sus expresiones y colocaciones como *Nuestro Señor* (en 27 casos) o *Señor Jesucristo* (5 veces), pasa a pertenecer a la categoría de la cultura religiosa.

El sustantivo *madre*, ya examinado por su reconversión de la subcategoría de la familia a la categoría de la cultura religiosa, en algunos casos se puede también atribuir a la subcategoría del nivel de formalidad, por ejemplo en expresiones como *de puta madre* (10 veces).

Algunos términos de la categoría de la familia en cambio resultan pertenecer por su uso a la subcategoría de la amistad. Es el caso de los términos *tío* y *tíos* que se encuentran exclusivamente utilizados como sinónimos de amigo/s. Finalmente, un caso esporádico pero curioso, el término *caballo*, elemento cultural inicialmente atribuido a la categoría de la flora y la fauna, resulta en algunos casos pertenecer a la categoría de la cultura material por su utilización, en el lenguaje coloquial, como sinónimo de la palabra *heroína*.

Ficha de los culturremas (de más de 50 ocurrencias) más frecuentes en el subcorpus de literatura traducidas del inglés al español (LIT_EN_ES)

| Culturrema | Categoría cultural | Ocurrencias | Posición en el listado de frecuencia | Comentarios en cuanto al uso |
|-------------------|--|--------------------|---|---|
| Madre | Cultura social, histórica y de folclore > Familia | 930 | 14 | Expresión: De puta madre x 8 (Welsh) Sobre todo utilización en sentido familiar Sentido religioso: Reina Madre de los Cielos x 4 Madre superiora x10 (Welsh) |
| Padre | Cultura social, histórica y de folclore > Familia | 313 | 84 | Solo en sentido familiar |
| Boy | Nombres propios | 264 | 109 | |
| Capullo | Cultura lingüística > Nivel de formalidad | 231 | 130 | |
| Tía | Cultura social, histórica y de folclore | 216 | 146 | |

| | | | | |
|--------|--|-----|-----|--|
| | > Familia | | | |
| Cabrón | Cultura lingüística > Nivel de formalidad | 215 | 147 | Cacho Cabrón x10 |
| Begbie | Nombres propios | 209 | 151 | |
| Puta | Cultura lingüística > Nivel de formalidad | 204 | 157 | Hostia puta x6 Ni puta idea x3 Quejarse como putas De una puta vez x8 |
| Señor | Cultura lingüística > Organización y papeles sociales | 196 | 170 | Solo como pronombres de cortesía |
| Usted | Cultura lingüística > Nivel de formalidad | 185 | 183 | Solo como pronombre de cortesía |
| Hija | Cultura social, histórica y de folclore > Familia | 184 | 184 | |
| Mierda | Cultura | 169 | 204 | Le importa una mierda x2 |

| | | | | |
|---------|--|-----|-----|---|
| | lingüística > Nivel de formalidad | | | Una puta mierda x2 Un montón de puta mierda x2 Exclamaciones: ¡Vete a la mierda! x11 |
| Dios | Cultura religiosa | 162 | 214 | Exclamación: ¡Dios mío! x30 Gracias a Dios x3 Exclamaciones: ¡Dios! x2 /¡Dios Santo! x4/ ¡Ay Dios! x4 (solo en Welsh) Dios la bendiga x5 Jurar por Dios x1 Dios el Cielo x3 |
| Puto | Cultura lingüística > Nivel de formalidad | 158 | 217 | |
| Familia | Cultura social, histórica y de folclore > Familia | 149 | 234 | |
| Señora | Cultura lingüística > Organización y papeles | 149 | 236 | Amable señora x2 Honorada señora x1 |

| | | | | |
|-----------|---|-----|-----|--|
| | sociales | | | |
| Hermano | Cultura social, histórica y de folclore > Familia | 143 | 249 | Expresiones: Como hermano y hermana x2 (Tan) Como una hermana que protege a su hermano menor (Tan) Con sentido familiar y de amistad |
| Hijo | Cultura social, histórica y de folclore > Familia | 142 | 250 | |
| Charlotte | Nombres propios | 116 | 303 | |
| Padres | Cultura social, histórica y de folclore > Familia | 116 | 305 | Solo en sentido familiar |
| Dolly | Nombres propios | 106 | 336 | |
| Candy | Nombres propios | 103 | 353 | |
| Billy | Nombres propios | 100 | 364 | |

| | | | | |
|---------|--|----|-----|---|
| Perro | Cultura geográfica y natural > Flora y fauna | 97 | 381 | |
| Tíos | Cultura social, histórica y de folclore > Familia | 94 | 398 | |
| Dolores | Nombres propios | 81 | 447 | |
| Jodido | Cultura lingüística > Nivel de formalidad | 81 | 449 | |
| Colega | Cultura social, histórica y de folclore > Amistad | 76 | 488 | Gran colega x2 Referido directamente a alguien x49 |
| Pub | Cultura geográfica y natural > Topónimos y referencias a lugares específicos | 73 | 509 | |

| | | | | |
|----------|---|----|-----|--|
| Señorita | Cultura lingüística > Nivel de formalidad | 69 | 541 | |
| Capullos | Cultura lingüística > Nivel de formalidad | 67 | 549 | |
| Joder | Cultura lingüística > Nivel de formalidad | 63 | 585 | |
| Conejos | Cultura geográfica y natural > Flora y fauna | 62 | 593 | |
| York | Cultura geográfica y natural > Topónimos y referencias a lugares específicos | 61 | 614 | |
| Hermanas | Cultura social, histórica y de folclore | 59 | 631 | |

| | | | | |
|-----------|--|----|-----|--------------------------------|
| | > Familia | | | |
| Papá | Cultura social, histórica y de folclore > Familia | 59 | 639 | |
| Hijos | Cultura social, histórica y de folclore > Familia | 57 | 668 | También con sentido de amistad |
| Cabrones | Cultura lingüística > Nivel de formalidad | 54 | 698 | |
| Crooks | Nombres propios | 54 | 699 | |
| Londres | Cultura geográfica y natural > Topónimos y referencias a lugares específicos | 53 | 715 | |
| Beardsley | Cultura geográfica y natural > | 52 | 723 | |

| | | | | |
|--------------|---|--------------|-----|--|
| | Topónimos y referencias a lugares específicos | | | |
| Cojones | Cultura lingüística > Nivel de formalidad | 51 | 735 | |
| Colegas | Cultura social, histórica y de folclore > Amistad | 51 | 736 | |
| Drogas | Cultura material | 51 | 740 | |
| Café | Cultura material | 50 | 747 | |
| Total | | 6.175 | | |

Tabla 22: culturemas más frecuentes en LIT_EN_ES

De la ficha presentada, resulta que las categorías culturales de culturemas más frecuentes (de más de 50 ocurrencias) del subcorpus de literatura traducida del inglés LIT_EN_ES son las que se resumen en la tabla a continuación:

| Categoría cultural > subcategoría | Ocurrencias | Types de culturemas por categoría | Porcentaje del total de los culturemas |
|---|--------------------|--|---|
| Cultura social, | 2.462 | 12 | 39,8% |

| | | | |
|--|--------------|-----------|-------------|
| histórica y de folclore > Familia | | | |
| Cultura lingüística> Nivel de formalidad | 1.478 | 11 | 23,9% |
| Nombres propios | 1.033 | 8 | 16,7% |
| Cultura social, histórica y de folclore > Organización y papeles sociales | 414 | 3 | 6,7% |
| Cultura geográfica y natural > topónimos y referencias a lugares específicos | 239 | 4 | 3,8% |
| Cultura religiosa | 162 | 1 | 2,6% |
| Cultura geográfica y natural > Flora y fauna | 159 | 2 | 2,5% |
| Cultura social, histórica y de folclore > Amistad | 127 | 2 | 2% |
| Cultura material | 101 | 2 | 1,6% |
| Total | 6.175 | 45 | 100% |

Tabla 23: categorías de culturemas más frecuentes en LIT_EN_ES

Para que la distribución de las categorías de los elementos culturales identificados resulte más clara, a continuación se presentan los datos recogidos en un gráfico visual:

Culturemas más frecuentes en el subcorpus de literatura traducidas del inglés al español (LIT_EN_ES)

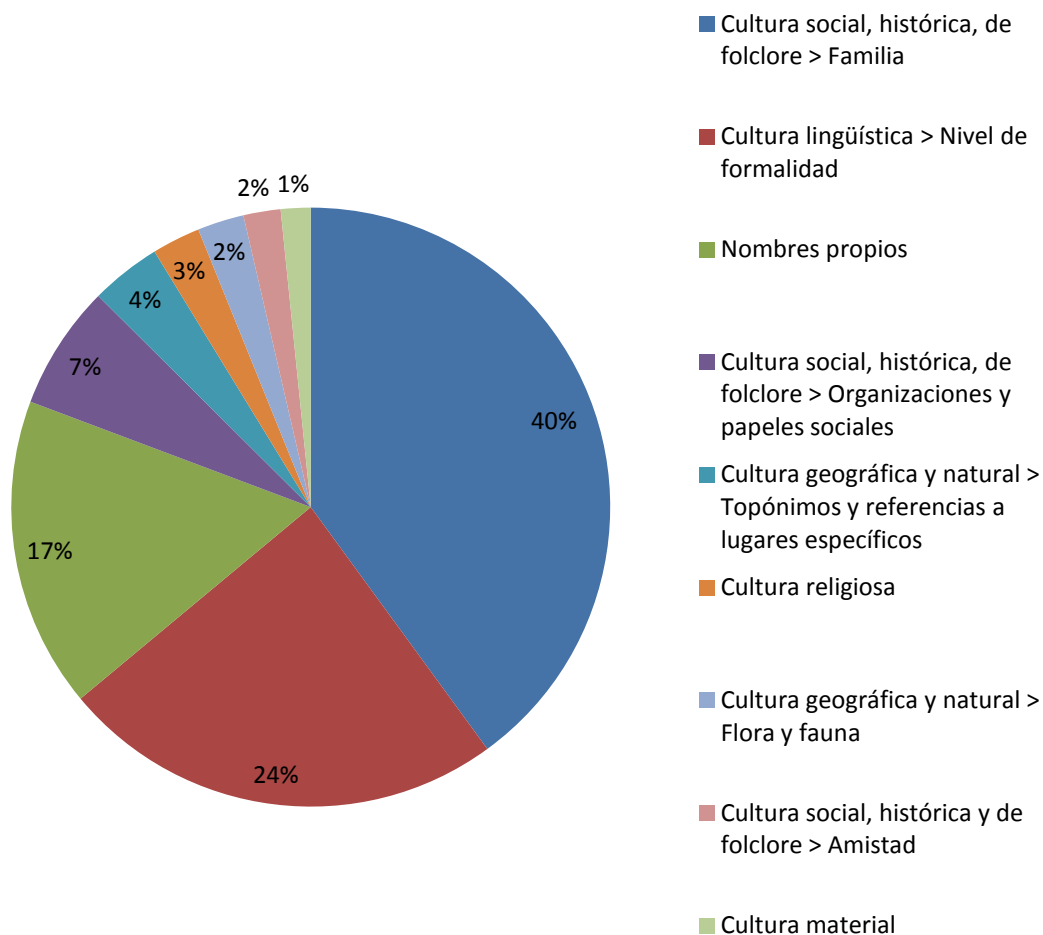


Tabla 24: Gráfico visual culturemas más frecuentes en el corpus d literatura traducida LIT_EN_ES

Cabe realizar algunos comentarios sobre la categorización anterior. Algunas categorías nos plantean aspectos reseñables evidentes. En primer lugar aparece de forma clara la cantidad de elementos culturales que dentro de la subcategoría del nivel de formalidad se identifican con insultos y expresiones que remiten a la falta de cortesía, con un total

de 1.293 culturemas, cantidad que corresponde al 87,4% de la subcategoría a la cual pertenecen. Algunos ejemplos son las expresiones y colocaciones como *ni puta idea* (3 veces), *¡hostia puta!* (6 veces), *cacho cabrón* (encontrado 10 veces) o *¡vete a la mierda!* (11 casos). Otros aspecto relevante desde el punto de vista cuantitativo es la dominación de la subcategoría de la familia que cuenta con más de un tercio del total de los culturemas encontrados (39,8%). Finalmente, entre los elementos la cultura religiosa, se ha identificado exclusivamente la palabras *Dios*, que se encuentra por la mayoría en expresiones y exclamaciones como *¡Dios mío!* (30 veces), *¡gracias a Dios!* (en 3 casos) o *¡Ay Dios!* (4 veces).

Como en el corpus de literatura LIT_ES analizado arriba, también de los resultados del análisis de este subcorpus LIT_EN_ES se observa un fenómeno interesante con respecto a la comparación entre la categoría cultural otorgada a algunos de los culturemas identificados (en la segunda columna de la tabla 21 presentada anteriormente) y sus expresiones y colocaciones (presentadas en la quinta columna de la misma ficha). A continuación se utiliza la misma tabla utilizada precedentemente para el corpus LIT_ES, para presentar la comparación:

| Elemento cultural | Categoría cultural | Comentarios en cuanto al uso | | Nueva categoría cultural (en algunos contextos) |
|--------------------------|---|--|---|--|
| Madre | Cultura social, histórica y de folclore > Familia | Expresión: De puta madre x 8 (Welsh) Sobre todo utilización en sentido familiar Sentido religioso: Reina | > | Cultura religiosa |

| | | | | |
|---------|--|---|---|--|
| | | Madre de los Cielos x 4 Madre superiora x10 (Welsh) | | |
| Hermano | Cultura social, histórica y de folclore > Familia | Expresiones: Como hermano y hermana x2 (Tan) Como una hermana que protege a su hermano menor (Tan) Con sentido familiar y de amistad | > | Cultura social, histórica y de folclore > Amistad |
| Hijos | Cultura social, histórica y de folclore > Familia | También con sentido de amistad | > | Cultura social, histórica y de folclore > Amistad |
| Tía | Cultura social, histórica y de folclore > Familia | | | Cultura social, histórica y de folclore > Amistad |
| Tíos | Cultura social, histórica y de folclore > Familia | | | Cultura social, histórica y de folclore > Amistad |

Tabla 25: cambio de categoría cultuemas LIT_EN_ES

De la comparación se desprende que algunos elementos culturales de la subcategoría de la familia, una vez estudiados en su contexto a través de las expresiones y colocaciones identificadas en el corpus, pertenecen a otras categorías, sobre todo a la categoría de la amistad (en 4 de los 5 casos resultantes), con los términos *hermano* o *hijos*, utilizados

por los personajes para referirse a amigos o personas cercanas sin ningún vínculo familiar, y los términos *tía* y *tíos* utilizados como sinónimos de amiga/os. Aunque en menor medida, algunos elementos de la categoría semántica de la familia se desplazan también a la categoría de la religión. En este caso destaca el término de género femenino *madre*, con expresiones como *Reina madre de los Cielos* (encontrado 4 veces) o (*Madre Superiora* 10 veces).

Ficha de los cultuemas (con más de 50 ocurrencias) más frecuentes en el subcorpus de literatura traducidas del italiano al español (LIT_IT_ES)

| Culturema | Categoría cultural > subcategoría | Ocurrencias | Posición en el listado de frecuencia | Comentarios en cuanto al uso |
|-----------|---|-------------|--------------------------------------|--|
| Usted | Cultura lingüística > nivel de formalidad | 532 | 28 | Utilizado solo como pronombre de cortesía |
| Padre | Cultura social, histórica y de folclore > familia | 498 | 33 | Mayoría= utilización en campo religioso: Padre todopoderoso, padre Santo x4, padre herbolario... Padre de familia x1 Utilización en sentido familiar: mi padre x21, tu padre x12 |
| Dios | Cultura | 322 | 83 | Dios lo quiera x4 |

| | | | | |
|----------|--|-----------|---------|--|
| | religiosa | | | Voluntad de Dios x5 Santo Dios x3 Como Dios manda x8 que Dios peroné x8 ¡Dios mío! x12 que Dios proteja x4 |
| Abad | Cultura religiosa | 318 | 86 | |
| Señor | Cultura social, histórica y de folclore> organización y papeles sociales | 255 + 199 | 120/160 | El Señor x25 Nuestro Señor x27 Señor Jesucristo x5 ¡Señor!/¡Señores (exclamaciones)x 6 Querido señor x6 El señor+ nombre propio x45 |
| Cristo | Cultura religiosa | 128 | 272 | Pobre Cristo x3 (Sciascia) |
| Concetta | Nombre propio | 104 | 337 | |
| Santo | Cultura religiosa | 104 | 338 | |
| Lisboa | Cultura geográfica y cultural > tóponimos y referencias a objetos | 89 | 399 | |

| | | | | |
|------------|---|----|-----|--|
| | específicos de una cultura | | | |
| Excelencia | Cultura social, histórica y de folclore > organización y papeles sociales | 82 | 435 | |
| Rey | Cultura social, histórica y de folclore > organización y papeles sociales | 82 | 437 | |
| Caballo | Cultúra geográfica y natural > flora y fauna | 79 | 451 | |
| Diablo | Cultura religiosa | 69 | 499 | |
| Hijo | Cultura social, histórica y de folclore > familia | 68 | 506 | En sentido religioso: Hijo de Dios x4 Hijo de María x2 Hijo de Eva x1 En sentido familiar: |

| | | | | |
|----------|--|----|-----|--|
| | | | | Expresiones: hijo de esa época x1 Hijo de puta x1 Hijo de la perdición x1 Exclamación ¡Hijos! x1 |
| Hija | Cultura social, histórica y de folclore > familia | 64 | 539 | |
| Madre | Cultura social, histórica y de folclore > familia | 58 | 606 | Madre de Dios x2/de Cristo x1/ de la Divina Gracia x5 Madre Iglesia x6 Madre naturaleza x1 Mayoría= utilización en sentido familiar |
| Aviñón | Cultura geográfica y cultural > Topónimos y referencias a objetos específicos de una cultura | 56 | 619 | |
| Hermanos | Familia | 56 | 626 | En sentido religioso: hermanos laicos x2 |

| | | | | |
|--------------|---|--------------|-----|---------------------------------|
| | | | | Venerables hermanos x3 |
| Señorita | Cultura social, histórica y de folclore > organización y papeles sociales | 56 | 633 | Solo como pronombre de cortesía |
| Total | | 3.219 | | |

Tabla 26: culturemas más frecuentes en LIT_IT_ES

De la ficha presentada resulta que las categorías culturales de culturemas más frecuentes en el subcorpus italiano-español LIT_IT_ES son:

| Categoría cultural > subcategoría | Ocurrencias | Types de culturemas por categoría | Porcentaje del total de los culturemas |
|---|--------------------|--|---|
| Cultura religiosa | 941 | 5 | 29,2% |
| Cultura social, histórica y de folclore > Familia | 744 | 5 | 23,1% |
| Cultura social, histórica y de folclore > Organización y papeles sociales | 674 | 4 | 20,9% |
| Cultura lingüística > Nivel de formalidad | 532 | 1 | 16,5% |

| | | | |
|--|--------------|-----------|-------------|
| Cultura geográfica y natural > Topónimos y referencias a lugares específicos | 145 | 2 | 4,5% |
| Nombres propios | 104 | 1 | 3,2% |
| Cultura geográfica y natural > Flora y fauna | 79 | 1 | 2,4% |
| Total | 3.219 | 19 | 100% |

Tabla 27: categorías de culturemas más frecuentes en LIT_IT_ES

Para que la distribución de las categorías de los elementos culturales identificados resulte más clara, a continuación se presentan los datos recogidos en un gráfico visual:

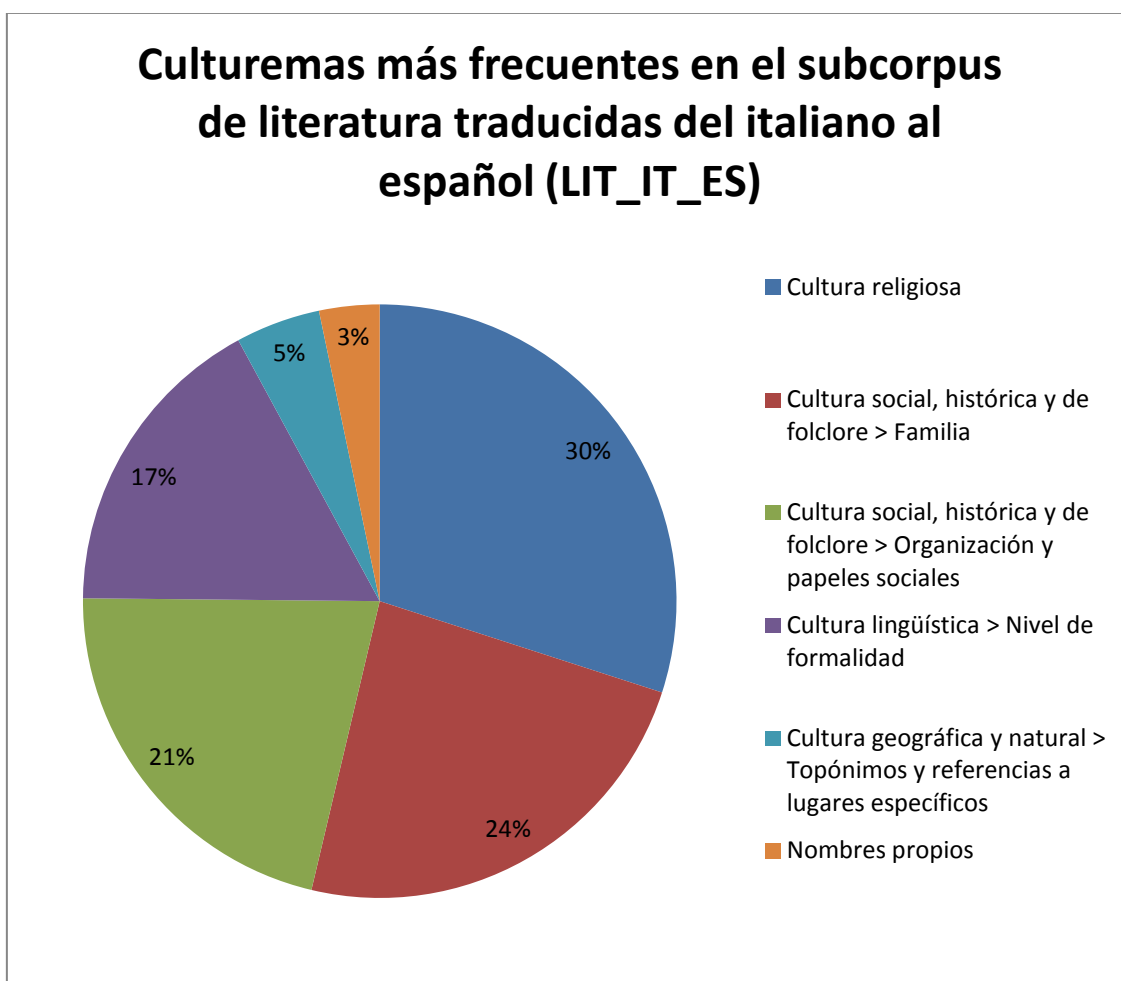


Tabla 28: gráfico visual culturemas más frecuentes culturemas literatura traducida del italiano al español

También en este subcorpus LIT_IT_ES, una vez analizados los resultados del análisis, resultan aspectos reseñables en cuanto a algunas categorías. Por ejemplo, parece oportuno examinar la categoría de la cultura religiosa que, con 941 elementos culturales, representa casi un tercio (el 29,2%) de la totalidad de los culturemas más frecuentes. La palabra más utilizada es *Dios*, con una frecuencia de 322 veces, sobre todo para indicar voluntad, autoridad y superioridad divina, y en casos menos frecuentes como parte de expresiones como *Dios lo quiera* (4 veces), *que Dios me perdone* (8 casos), o como *Dios manda!* (en 8 ocasiones). Otros culturemas de esta categoría encontrados son *abad* (318 veces); *Cristo* (128 veces), término también incluido en expresiones, como *pobre Cristo* (3 casos); *santo* (104 veces); y finalmente el término *diablo* (69 casos). Además de los resultados se puede observar que, también por la cantidad de culturemas diferentes, esta categoría resulta ser la más fructífera ya que incluye 5 de las 19 palabras culturales identificadas.

Además es imprescindible subrayar otro fenómeno resultante de la observación de los datos presentados en la ficha obtenida que ya se ha estudiado para los corpus analizados anteriormente. Es interesante comparar la categoría cultural (en la segunda columna de la ficha) con las expresiones y colocaciones (en la quinta columna) de algunos de los elementos culturales identificados. Los resultados de dicha comparación se presenta a continuación:

| Elemento cultural | Categoría cultural | Expresiones, colocaciones, etc. | | Nueva categoría semántica (en algunos casos) |
|--------------------------|---|---|---|---|
| Padre | Cultura social, histórica y de folclore > Familia | utilización en campo religioso x299 (Padre todopoderoso, Padre Santo x4, padre herbolario, ...) Padre de familia Utilización en sentido familiar: mi padre x21, tu padre x12 | > | Cultura religiosa |
| Señor | Cultura social, histórica y de folclore > Organización y papeles sociales | El Señor x25 Nuestro Señor x27 Señor Jesucristo x5 ¡Señor!/¡Señores (exclamaciones)x 6 Querido señor x6 El señor+ nombre propio x45 | > | Cultura religiosa |
| Hijo | Cultura social, histórica y de folclore > Familia | En sentido religioso: Hijo de Dios x4 Hijo de María x2 Hijo de Eva En sentido familiar Expresiones: hijo de esa época/de puta/ de la perdición Hijas de las delicias> Exclamación ¡Hijos! | > | Cultura religiosa |
| Madre | Cultura social, | Madre de Dios/de Cristo/ | > | Cultura religiosa |

| | | | | |
|----------|--|--|---|-------------------|
| | histórica y de folclore > Familia | del Señor/ de la Divina Gracia x5 Madre Iglesia x6 Madre naturaleza x1 Mayoría= utilización en sentido familiar | | |
| Hermanos | Cultura social, histórica y de folclore > Familia | En sentido familiar: mi/tu/su hermano Hermano mayor En sentido religioso: hermanos laicos x2 Hermano del orden de... Venerables hermanos x3 Como título religioso: Hermano + nombre propio | > | Cultura religiosa |

Tabla 29: cambio de categoría culturemas LIT_IT_ES

Como se puede observar en la tabla, después de comparar la categoría cultural inicialmente asignada a estos culturemas con las expresiones y las colocaciones en las cuales se han utilizado los elementos en cuestión, resulta que efectivamente los elementos culturales analizados pertenecen a otra categoría.

En concreto, ocurre que casi todos los elementos culturales inicialmente insertados en la categoría semántica de la familia se han revelado como pertenecientes, en algunos casos, a la categoría semántica de la religión. Este fenómeno se detecta particularmente en los sustantivos de género masculino: *padre*, relacionado con la religión 299 veces en un total de 498 apariciones, referido a los sacerdotes, también llamados *padres de la Iglesia* (1 vez), o como exclamación en *¡Padre Santo!* (4 veces); *hermano*, encontrado en el subcorpus un total de 56 veces, que también aparece referido a los sacerdotes o a

su forma de relacionarse con los demás en entornos católicos, por ejemplo *venerables hermanos* (3 casos) o *hermanos laicos* (2 casos). Entre los sustantivos de género femenino en cambio, el elemento cultural más utilizado con acepción religiosa es *madre*, encontrado en total en 58 ocasiones, de las cuales 6 hacen referencia a la institución de la Iglesia, *madre iglesia* (6 veces), o a la Virgen con expresiones como *madre de Dios* (2 veces), *madre de la Divina Gracia* (5 veces) o *madre de Cristo* (1 vez). Aunque en una sola ocasión, el culturema *madre* ha aparecido, también, relacionado con la naturaleza (*Madre Naturaleza*). Asimismo, entre los elementos que pertenecen a la subcategoría de la organización y papeles sociales, hay un culturema que en algunas ocasiones resulta más relacionado con la categoría de la cultura religiosa: la palabra *Señor*, referida a Dios en 47 casos (el 14,5% del total de las veces en las cuales el término se ha encontrado en el subcorpus) y, a menudo, utilizada como exclamación, por ejemplo *¡Señor!* (2 casos).

Como conclusión de este apartado se puede considerar que, tanto en el corpus LIT_ES como en los dos subcorpus LIT_EN_ES y LIT_IT_ES que lo componen, existen culturemas que parecen pertenecer a una categoría pero, después de un análisis de su contexto, se ubican en una categoría cultural diferente. Esto enfatiza la indisoluble dependencia de los culturemas de su entorno y la consiguiente necesidad de estudiar estos elementos en su contexto textual.

4.2 Análisis comparativos

Como ya se ha explicado en el apartado 3.2.5.2, también esta segunda etapa del análisis se divide en dos fases. En primer lugar se desarrollarán comparaciones entre los textos traducidos y los respectivos originales, con el fin de identificar y clasificar las técnicas

de traducción utilizadas por cada culturema identificado. Sucesivamente se realizará un análisis comparativo externo, cruzando los corpus de este estudio con corpus de referencia para alcanzar conocimientos sobre el peso idiosincrásico de los culturemas.

4.2.1 Análisis comparativo interno

El objetivo de esta fase es comparar los culturemas identificados en la fase anterior en los subcorpus de literatura traducida LIT_EN_ES y LIT_IT_ES con los de los respectivos textos originales, para identificar, a través de la comparación, las técnicas de traducción utilizadas y clasificarlas según el marco teórico propuesto en el apartado 2.5 del capítulo 2.

Para alcanzar dicho objetivo, se utiliza el programa *Intertext Editor* que permite alinear segmentos del texto original con segmentos del texto traducido. En el caso de este estudio, el uso del alineado automático arroja resultados escasos porque los textos elegidos para su estudio se han bajado directamente de la web, el OCR es de baja calidad y las líneas vacías son diseminadas por todo el texto en lugar de sistemáticamente al final de párrafos o capítulos. De modo que, después de haber seleccionado las opciones *paragraph break: empty line* y *automatic sentence splitter*, es necesario revisar la alineación y corregirla manualmente. Los ajustes que se han necesitado con más frecuencia son la unión o la división de segmentos y la corrección de errores de ortografía o cambios de letras debidos a la baja calidad del OCR. De este modo se ha obtenido un alineado en el que no siempre se mantiene la relación 1:1 entre segmentos originales y traducidos, pero que resulta igualmente útil para el propósito de este trabajo, ya que los segmentos más largos ocupan de unas 4 o 5 líneas, lo cual se considera una medida apta para buscar y encontrar rápidamente en los textos originales

los culturemas correspondientes a los identificados en el corpus de literatura traducida. Desafortunadamente, el programa *Intertext Editor* tiene el límite de alinear solo un texto a la vez, lo cual ha ralentizado el estudio.

Finalmente, es importante señalar que durante el proceso de alineación se ha descubierto que, en algunas de las obras originales, faltan partes de texto, presentes en cambio en las correspondientes obras traducidas, o viceversa. En la mayoría de los casos, la causa de estas ausencias es la descarga de textos incompletos de la web, y solo en algunas ocasiones las omisiones parecen ser una elección del traductor. Como es necesario considerar este factor para los fines del análisis, a continuación se presenta el número de segmentos que faltan por cada obra analizada, ordenadas según el idioma y el años de publicación:

- Steinbeck: en la obra original faltan 100 segmentos (985-1.016, 2.830-2.899) presentes en la obra traducida;
- Nabokov: La prefacio (segmentos 1-33) aparece solo en la obra original y falta en la traducción, al revés que el epílogo (desde el 5.484 hasta el final) que solo está en la traducción.
- Tan: en el texto traducido al español faltan 13 segmentos (1.406, 2.104-2.110, 2.983-2.989), presentes en el texto original;
- Welsh: el texto traducido ha sufrido muchas eliminaciones (segmentos 475, 731, 1.404, 2.332, 5.123). Estas ausencias parecen ser elecciones del traductor ya que las frases se siguen correctamente por sentido y gramática, razón por la cual se excluye la opción de un texto incompleto;

- Calvino: en el texto original faltan 75 segmentos (870-890, 961-972, 865-884, 956-966, 1.733-1.740, 2.067-2.075) presentes en el texto traducido;

- Tabucchi: en el texto original falta el epílogo (segmentos 1.903-1.934), presente en el texto traducido;

Siguiendo los pasos presentados en el capítulo de metodología (véase el apartado 3.3.2.2), una vez alineados los textos, para realizar el análisis comparativo interno se acometerán las siguientes tareas:

- se localizan en el alineado los elementos culturales identificados en las traducciones (10.951 elementos culturales, 6.810 en el subcorpus LIT_EN_ES y 4.141 en el subcorpus LIT_IT_ES) por medio de la opción de *Intertext* “*find in the right column*” y se localizan los elementos correspondientes en los respectivos segmentos de los textos originales;

- los elementos culturales resultantes en las obras originales se anotan junto a las respectivas traducciones en español;

- comparando el elemento cultural de la obra traducida con el de la obra original, se establece la técnica traductora utilizada que se clasifica según la propuesta presentada en el apartado 2.5. Por ejemplo, el elemento cultural *cabrón*, identificado en las obras traducidas, ha resultado corresponder en las obras en lengua original a los términos *caprone* (en italiano), *cunt* o *sonuvabitch* (en inglés). En el caso del italiano la técnica de traducción utilizada se ha clasificado como traducción literal (ya que no se han utilizado equivalentes culturales como *cornuto* o *bastardo*) y en los casos del inglés como equivalente cultural.

Para establecer la naturaleza léxica de los culturemas analizados, útil para establecer los métodos de traducción empleados, se utilizan unos diccionarios de referencias que permiten establecer qué términos son parte integrante de un idioma, cuáles son préstamos y cuáles están ausentes en el diccionario y, por consiguiente, son transposiciones de una lengua a otra. Los diccionarios utilizados, todos en versión online, son el DRAE como referencia para el español, el *Oxford Dictionary* para el inglés y el diccionario *Treccani* para el italiano.

Los resultados del análisis comparativo interno se reúnen en la ficha que se presenta a continuación, en la cual se indica la técnica utilizada por el traductor (en la primera columna), la aparición de dicha técnica en cada uno de los subcorpus analizados LIT_EN_ES y LIT_IT_ES (en las segunda y tercera columnas) y la aparición total de cada técnica (en la cuarta columna).

| Técnica de traducción | Aparición por corpus | | Aparición total |
|-----------------------|---------------------------------------|---|-----------------|
| | Subcorpus inglés-español LIT_EN_ES | Subcorpus italiano-español LIT_IT_ES | |
| Transposición | 2.747 | 1.030 | 3.777 |
| Préstamo | 471 | 63 | 534 |
| Naturalización | 411 | 311 | 722 |
| Traducción literal | 2.019 | 2.602 | 4.621 |
| Neutralización | 4 | 0 | 4 |
| Hiperónimo | 17 | 2 | 19 |

| | | | |
|----------------------------------|--------------|--------------|---------------|
| Doblete | 7 | 2 | 9 |
| Traducción estándar aceptada | 0 | 1 | 1 |
| Paráfrasis explicativa | 0 | 1 | 1 |
| Notas a pie de página | 16 | 7 | 23 |
| Supresión de informaciones | 4 | 15 | 19 |
| Omisión | 0 | 0 | 0 |
| Equivalente cultural o funcional | 1.035 | 90 | 1.125 |
| Creación autónoma | 52 | 2 | 54 |
| Total | 6.783 | 4.126 | 10.909 |

Tabla 30: técnicas de traducción literatura traducida LIT_ES, LIT_EN_ES y LIT_IT_ES

Como se observa comparando los números totales de técnicas de traducción que se localizan al final de la ficha con el número total de culturemas encontrados en los corpus, algunas técnicas traductorales se quedan excluidas de la categorización propuesta en el marco teórico del apartado 2.4 del presente estudio. En efecto, en algunos casos se ha identificado una falta de equivalencia en el plano semántico o estilístico, por ejemplo el término *cheerleader* traducido con la expresión *jugador de fútbol* se ha considerado una falta de equivalencia semántica porque el significado varía entre original y traducción, y el término *sterco* traducido con la palabra *mierda* se ha considerado una falta de equivalencia estilística, ya que en la obra original se ha utilizado un término menos vulgar que en la obra traducida. Esta técnica, que no se había considerado en la clasificación establecida precedentemente, se denominará “inequivalencia” y se añadirá a la clasificación de técnicas de traducción considerada hasta ahora. Los resultados de

esta nueva categoría para los subcorpus LIT_EN_ES y LIT_IT_ES (inglés-español e italiano-español) se presentan a continuación:

| Técnica de traducción | Aparición por lengua | | Aparición total |
|------------------------------|-----------------------------|-------------------|------------------------|
| | del italiano | del inglés | |
| Inequivalencia | 6 | 3 | 9 |

Tabla 31: técnica de traducción inequivalencia en los corpus de literatura traducida LIT_ES, LIT_EN_ES y LIT_IT_ES

La identificación de esta técnica podría ser la comprobación de una hipótesis parcial formulada en el apartado precedente acerca de la ausencia de la categoría “interferencia cultural” y sus subcategorías (véase los resultados de la clasificación de los culturemas en categorías culturales en el apartado 4.1.2). La técnica de traducción clasificada como inequivalencia puede servir para rellenar la laguna creada por la falta de culturemas en la categoría de la interferencia cultural. No cabe duda de que existen faltas de equivalencias entre los textos originales y las traducciones en cuanto a la traducción de los elementos culturales, sin embargo, por su naturaleza de técnica de traducción más que de categoría cultural, estas características solo se han podido encontrar a través de una comparación entre traducciones y originales.

Además en la ficha que resume las técnicas de traducción identificadas faltan las técnicas de 34 culturemas, respectivamente 14 del corpus inglés-español LIT_EN_ES y 10 del corpus de traducciones del italiano LIT_IT_ES. Estas técnicas no se han conseguido identificar por la ausencia de extractos (tanto de originales como de

traducciones) que se ha explicado al comienzo del presente apartado referido al proceso de alineación de los textos.

En resumen, de los datos se desprende que las técnicas de traducción más utilizadas en los dos subcorpus LIT_EN_ES y LIT_IT_ES, son las presentadas en orden de utilización en la tabla a continuación:

| Subcorpus de las obras traducidas del inglés LIT_EN_ES | Subcorpus de las obras traducidas del italiano LIT_IT_ES |
|---|---|
| Transposición 56% | Traducción literal 62% |
| Traducción literal 41,2% | Transposición 24,1% |
| Préstamo 9,6% | Naturalización 7,4% |
| Naturalización 8,4% | Préstamo 1,5% |

Tabla 32: técnicas de traducción más utilizadas en los subcorpus LIT_EN_ES y LIT_IT_ES

Entre las técnicas menos utilizadas se encuentran la paráfrasis explicativa y la traducción estándar aceptada, ambas encontradas solo en el subcorpus LIT_IT_ES. Finalmente es interesante notar que la categoría de la omisión no ha resultado ser nada fructífera, ya que dicha técnica no se ha identificado en ninguna de las obras analizadas. De esta fase del análisis se puede concluir que las técnicas de traducción más utilizadas son las exotizantes, que se acercan más a la cultura original.

4.2.2 Análisis comparativo externo

Como ya se ha explicado en el capítulo de metodología (véase el apartado 3.2.5.2.2), el análisis comparativo externo permitirá establecer y comparar el peso de los culturemas como palabras claves en la literatura traducida y en la literatura en lengua original. Además por medio de este análisis será posible calibrar si la traducción afecta a la idiosincrasia de los textos y en qué medida lo hace.

Siguiendo el orden presentado en el capítulo de metodología (apartado 3.2.5.2.2), el análisis se desarrollará mediante las siguientes tareas:

- por medio de la comparación de nuestro corpus con el corpus de referencia CREA, se crea la lista de palabras claves relativa al corpus de literatura traducida LIT_ES;
- una vez obtenida dicha lista, se consideran y se analizan las primeras 100 palabras (que se pueden consultar en el anexo 6 al final del presente estudio), hasta identificar entre ellas la presencia de culturemas.

Los culturemas identificados de dichos análisis se presentan en las fichas a continuación ordenados según su valor de *keyness*:

Elementos culturales identificados entre los 100 primeros términos de la lista de palabras clave del corpus de literatura traducida global LIT_ES

| Orden en el listado de palabras clave | Palabra clave | Frecuencia | <i>Keyness</i> |
|--|----------------------|-------------------|-----------------------|
|--|----------------------|-------------------|-----------------------|

| | | | |
|----|-----------|-----|-----------|
| 5 | Abad | 318 | 3.224.259 |
| 12 | Capullo | 232 | 2.352.290 |
| 16 | Cabrón | 217 | 2.200.202 |
| 18 | Begbie | 209 | 2.119.089 |
| 23 | Senor | 199 | 2.017.697 |
| 32 | Puto | 158 | 1.601.990 |
| 46 | Curley | 124 | 1.257.258 |
| 50 | Charlotte | 116 | 1.176.145 |
| 59 | Dolly | 106 | 1.074.753 |
| 60 | Concetta | 104 | 1.054.475 |
| 66 | Tíos | 99 | 1.003.779 |
| 67 | Tío | 453 | 998.540 |
| 81 | Jodido | 81 | 821.274 |
| 87 | Pub | 74 | 750.299 |
| 92 | Capullos | 68 | 689.464 |

| | | | |
|--------------|--|--------------|--|
| Total | | 2.558 | |
|--------------|--|--------------|--|

Tabla 33: culturemas clave en LIT_ES

El mismo análisis se efectuará para el subcorpus de literatura traducida del inglés LIT_EN_ES siguiendo los siguientes pasos:

- por medio de la comparación de nuestro subcorpus traducido del inglés con el corpus de referencia CREA, se crea el listado de palabras claves relativo al subcorpus de literatura traducida del inglés LIT_EN_ES;
- una vez obtenida dicha lista, se consideran y se analizan las primeras 100 palabras (que se pueden consultar en el anexo 7 al final del presente estudio), para identificar entre ellas la presencia de culturemas.

Los resultados del análisis del peso idiosincrásico de los culturemas en el subcorpus LIT_EN_ES se presentan en la ficha a continuación:

Elementos culturales identificados entre los 100 primeros términos de la lista de palabras clave del subcorpus de literatura traducida del inglés LIT_EN_ES

| Orden en el listado de palabras clave | Palabra clave | Frecuencia | Keyness |
|--|----------------------|-------------------|----------------|
| 7 | Capullo | 231 | 2.668.562 |
| 9 | Cabrón | 215 | 2.483.726 |

| | | | |
|----|-----------|-----|-----------|
| 10 | Begbie | 209 | 2.414.413 |
| 13 | Madre | 930 | 1.974.593 |
| 14 | Puto | 158 | 1.825.250 |
| 24 | Tío | 413 | 1.367.001 |
| 26 | Charlotte | 116 | 1.340.057 |
| 29 | Dolly | 106 | 1.224.535 |
| 33 | Tíos | 94 | 1.085.908 |
| 39 | Jodido | 81 | 935.729 |
| 43 | Pub | 73 | 843.312 |
| 46 | Puta | 204 | 802.902 |
| 49 | Capullos | 67 | 773.998 |
| 52 | Joder | 63 | 727.790 |
| 60 | Cabrones | 54 | 623.820 |
| 64 | Mierda | 169 | 616.373 |
| 69 | Tía | 216 | 592.350 |

| | | | |
|--------------|----------|--------------|---------|
| 70 | Cojones | 51 | 589.163 |
| 84 | Carlson | 45 | 519.850 |
| 91 | Hijoputa | 43 | 496.745 |
| Total | | 3.538 | |

Tabla 34: culturemas clave en LIT_EN_ES

Finalmente se repetirán las mismas tareas para el subcorpus de literatura traducida del italiano LIT_IT_ES siguiendo las tareas presentadas a continuación:

- por medio de la comparación de nuestro subcorpus traducido del italiano con el corpus de referencia CREA, se crea el listado de palabras claves relativo al subcorpus de literatura traducida del inglés LIT_IT_ES;

- una vez obtenida dicha lista, se consideran y se analizan las primeras 100 palabras (que se pueden consultar en el anexo 12 al final del presente estudio), para identificar entre ellas la presencia de culturemas.

A continuación se presenta la ficha con los resultados de análisis idiosincrásico de los culturemas del subcorpus de traducciones del italiano LIT_IT_ES:

Elementos culturales identificados entre los 100 primeros términos de la lista de palabras clave del subcorpus de literatura traducida del italiano LIT_IT_ES

| Orden en el listado de palabras | Palabra clave | Frecuencia | Keyness |
|--|----------------------|-------------------|----------------|
|--|----------------------|-------------------|----------------|

| clave | | | |
|--------------|-------------|--------------|-----------|
| 3 | Abad | 318 | 3.652.765 |
| 10 | Senor | 199 | 2.285.850 |
| 52 | Príncipe | 249 | 748.058 |
| 59 | Aviñón | 56 | 643.254 |
| 66 | Usted | 566 | 610.124 |
| 79 | Palermo | 46 | 528.387 |
| 81 | Padre | 498 | 518.235 |
| 83 | Sicilia | 45 | 516.901 |
| 88 | Café | 43 | 493.927 |
| 92 | Limonada | 42 | 482.441 |
| 98 | Fratricelli | 40 | 459.467 |
| Total | | 2.102 | |

Tabla 35: culturemas claves en LIT_IT_ES

El análisis sigue, como se ha ilustrado en el apartado 3.2.5.2.2 del capítulo de metodología, con la aplicación de la misma metodología comparativa externa a los dos subcorpus de obras originales siguiendo las mismas tareas consignadas antes:

- se crea la lista de palabras clave del subcorpus de obras en lengua inglesa por medio de la comparación de esta con el corpus de referencia BoE;
- una vez obtenida la lista de palabras clave, se consideran las primeras 100 palabras (que se pueden consultar en el anexo 9) y se analiza el peso idiosincrásico de los culturemas en el subcorpus de obras originales inglesas;
- se crea la lista de palabras clave del subcorpus de obras en lengua italiana por medio de la comparación de esto con el corpus de referencia de *Leeds Italian Internet Frequency List*;
- una vez obtenida la lista de palabras clave, se consideran las primeras 100 palabras (véase el anexo 10) y se analiza el peso idiosincrásico de los culturemas en el subcorpus de obras originales italianas.

Los culturemas identificados entre las primeras 100 palabras de los listados de palabras claves de los dos subcorpus de literatura en lengua original se pueden consultar en las fichas mostradas a continuación:

Elementos culturales identificados entre los 100 primeros términos de la lista de palabras clave del subcorpus de obras originales ingleses

| Orden en el listado de palabras clave | Palabra clave | Frecuencia | Keyness |
|--|----------------------|-------------------|----------------|
| 9 | Fuckin | 965 | 13.226.912 |

| | | | |
|--------------|--------|--------------|------------|
| 10 | Ay | 903 | 12.377.100 |
| 15 | Cunt | 590 | 8.086.920 |
| 45 | Don | 659 | 3.710.479 |
| 54 | Lolita | 236 | 3.234.768 |
| 63 | Begbie | 205 | 2.809.862 |
| 91 | Mother | 862 | 1.859.868 |
| Total | | 4.420 | |

Tabla 36: cultuemas clave obras originales en lengua inglés

Elementos culturales identificados entre los 100 primeros términos de la lista de palabras clave del subcorpus de obras originales italianas

| Orden en el listado de palabras clave | Palabra clave | Frecuencia | Keyness |
|--|----------------------|-------------------|----------------|
| 6 | Don | 525 | 785.959 |
| 9 | Dottor | 331 | 566.936 |
| 12 | L'abate | 227 | 452.578 |

| | | | |
|--------------|-----------|--------------|---------|
| 17 | Príncipe | 246 | 371.213 |
| 53 | Concetta | 105 | 209.342 |
| 74 | Padre | 297 | 144.175 |
| 79 | Signorina | 67 | 133.580 |
| 100 | Lisboa | 54 | 107.662 |
| Total | | 1.852 | |

Tabla 37 : culturemas clave obras originales en lengua italiana

En primer lugar es importante precisar que todos los elementos culturales considerados arriba, tanto de los corpus formados por textos traducidos como de los formados por textos originales, tienen un valor de *keyness* superior al 15,13 por lo cual su significación estadística se situaría en el percentil de 99,99 y su valor p sería inferior al 0.0001. Ello quiere decir, como es bien sabido, que la probabilidad de que se cumpla la hipótesis nula (y que los datos se deban a la casualidad) es ínfima, del 0,01%).

Analizando las fichas con los resultados del análisis parece evidente que, en el corpus de obras traducidas, se encuentra un mayor número de culturemas diferentes (15 en el corpus de literatura traducida global LIT_ES, 20 en el subcorpus de literatura traducida del inglés LIT_EN_ES y 11 en el subcorpus de literatura traducida del italiano LIT_IT_ES) con respecto al corpus de obras originales (7 culturemas en las obras originales ingleses y 8 culturemas en las obras originales italianas). Este resultado podría ser un indicio de la importante naturaleza cultural de la literatura traducida, en la

cual se encuentra un mayor número de elementos culturales: como los culturemas son definatorios de una cultura, en las obras traducidas se utilizan y se respetan más que en las lenguas originales y que parecen apuntar a una mayor variación de culturemas en traducción que en originales. El número total de culturemas encontrados entre las primeras 100 palabras de la lista de palabras clave, en cambio, es superior en el corpus de obras en lengua original (6.272) que en el corpus de literatura traducida LIT_ES (2.558 culturemas en total). Esto podría deberse a que muchos de los culturemas empleados en la literatura original son moneda común en la lengua cotidiana mientras que los culturemas empleados en la literatura traducida, al estar vinculados a la cultura de partida, resultan más exóticos y por ende, clave frente a al uso de la lengua cotidiana.

Del análisis se desprenden también otros fenómenos interesantes. Por un lado, la presencia de algunos elementos culturales con la misma frecuencia tanto en el corpus de literatura traducida global LIT_ES como en uno de los subcorpus, hecho que implica la utilización de dicho elementos solo en uno de los idiomas analizados. Así que los términos *begbie* (encontrado 209 veces), *puto* (encontrado 158 veces) y *Charlotte* (encontrado 116 veces) tienen la misma frecuencia tanto en el corpus de literatura LIT_ES como en el subcorpus inglés-español LIT_EN_ES y por consecuencia resultan ser términos presentes solo en dicho subcorpus. Lo mismo resulta por las palabras *abad* (encontrado 318 veces) y *senor* (encontrado 199 veces) en el corpus de traducciones del italiano LIT_IT_ES.

Por otro lado, comparando las fichas de los subcorpus de literatura traducida LIT_EN_ES y LIT_IT_ES con las de los respectivos subcorpus de literatura en lengua original, se identifican algunos elementos comunes:

| Corpus literatura traducida | Subcorpus | Frecuencia | Corpus literatura en lengua original correspondiente | Frecuencia |
|------------------------------------|------------------|-------------------|---|-------------------|
| Madre | LIT_EN_ES | 930 | Mother | 862 |
| Begbie | LIT_EN_ES | 209 | Begbie | 205 |
| Príncipe | LIT_IT_ES | 249 | Principe | 246 |
| Padre | LIT_IT_ES | 498 | Padre | 297 |

Tabla 38: comparación cultuemas clave literatura traducida y obras originales

Comparando el número de frecuencia de los elementos culturales detectados entre las 100 primeras palabras claves de los subcorpus de literatura traducida LIT_EN_ES y LIT_IT_ES con los encontrados en los sobcorpus de obras en lengua original, se observa que, en literatura traducida, el número de frecuencia de cada cultuema es mayor que en la literatura en lengua original. Este dato demuestra una vez más la mayor utilización de los cultuemas en la literatura traducida que en la literatura en lengua original.

Por lo que se refiere a las categorías de cultuemas encontradas en las listas de palabras clave, no parece haber un patrón específico para ninguno de los dos corpus analizados LIT_EN_ES y LIT_IT_ES.

5 CONCLUSIONES

5.1 Recapitulación del trabajo, de las hipótesis y de los objetivos

En el presente trabajo y como base para ulteriores estudios futuros (en particular para la tesis doctoral), nos hemos propuesto corroborar o refutar la hipótesis general, presentada en el apartado 3.2.1:

Una metodología basada en corpus facilita el proceso de identificación de áreas proclives al uso de culturemas y propicia el análisis cuantitativo y cualitativo del tratamiento de los culturemas identificados.

Junto con esta hipótesis, partíamos de los objetivos principales explicados también en el apartado 3.2.1, que pueden sintetizarse en: obtener un conocimiento más profundo del corpus utilizado que pueda ser útil como base para ulteriores investigaciones, demostrar la operatividad de las clasificaciones propuestas para los culturemas y para las técnicas de traducción utilizadas, y demostrar la efectividad del método utilizado para que se pueda emplear en trabajos futuros, específicamente en una tesis doctoral.

Con estos fines en mente, el presente estudio se ha dividido en dos partes: una primera parte teórica y una segunda parte de análisis. En la primera parte (i) se han presentado las nociones de cultura y culturemas como objetos de estudio, analizándose las definiciones y las clasificaciones propuestas por algunos de los autores del ámbito y explicándose la diferencia entre marcas culturales y pseudo-marcas culturales (cap. 1); (ii) se han revisado las posibles técnicas para traducir los elementos culturales y trasponerlos de una cultura a otra, desde diversos puntos de vista (cap.2); (iii) se ha presentado la lingüística de corpus como metodología, resumiendo sus orígenes, sus objetivos, las ventajas y los límites de su aplicación y las herramientas que se utilizan

para desarrollar análisis basados en corpus (cap. 3, apartado 3.1). En esta primera parte teórica, por tanto, se ha iniciado al lector en los aspectos básicos del presente estudio resumiendo los conceptos fundamentales en los que este se basa y, a partir de la teoría, se ha propuesto el marco teórico que se utilizará durante el análisis para la clasificación de los culturemas y de sus técnicas de traducción (apartados 1.3 y 2.5).

En la segunda parte de análisis se han utilizado los conocimientos explicados antes para desarrollar un estudio propio sobre la naturaleza y la traducción de los elementos culturales de un archivo de textos de literatura traducida (LIT_ENIT_ES). Así se han presentado los corpus utilizados (cap. 3, apartado 3.2.2), se ha explicado la metodología adoptada (apartado 3.2) y se ha procedido con el análisis descriptivo-comparativo (capítulo 4) con el fin de corroborar la tesis presentada y cumplir los objetivos propuestos en el apartado 3.2.1. Dicho análisis se ha organizado en dos etapas cada una de ellas dividida, a su vez, en dos fases. La primera etapa contiene el análisis descriptivo, con el fin de obtener un conocimiento más profundo y detallado del corpus utilizado. Esta etapa descriptiva se divide en dos fases: la fase de identificación y clasificación de culturemas en categorías culturales precedentemente establecidas (véase el apartado 1.3) y la fase de identificación de los culturemas más frecuentes (con una ocurrencia mayor o igual a 50) en el corpus analizado (LIT_ES). En la segunda etapa se desarrolla el análisis comparativo también siguiendo dos fases. En la primera fase se acomete un análisis comparativo interno, cotejando los textos traducidos con las respectivas obras originales, mediante los alineados de los textos, para identificar las técnicas de traducción de los culturemas y clasificarlas según las categorías establecidas en el apartado 2.5. En la segunda fase se realiza un análisis comparativo externo, cruzando los corpus utilizados en el presente estudio con corpus de referencia, para

obtener listas de palabras claves y, en consecuencia, comparar el peso idiosincrásico de los culturemas identificados en las obras traducidas con el peso idiosincrásico que presentan los elementos culturales de las obras originales.

5.2 Interpretación de los resultados

La interpretación de los resultados obtenidos se presenta en este apartado siguiendo la estructura del análisis que se acaba de explicar. Como se ha aclarado arriba, la primera etapa del análisis, la etapa descriptiva, se ha dividido en dos fases. En la primera, se han identificado los culturemas y se han clasificado siguiendo las categorías culturales propuestas anteriormente (apartado 1.3). De esta fase se desprende que las categorías de culturemas más utilizados en nuestro corpus de literatura traducida son las de la cultura social, histórica y de folclore, cultura geográfica y natural, nombres propios, cultura lingüística y cultura religiosa. Asimismo las subcategorías de culturemas más utilizadas resultan ser la de los topónimos y referencias a lugares específicos, la de flora y fauna (ambas incluidas en la categoría de la cultura geográfica y natural), la de la organización y papeles sociales (como parte de la categoría cultural de la cultura social, histórica y de folclore) y la del nivel de formalidad (perteneciente a la categoría de la cultura lingüística) (véase los resultados relativos a esta fase en el apartado 4.1.1).

De los resultados de esta primera fase descriptiva se puede concluir que el método empleado para la identificación de los culturemas resulta operativo y eficaz y, por esta misma razón, se puede utilizar en investigaciones futuras. Sin embargo, parece necesario añadir a este método las debidas mejoras ya que, así como se ha presentado, se prevé un proceso largo, con un amplio número de elementos que deben analizarse

manualmente y, por consiguiente, con una alta probabilidad de cometer errores o ralentizar el ritmo de las tareas analíticas.

Asimismo la clasificación adoptada para categorizar los culturemas es perfectamente operativa y aplicable a futuros trabajos basados en corpus. No obstante, nuestro estudio parece sugerir la pertinencia de algunos retoques en el marco teórico. El trabajo descriptivo con el archivo LIT_ES refleja la necesidad de agregar una categoría de culturemas no contemplada en la bibliografía especializada (nombres propios) y hace evidente la redundancia de la categoría interferencia cultural (que se convertirá en la técnica de traducción de inequivalencia). Estos resultados corroboran en parte la hipótesis general presentada en el apartado 3.2.1. Por lo demás, los deslizamientos de categorías de culturemas, determinados por el contexto e identificados en esta fase del análisis, reflejan la necesidad de combinar el método cuantitativo de los análisis automatizados con el método cualitativo del análisis manual para rectificar posibles errores.

En la segunda fase de la etapa descriptiva del estudio, se han identificado y clasificado los culturemas más frecuentes del corpus analizado (con una ocurrencia igual o mayor de 50). De esta fase se desprende, por un lado, que la mayoría de los culturemas más frecuentes pertenecen a las mismas categorías de los culturemas más utilizados en el corpus, que ya se han identificado en la fase precedente del análisis. De este modo las subcategorías más frecuentes resultan ser la de la familia, de la organización y papeles sociales (que forman parte de la categoría de la cultura social, histórica y de folclore), la del nivel de formalidad (parte de la categoría de la cultura lingüística), y la categoría de la cultura religiosa. Por otro lado, no se han identificado, entre los culturemas más frecuentes, aquellos elementos culturales que pertenecen a la categoría de la cultura

geográfica y natural, categoría identificada entre las más utilizadas en el análisis previo. Este resultado demuestra que dichas categorías incluyen muchos *culturemas* diferentes (*types*) pero que la ocurrencia de cada uno no es muy elevada.

De los resultados de esta segunda fase descriptiva puede concluirse que la utilización conjunta de estas dos fases del análisis descriptivo (la identificación de los *culturemas* utilizados en el corpus y la de los más frecuentes) es útil tanto desde un punto de vista cualitativo, porque permite identificar un abanico de *culturemas* más amplio, como desde un punto de vista cuantitativo, ya que permite localizar aquellos elementos culturales con un impacto más recurrente en todo el corpus. Ambos datos son importantes a la hora de describir un archivo de corpus y, como proporcionan un conocimiento más profundo del corpus, han permitido cumplir el primero de los objetivos establecidos.

También el análisis comparativo se ha dividido en dos fases, una interna y otra externa. Tras el análisis comparativo interno, dedicado a la identificación de las técnicas traductoras por medio del contraste de textos meta y fuente y a la sucesiva clasificaciones de dichas técnicas en categorías, se observa que las técnicas de traducción más utilizadas son la traducción literal, la transposición, el préstamo y la naturalización, todas técnicas exotizantes, según el continuum entre exotismo y adaptación cultural propuesto y explicado en el capítulo 2.

Los resultados de esta primera fase comparativa permiten corroborar totalmente la hipótesis principal y cumplir los objetivos, demostrando la operatividad y la funcionalidad tanto de la metodología de identificación de técnicas de traducción y *culturemas* adoptada como de la clasificación de técnicas de traducción propuestas. No

obstante, el estudio parece sugerir la pertinencia de algunos retoques en el marco teórico. El trabajo comparativo con el archivo LIT_ES refleja la necesidad de agregar la categoría de inequivalencia entre las técnicas de traducción (que viene a sustituir a la interferencia cultural de las categorías de culturemas). El estudio, por tanto, muestra que resulta difícil detectar, de forma descriptiva y monolingüe, la interferencia cultural y que esta más bien se localiza en trabajos comparativos bi/multilingües. Asimismo, tras este análisis comparativo interno, la categoría de la omisión ha resultado no ser operativa por una carencia metodológica: para identificar los casos de omisión de culturemas en el trasvase de la literatura en lengua original a la literatura traducida, sería necesario un análisis en sentido inverso, o sea habría que identificar en primer lugar los elementos culturales en las obras originales y luego, por medio de programas como *Intertext Editor*, localizarlos en el alineado correspondiente (y, sin embargo, nosotros hemos partido de los textos meta, que hemos contrastado con sus respectivos originales). Además, la comparación de traducciones y originales (desarrollada con *Intertext Editor*), ha puesto en evidencia, como ya se ha explicado más detalladamente en el apartado 4.2.1, algunas disparidades por la ausencia, en algunos casos, de extractos de partida o de llegada. Este fenómeno remarca la necesidad de revisar manualmente los corpus analizados automáticamente para rectificar y corregir errores y enfatiza la consiguiente conveniencia de conjugar métodos cuantitativos y cualitativos para alcanzar resultados fiables.

El objetivo principal del análisis comparativo externo ha sido la identificación del peso idiosincrásico de los culturemas de la literatura traducida con respecto a los de la literatura original. De los resultados se desprende una mayor variedad de culturemas en la literatura traducida (establecida por la mayor cantidad de *types* y la menor cantidad de

tokens totales), que impulsa a considerar a los culturemas, que por naturaleza son elementos definitorios de una cultura específica, como elementos clave de la literatura traducida. Los resultados demuestran también que el mismo culturema se utiliza más en el corpus de literatura traducida que en el corpus de obras en lengua original, hecho que sugiere una mayor explicitación y una más amplia utilización de los culturemas en la literatura traducida, que por lo tanto se convierte en un importante medio de exportación/importación de culturemas; estos últimos adquieren una relevancia evidente en la literatura y en la cultura receptora (para los resultados detallados del análisis externo véase el apartado 4.2.2).

En conclusión, gracias al análisis descriptivo-comparativo se ha obtenido un mayor conocimiento del corpus analizado y del objeto de investigación del presente estudio, cumpliendo con los objetivos de partida. Además, la metodología de investigación y las clasificaciones propuestas han resultado útiles y operativas y es por esto por lo que, para investigaciones futura, proponemos replicar la estructura metodológica descriptiva-comparativa del presente estudio, corroborando así la hipótesis principal de inicio. Finalmente, ha quedado claro que el análisis cuantitativo es complementario al análisis cualitativo, por lo cual es imprescindible la unión de métodos automatizados, como el análisis de corpus, con métodos manuales que corrigen y perfeccionan los resultados obtenidos.

5.3 Propuestas para investigaciones futuras

Los resultados de los análisis descriptivos y comparativos desarrollados en el capítulo 4 se prestan a interpretaciones que pueden servir de inspiración para investigaciones

futuras. El análisis de los culturemas y su tratamiento en corpus de literatura traducida se puede continuar en diferentes direcciones.

Así, una vez ampliado el corpus por número y criterios de elección de las obras para aumentar su representatividad, se puede investigar sobre la comparación entre los subcorpus de literatura traducida del inglés y del italiano para identificar elementos comunes o divergentes y relacionar los resultados de dicha comparación con otros temas.

Al hilo de lo dicho arriba, se podría investigar sobre la influencia del par de idiomas analizado sobre la naturaleza, la clasificación y el tratamiento de los elementos culturales, siguiendo la hipótesis siguiente: la lengua de origen influye decisivamente en el tipo (categorías y subcategorías) de culturemas que se utilizan en las literaturas originales y traducidas, con independencia de la temática de estas.

Otra posibilidad podría ser analizar el tratamiento de los elementos culturales traducidos desde diferentes idiomas para estudiarlos en relación con los universales lingüísticos, por ejemplo basándose en la hipótesis siguiente: algunos universales de traducción son constructos cuestionables (quizá vigentes en trasvases en los que el destino es el producto en inglés). Las categorías y subcategorías de las técnicas que se aplican en las traducciones dependen de la relación entre lengua origen y meta y del momento del trasvase.

Y sin embargo, de los resultados del análisis comparativo externo resulta que, en la literatura traducida los culturemas clave aparecen en mayor medida; es decir, aparecen más veces. Así que, por su mayor utilización, , la literatura traducida parece ser más explícita que la literatura en lengua original, siguiendo así el universal de explicitación.

Otras posibilidades de investigación incluirían la comparación de un archivo LIT_ENIT_ES mejorado (y de sus elementos culturales identificados y clasificados), con otros corpus para establecer afinidades y divergencias. De hecho, esta es la opción elegida para una futura tesis doctoral, en la que nos proponemos comparar un corpus de literatura traducida con otro de literatura de viaje, con el fin de establecer la semejanza entre ambos. Este estudio futuro se basa en teorías según las cuales la traducción puede equipararse a un viaje. El trabajo, por tanto, se identifica con los postulados, entre otros, de Harish Trivedi (2005:sec. 2), que considera la traducción *the process and condition of human migracy*, y de Ovidi Carbonell (en comunicación personal), que compara la traducción con una experiencia viajera (en palabras de Carbonell, un *viaje transcultural*) en la cual se transmite el conocimiento del “otro”.

ANEXOS

ANEXO 1

STOP LIST UTILIZADA PARA ELIMINAR LAS PALABRAS GRAMATICALES DE LAS LISTAS DE PALABRAS ANALIZADAS

| | | |
|----------|----------|----------|
| al | no | has |
| aunque | como | ha |
| con | o | hemos |
| de | porque | habeis |
| del | porqué | han |
| desde | qué | había |
| el | un | habías |
| ella | una | habíamos |
| ellas | unos | habían |
| ellos | unas | habiaís |
| en | mi | hubo |
| hacia | tu | es |
| hasta | su | está |
| la | mis | estoy |
| las | tus | soy |
| los | sus | eres |
| mas | nuestro | somos |
| nosotros | vuestro | sois |
| para | nuestros | son |
| pero | vuestros | estás |
| por | mio | estamos |
| que | tuyo | estais |
| sin | suyo | están |
| vosotros | mios | estuvo |
| yo | tuyos | estuve |
| él | suyos | |
| y | he | |

ANEXO 2

ELEMENTOS CULTURALES ELIMINADOS DESPUÉS DEL ANÁLISIS DE LOS CULTUREMAS EN SU CONTEXTO

| Elemento | Razón de la eliminación del elemento del listado de culturemas |
|-----------------|--|
| Abastecimientos | Por no tener naturaleza cultural |
| Aberwerfen | Palabra contenida en una frase en alemán |
| Ackley | Nombre propio de protagonista |
| Adae | Palabra contenida en una frase en latín |
| Adaeu | Palabra contenida en una frase en latín |
| Adamnano | Palabra contenida en una frase en latín |
| Adamítico | Por no tener naturaleza cultural (utilizado como adjetivo) |
| Adictos | Por no tener naturaleza cultural (utilizado como adjetivo) |
| Affirmavit | Palabra contenida en una frase en latín |
| Alberti | Palabra contenida en una frase en latín |
| Albus | Palabra contenida en una frase en latín |
| Alchernicus | Palabra contenida en una frase en latín |
| Alfonso | Nombre propio de protagonista |
| Alpino | Por no tener naturaleza cultural (utilizado como adjetivo) |
| And | Por ser una palabra gramatical |
| Anni | Palabra contenida en una frase en latín |
| Anno | Palabra contenida en una frase en latín |
| Argentina | Por no tener naturaleza cultural (utilizado como adjetivo) |
| Attention | Palabra contenida en una frase en francés |
| Augustinianorum | Palabra contenida en una frase en latín |
| Augusto | Por no tener naturaleza cultural (utilizado como adjetivo) |
| Aut | Palabra contenida en una frase en latín |
| Autre | Palabra contenida en una frase en latín |
| Aux | Palabra contenida en una frase en latín y en francés |
| Ave | Por no tener naturaleza cultural (utilizado como sustantivo en el sentido de pájaro) |
| Avec | Palabra contenida en una frase en francés |
| Babylonia | Palabra contenida en una frase en latín |
| Baby | Palabra contenida en una frase en inglés |
| Baff | Onomatopeya |
| Bang | Onomatopeya |
| Bath | Palabra contenida en una frase en inglés |
| Bayeta | Por no tener naturaleza cultural (utilizado como sustantivo) |
| Bretón | Por no tener naturaleza cultural (utilizado como adjetivo) |
| Bribón | Por no tener naturaleza cultural (utilizado como adjetivo) |
| Canónico | Por no tener naturaleza cultural (utilizado como adjetivo) |
| Capullo | Por no tener naturaleza cultural (utilizado con sentido de envoltura) |
| Charta | Palabra contenida en una frase en latín |

| | |
|-------------|---|
| Chinky | Nombre propio de protagonista |
| Clic | Onomatopeya |
| Concepción | Por no tener naturaleza cultural (utilizado sin relación a la religión) |
| Curado | Por no tener naturaleza cultural (utilizado como adjetivo) |
| Dedal | Por no tener naturaleza cultural (utilizado como sustantivo) |
| Deficiencia | Por no tener naturaleza cultural (utilizado en el sentido de falta) |
| Deficiente | Por no tener naturaleza cultural (utilizado en el sentido de incompletez) |
| Dondequiera | Por no tener naturaleza cultural |
| Florida | Por no tener naturaleza cultural (utilizado como adjetivo) |
| Garage | Palabra contenida en una frase en inglés |
| Garaje | Por no tener naturaleza cultural (utilizado como sustantivo) |
| Garden | Palabra contenida en una frase en inglés |
| Indias | Por no tener naturaleza cultural (utilizado como adjetivo) |
| Lolitas | Nombre propio de protagonista |
| Magenta | Por no tener naturaleza cultural (utilizado para definir un color) |
| Marketing | Palabra contenida en una frase en inglés |
| Nazca | Por no tener naturaleza cultural (utilizado como verbo) |
| Nevada | Por no tener naturaleza cultural (utilizado como adjetivo) |
| Pipí | Onomatopeya |
| Senatores | Por no tener naturaleza cultural |
| Swing | Onomatopeya |
| Taller | Por no tener naturaleza cultural (utilizado en el sentido de lugar) |
| Trans | Por no tener naturaleza cultural (utilizado como prefijo) |

ANEXO 3

PALABRAS CON FRECUENCIA MAYOR O IGUAL A 50 DEL CORPUS LIT_ES

| | | | | | | | | |
|----|-------|----------|----|-----|-----------|-----|-----|-----------|
| 1 | 41707 | de | 40 | 931 | siempre | 82 | 658 | sido |
| 2 | 10595 | se | | | | 83 | 657 | parte |
| 3 | 7884 | me | 41 | 927 | algo | 84 | 649 | sé |
| 4 | 7236 | lo | 42 | 918 | otro | 85 | 645 | dije |
| 5 | 4123 | le | 43 | 917 | otra | 86 | 636 | aún |
| 6 | 3367 | era | 44 | 916 | poco | 87 | 634 | tal |
| 7 | 3265 | más | 45 | 896 | uno | 88 | 632 | mucho |
| 8 | 3083 | si | 46 | 889 | ser | 89 | 621 | dice |
| 9 | 2525 | cuando | 47 | 869 | fue | 90 | 615 | aquella |
| 10 | 2177 | dijo | 48 | 868 | este | 91 | 611 | menos |
| 11 | 2050 | estaba | 49 | 849 | hay | 92 | 609 | libro |
| 12 | 2022 | ya | 50 | 846 | cosas | 93 | 602 | tanto |
| 13 | 1928 | sobre | 51 | 846 | mí | 94 | 596 | hace |
| 14 | 1916 | todo | 52 | 833 | esa | 95 | 594 | puede |
| 15 | 1736 | te | 53 | 811 | padre | 96 | 591 | hombre |
| 16 | 1476 | muy | 54 | 806 | antes | | | |
| 17 | 1417 | sólo | 55 | 801 | casa | 97 | 585 | durante |
| 18 | 1407 | vez | 56 | 800 | hacer | | | |
| 19 | 1351 | eso | 57 | 787 | donde | 98 | 584 | mejor |
| 20 | 1311 | todos | 58 | 778 | cómo | 99 | 580 | eran |
| 21 | 1305 | dos | 59 | 776 | aquí | 100 | 575 | otros |
| 22 | 1291 | ahora | 60 | 772 | ojos | 101 | 567 | esto |
| 23 | 1273 | después | 61 | 770 | verdad | 102 | 560 | allí |
| | | | 62 | 764 | mismo | 103 | 553 | todas |
| 24 | 1241 | también | 63 | 761 | años | 104 | 551 | dicho |
| | | | 64 | 751 | usted | 105 | 533 | gran |
| 25 | 1205 | nada | 65 | 746 | guillermo | 106 | 532 | tú |
| 26 | 1194 | bien | 66 | 743 | decir | 107 | 529 | |
| 27 | 1185 | nos | 67 | 733 | nunca | | | demasiado |
| 28 | 1121 | entre | 68 | 728 | tiempo | 108 | 527 | casi |
| 29 | 1113 | tenía | 69 | 718 | día | 109 | 522 | les |
| 30 | 1092 | así | 70 | 715 | cada | 110 | 516 | voz |
| 31 | 1053 | mientras | 71 | 713 | tiene | 111 | 514 | cosa |
| | | | 72 | 699 | momento | 112 | 513 | mujer |
| 32 | 1023 | ese | 73 | 689 | ver | 113 | 506 | veces |
| 33 | 1016 | tan | 74 | 688 | cabeza | 114 | 504 | parece |
| 34 | 988 | madre | 75 | 687 | luego | 115 | 504 | sino |
| 35 | 986 | entonces | 76 | 686 | hecho | 116 | 503 | modo |
| | | | 77 | 684 | mano | 117 | 498 | lado |
| 36 | 964 | ni | 78 | 680 | vida | 118 | 498 | toda |
| 37 | 954 | sí | 79 | 671 | podía | 119 | 488 | bajo |
| 38 | 942 | pereira | 80 | 658 | aquel | 120 | 484 | dios |
| 39 | 937 | esta | 81 | 658 | noche | 121 | 482 | nadie |

| | | | | | | | | |
|-----|-----|----------|-----|------------|-----------|-----|------------|----------|
| 122 | 480 | mundo | 168 | 362 | mañana | 209 | 306 | cama |
| 123 | 478 | pues | | | | 210 | 303 | rostro |
| 124 | 475 | quien | 169 | 360 | cuerpo | 211 | 301 | dentro |
| 125 | 474 | haber | 170 | 360 | incluso | 212 | 297 | iglesia |
| 126 | 474 | tres | 171 | 358 | saber | 213 | 296 | caso |
| 127 | 472 | fuera | 172 | 355 | don | 214 | 294 | hombres |
| 128 | 468 | estaban | 173 | 355 | hubiera | | | |
| | | | | | | 215 | 293 | través |
| 129 | 468 | quizá | 174 | 353 | cual | 216 | 292 | joven |
| 130 | 465 | puerta | 175 | 351 | fin | 217 | 291 | alguna |
| 131 | 457 | quería | 176 | 351 | mal | 218 | 289 | además |
| 132 | 453 | tío | 177 | 351 | pronto | | | |
| 133 | 451 | señor | 178 | 350 | nuestra | 219 | 289 | muerte |
| 134 | 450 | tengo | 179 | 346 | lennie | 220 | 288 | dónde |
| 135 | 449 | george | 180 | 345 | iba | 221 | 286 | claro |
| 136 | 444 | quién | 181 | 344 | junto | 222 | 285 | cuarto |
| 137 | 439 | hacía | 182 | 343 | cierto | 223 | 284 | cuanto |
| 138 | 424 | cuenta | 183 | 343 | | 224 | 279 | tipo |
| 139 | 423 | nuevo | | habitación | | 225 | 277 | estos |
| 140 | 420 | hablar | 184 | 343 | misma | 226 | 277 | nombre |
| 141 | 416 | sea | 185 | 338 | estar | | | |
| 142 | 409 | bueno | 186 | 337 | vi | 227 | 276 | siquiera |
| 143 | 408 | habría | 187 | 337 | viejo | | | |
| 144 | 408 | sabía | 188 | 336 | estado | 228 | 274 | lolita |
| 145 | 405 | creo | 189 | 335 | quiero | 229 | 274 | quiere |
| 146 | 403 | esos | 190 | 334 | gente | 230 | 274 | volvió |
| 147 | 402 | contra | 191 | 334 | primera | 231 | 272 | luz |
| 148 | 402 | manos | | | | 232 | 270 | boca |
| 149 | 400 | alguien | 192 | 332 | lugar | 233 | 270 | persona |
| 150 | 396 | visto | 193 | 330 | demás | | | |
| 151 | 395 | palabras | 194 | 330 | tener | 234 | 270 | preguntó |
| | | | 195 | 329 | muchos | 235 | 269 | duda |
| 152 | 394 | libros | | | | 236 | 268 | dio |
| 153 | 392 | parecía | 196 | 329 | podría | 237 | 267 | pasado |
| 154 | 391 | hizo | 197 | 328 | esas | 238 | 265 | boy |
| 155 | 391 | sabe | 198 | 326 | bastante | 239 | 265 | pie |
| 156 | 390 | tarde | | | | 240 | 262 | cuatro |
| 157 | 388 | ante | 199 | 325 | cualquier | 241 | 261 | encima |
| 158 | 382 | punto | 200 | 325 | decía | 242 | 261 | hora |
| 159 | 377 | da | 201 | 325 | días | 243 | 260 | algunos |
| 160 | 376 | mesa | 202 | 324 | todavía | | | |
| 161 | 371 | puedo | 203 | 318 | abad | 244 | 259 | |
| 162 | 370 | ir | 204 | 313 | agua | | biblioteca | |
| 163 | 366 | otras | 205 | 312 | digo | 245 | 259 | muchas |
| 164 | 365 | sabes | 206 | 312 | embargo | | | |
| 165 | 364 | cara | | | | 246 | 257 | estas |
| 166 | 363 | algún | 207 | 310 | manera | 247 | 256 | tras |
| 167 | 363 | va | 208 | 309 | buena | 248 | 254 | idea |

| | | | | | | | | |
|-----|-----|----------|-----|-----|-----------|-----|-----|----------|
| 249 | 254 | pelo | 291 | 221 | sol | 331 | 202 | camino |
| 250 | 254 | poder | 292 | 220 | pregunté | 332 | 202 | daba |
| 251 | 254 | puso | | | | 333 | 202 | detrás |
| 252 | 251 | príncipe | 293 | 219 | pensar | 334 | 202 | personas |
| | | | 294 | 219 | renton | | | |
| 253 | 250 | puesto | 295 | 217 | cabrón | 335 | 202 | pies |
| 254 | 249 | tienes | 296 | 215 | último | 336 | 200 | tierra |
| 255 | 248 | debía | 297 | 213 | alrededor | 337 | 199 | largo |
| 256 | 248 | forma | | | | 338 | 199 | segundo |
| 257 | 248 | hija | 298 | 213 | final | | | |
| 258 | 247 | dinero | 299 | 213 | llevaba | 339 | 199 | senor |
| 259 | 246 | miró | 300 | 213 | mirada | 340 | 198 | familia |
| 260 | 246 | solo | 301 | 213 | seguida | 341 | 198 | par |
| 261 | 244 | historia | | | | 342 | 198 | primero |
| | | | 302 | 213 | sitio | | | |
| 262 | 244 | sería | 303 | 212 | ciudad | 343 | 197 | atrás |
| 263 | 243 | palabra | 304 | 212 | frente | 344 | 197 | llegar |
| 264 | 243 | voy | 305 | 212 | oh | 345 | 195 | mayor |
| 265 | 242 | medio | 306 | 212 | sangre | 346 | 195 | pequeña |
| 266 | 241 | amor | 307 | 211 | dado | | | |
| 267 | 241 | cinco | 308 | 211 | muerto | 347 | 195 | spud |
| 268 | 240 | orden | 309 | 211 | mujeres | 348 | 194 | dar |
| 269 | 238 | debe | | | | 349 | 194 | mira |
| 270 | 238 | recuerdo | 310 | 211 | puta | 350 | 194 | ninguna |
| | | | 311 | 211 | seguro | | | |
| 271 | 238 | sentido | 312 | 210 | hijo | 351 | 194 | nueva |
| 272 | 238 | sick | 313 | 210 | leer | 352 | 194 | salir |
| 273 | 235 | pobre | 314 | 210 | niña | 353 | 194 | ti |
| 274 | 234 | aire | 315 | 210 | trabajo | 354 | 193 | calle |
| 275 | 234 | cambio | 316 | 209 | begbie | 355 | 193 | cerca |
| 276 | 234 | doctor | 317 | 209 | monteiro | 356 | 193 | señora |
| 277 | 233 | buen | 318 | 209 | puedes | 357 | 193 | tenido |
| 278 | 233 | habia | 319 | 209 | suelo | 358 | 192 | hoy |
| 279 | 232 | capullo | 320 | 209 | tienen | 359 | 190 | acaso |
| 280 | 231 | tancredi | 321 | 208 | ahí | 360 | 190 | color |
| | | | 322 | 208 | hubiese | 361 | 190 | tampoco |
| 281 | 230 | aquellos | | | | | | |
| | | | 323 | 208 | realidad | 362 | 189 | abajo |
| 282 | 229 | vamos | | | | 363 | 187 | abadía |
| 283 | 228 | cocina | 324 | 205 | monjes | 364 | 187 | dejar |
| 284 | 226 | paso | 325 | 204 | amigo | 365 | 187 | hermano |
| 285 | 226 | tía | 326 | 204 | eh | | | |
| 286 | 225 | negro | 327 | 204 | quieres | 366 | 187 | razón |
| 287 | 225 | os | 328 | 203 | gaetano | 367 | 187 | sentía |
| 288 | 223 | éste | | | | 368 | 187 | silencio |
| 289 | 222 | grandes | 329 | 203 | pregunto | | | |
| | | | | | | 369 | 187 | vista |
| 290 | 221 | rossi | 330 | 203 | primer | 370 | 187 | único |

| | | | | | | | | |
|-----|-----|--------------|-----|-----|------------|-----|-----|-----------|
| 371 | 186 | haya | 409 | 171 | querido | 446 | 156 | mirando |
| 372 | 186 | llegado | | | | | | |
| 373 | 186 | pensé | 410 | 170 | aspecto | 447 | 155 | especie |
| 374 | 185 | chica | | | | 448 | 155 | esposa |
| 375 | 185 | hotel | 411 | 170 | ningún | 449 | 154 | dedos |
| 376 | 185 | posible | 412 | 169 | diez | 450 | 154 | importa |
| 377 | 185 | según | 413 | 169 | labios | | | |
| 378 | 185 | siguiente | 414 | 169 | novela | 451 | 154 | sensación |
| 379 | 184 | año | 415 | 168 | clase | 452 | 153 | padres |
| 380 | 184 | blanco | 416 | 168 | conmigo | 453 | 151 | piernas |
| 381 | 184 | pasar | | | | 454 | 151 | será |
| 382 | 184 | propia | 417 | 168 | | 455 | 151 | tuvo |
| 383 | 184 | sostiene | | | malaquías | 456 | 150 | cueello |
| | | | 418 | 168 | tomar | 457 | 150 | delante |
| 384 | 183 | corazón | 419 | 168 | tono | 458 | 150 | tenemos |
| | | | 420 | 167 | ve | | | |
| 385 | 182 | expresión | 421 | 167 | volver | 459 | 149 | fuese |
| 386 | 182 | fondo | 422 | 165 | acuerdo | 460 | 149 | llama |
| 387 | 182 | horas | | | | 461 | 149 | pregunta |
| 388 | 181 | gracias | 423 | 164 | fuerza | | | |
| 389 | 181 | veía | 424 | 162 | cielo | 462 | 148 | carne |
| 390 | 180 | ventana | 425 | 162 | fabrizio | 463 | 148 | diciendo |
| | | | | | | | | |
| 391 | 179 | haciendo | 426 | 162 | maestro | 464 | 148 | gusta |
| 392 | 178 | brazo | | | | 465 | 148 | haze |
| 393 | 178 | ido | 427 | 162 | seguí | 466 | 148 | instante |
| 394 | 177 | empezó | 428 | 162 | sentado | | | |
| | | | | | | 467 | 147 | ah |
| 395 | 177 | mente | 429 | 161 | debería | 468 | 147 | última |
| 396 | 176 | apenas | 430 | 161 | sonrisa | 469 | 146 | cierta |
| 397 | 176 | siento | 431 | 160 | | 470 | 146 | di |
| 398 | 175 | | | | berengario | 471 | 146 | escribir |
| | | realmente | 432 | 160 | ello | | | |
| 399 | 174 | | 433 | 159 | allá | 472 | 146 | san |
| | | precisamente | 434 | 159 | jamás | 473 | 146 | tenían |
| 400 | 173 | mierda | 435 | 159 | miedo | 474 | 145 | director |
| 401 | 173 | | 436 | 159 | mío | | | |
| | | muchacha | 437 | 159 | pareció | 475 | 145 | piso |
| 402 | 172 | algunas | 438 | 159 | rato | 476 | 145 | pueden |
| | | | 439 | 158 | puto | 477 | 145 | sala |
| 403 | 172 | ejemplo | 440 | 157 | alma | 478 | 144 | edad |
| | | | 441 | 157 | brazos | 479 | 144 | hacerlo |
| 404 | 172 | lejos | 442 | 157 | entrar | 480 | 144 | pared |
| 405 | 172 | lengua | 443 | 157 | pensaba | 481 | 144 | piel |
| 406 | 172 | oído | | | | 482 | 144 | páginas |
| 407 | 172 | papa | 444 | 156 | hablaba | | | |
| 408 | 172 | veo | | | | 483 | 143 | digamos |
| | | | 445 | 156 | lector | | | |

| | | | | | | | | |
|-----|-----|--------------|-----|-----|---------------|-----|-----|------------|
| 484 | 143 | jorge | 523 | 134 | semana | 560 | 125 | quienes |
| 485 | 143 | pasa | | | | | | |
| 486 | 143 | perro | 524 | 134 | severino | 561 | 125 | señorita |
| 487 | 142 | caballo | | | | | | |
| 488 | 142 | | 525 | 133 | pesar | 562 | 125 | venancio |
| | | respondio | 526 | 132 | angelica | 563 | 124 | carta |
| 489 | 142 | vino | | | | 564 | 124 | curley |
| 490 | 141 | capaz | 527 | 132 | arriba | 565 | 124 | dejó |
| 491 | 141 | cardoso | 528 | 132 | et | 566 | 124 | |
| | | | 529 | 132 | vio | | | encuentro |
| 492 | 141 | encontrar | 530 | 131 | alto | 567 | 124 | escalera |
| 493 | 141 | juego | 531 | 131 | hablando | | | |
| 494 | 140 | cristo | 532 | 130 | fuego | 568 | 124 | fuerte |
| 495 | 140 | llegó | 533 | 130 | interior | 569 | 124 | grande |
| 496 | 140 | meses | 534 | 130 | pude | 570 | 124 | gritó |
| 497 | 140 | pequeño | 535 | 130 | sueño | 571 | 124 | modos |
| | | | 536 | 129 | comida | 572 | 124 | número |
| 498 | 140 | tuve | 537 | 129 | debo | | | |
| 499 | 139 | humbert | 538 | 129 | peor | 573 | 124 | seis |
| | | | 539 | 129 | | 574 | 124 | tiempos |
| 500 | 139 | sombra | | | probablemente | | | |
| 501 | 139 | ubertino | 540 | 129 | slim | 575 | 123 | fueron |
| | | | 541 | 128 | | 576 | 123 | lectura |
| 502 | 138 | deseo | | | automóvil | 577 | 123 | marido |
| 503 | 138 | edificio | 542 | 128 | dolor | 578 | 123 | suerte |
| | | | 543 | 128 | hacen | 579 | 123 | viaje |
| 504 | 138 | fui | 544 | 128 | niños | 580 | 122 | azul |
| 505 | 138 | general | 545 | 128 | seguir | 581 | 122 | escuela |
| 506 | 138 | propio | 546 | 128 | sigue | 582 | 122 | rents |
| 507 | 137 | bernardo | 547 | 128 | viento | 583 | 122 | sola |
| 508 | 137 | efecto | 548 | 127 | hice | 584 | 122 | vivir |
| 509 | 137 | escrito | 549 | 127 | hombros | 585 | 121 | favor |
| 510 | 137 | mía | | | | 586 | 121 | mamá |
| 511 | 137 | problema | 550 | 127 | ludmilla | 587 | 121 | vaya |
| 512 | 136 | aquellas | | | | 588 | 120 | acerca |
| | | | 551 | 127 | pudiera | 589 | 120 | cuales |
| 513 | 136 | | | | | 590 | 120 | perdido |
| | | conversación | 552 | 127 | ésta | | | |
| 514 | 136 | grupo | 553 | 126 | | 591 | 120 | phoebe |
| 515 | 136 | imagen | | | encontrado | 592 | 119 | igual |
| 516 | 136 | principio | 554 | 126 | papel | 593 | 119 | |
| 517 | 135 | amigos | 555 | 126 | partes | | | lentamente |
| 518 | 135 | comer | 556 | 126 | pueblo | 594 | 119 | niño |
| 519 | 134 | china | 557 | 126 | vieja | 595 | 119 | nuestras |
| 520 | 134 | cillerero | 558 | 125 | habla | | | |
| | | | 559 | 125 | | 596 | 119 | oro |
| 521 | 134 | cuya | | | naturaleza | 597 | 119 | única |
| 522 | 134 | dejado | | | | | | |

| | | | | | | | | |
|-----|-----|-----------|-----|-----|--------------|-----|-----|--------------|
| 598 | 118 | derecha | 635 | 111 | adso | 671 | 105 | dicen |
| | | | 636 | 111 | atención | 672 | 105 | ganas |
| 599 | 118 | mil | | | | 673 | 105 | loco |
| 600 | 118 | pocos | 637 | 111 | cuyo | 674 | 105 | mala |
| 601 | 118 | santo | 638 | 111 | difícil | 675 | 105 | piedra |
| 602 | 118 | torno | 639 | 111 | espera | 676 | 104 | conchetta |
| 603 | 117 | despues | 640 | 111 | | | | |
| | | | | | estábamos | 677 | 104 | espalda |
| 604 | 117 | falta | 641 | 111 | | | | |
| 605 | 117 | mar | | | habitaciones | 678 | 104 | leído |
| 606 | 117 | miraba | 642 | 111 | | 679 | 104 | olor |
| 607 | 117 | nariz | | | respondió | 680 | 104 | puse |
| 608 | 117 | pecho | 643 | 111 | tommy | 681 | 104 | vestido |
| 609 | 117 | quisiera | 644 | 111 | época | 682 | 103 | candy |
| | | | 645 | 110 | buenas | 683 | 103 | colegio |
| 610 | 116 | bolsillo | 646 | 110 | debajo | 684 | 103 | espejo |
| | | | 647 | 110 | dirección | 685 | 103 | ocasión |
| 611 | 116 | charlotte | 648 | 110 | pienso | | | |
| | | | 649 | 110 | policía | 686 | 103 | relación |
| 612 | 116 | justo | 650 | 110 | terrible | | | |
| 613 | 116 | lleno | 651 | 109 | | 687 | 103 | supe |
| 614 | 116 | media | | | comprendí | 688 | 102 | baño |
| 615 | 116 | risa | 652 | 109 | memori | 689 | 102 | buscar |
| 616 | 116 | scalambri | | | | 690 | 102 | cabo |
| 617 | 116 | supongo | 653 | 109 | teléfono | 691 | 102 | |
| | | | | | | | | esperando |
| 618 | 115 | autor | 654 | 108 | anterior | 692 | 102 | pasos |
| 619 | 115 | lleva | | | | 693 | 102 | pueda |
| 620 | 115 | podido | 655 | 108 | buenos | 694 | 102 | rojo |
| 621 | 115 | sentí | 656 | 108 | esperaba | 695 | 102 | trataba |
| 622 | 114 | decirle | | | | 696 | 102 | ventanas |
| 623 | 114 | llamaba | 657 | 108 | michele | | | |
| | | | | | | 697 | 102 | volví |
| 624 | 114 | salina | 658 | 108 | oye | 698 | 101 | ambos |
| 625 | 113 | adelante | 659 | 108 | siguió | 699 | 101 | caer |
| | | | 660 | 108 | van | 700 | 101 | d |
| 626 | 113 | menudo | 661 | 108 | vuelta | 701 | 101 | |
| | | | 662 | 107 | ei | | | encontraba |
| 627 | 113 | oí | 663 | 107 | montón | 702 | 101 | malo |
| 628 | 113 | polvo | | | | 703 | 100 | billy |
| 629 | 112 | bencio | 664 | 106 | baja | 704 | 100 | cartas |
| 630 | 112 | conocido | 665 | 106 | chico | 705 | 100 | coche |
| 631 | 112 | dientes | 666 | 106 | dolly | 706 | 100 | entró |
| 632 | 112 | | 667 | 106 | hijos | 707 | 100 | madera |
| | | muchacho | 668 | 106 | pasó | 708 | 100 | obra |
| 633 | 112 | parecían | 669 | 106 | segunda | 709 | 100 | |
| | | | | | | | | pensamientos |
| 634 | 112 | pensando | 670 | 106 | vas | 710 | 100 | salvatore |

| | | | | | | | | |
|-----|-----|--------------|-----|----|------------|-----|----|----------------|
| 711 | 100 | vuelve | 750 | 95 | ojo | 795 | 89 | chino |
| 712 | 99 | acababa | 751 | 95 | placer | 796 | 89 | |
| | | | 752 | 95 | verde | | | comisario |
| 713 | 99 | acercó | 753 | 95 | vuelto | 797 | 89 | dolores |
| 714 | 99 | adelmo | 754 | 94 | cree | 798 | 89 | gesto |
| 715 | 99 | espíritu | 755 | 94 | | 799 | 89 | lisboa |
| | | | | | cualquiera | 800 | 89 | marcha |
| 716 | 99 | frío | 756 | 94 | culo | 801 | 89 | mei |
| 717 | 99 | | 757 | 94 | vacío | 802 | 89 | |
| | | importante | 758 | 93 | abierta | | | scriptorium |
| 718 | 99 | llena | 759 | 93 | blanca | 803 | 88 | aun |
| 719 | 99 | presencia | 760 | 93 | centro | 804 | 88 | chaqueta |
| 720 | 99 | quedó | 761 | 93 | darle | 805 | 88 | hermana |
| 721 | 99 | secreto | 762 | 93 | diablo | | | |
| 722 | 99 | tíos | 763 | 93 | esperar | 806 | 88 | hermanos |
| 723 | 98 | cena | 764 | 93 | habrá | 807 | 88 | llevar |
| 724 | 98 | culpa | 765 | 93 | oír | 808 | 88 | resulta |
| 725 | 98 | menor | 766 | 93 | rey | 809 | 88 | suficiente |
| 726 | 98 | mirar | 767 | 93 | ves | 810 | 88 | treinta |
| 727 | 98 | trata | 768 | 92 | alta | 811 | 87 | cuidado |
| 728 | 97 | cuál | 769 | 92 | creía | | | |
| 729 | 97 | deja | 770 | 92 | fray | 812 | 87 | izquierda |
| 730 | 97 | | 771 | 92 | resto | 813 | 87 | |
| | | exactamente | 772 | 92 | salió | | | problemas |
| 731 | 97 | | 773 | 91 | causa | 814 | 87 | pudo |
| | | franciscanos | 774 | 91 | cincuenta | 815 | 87 | viene |
| 732 | 97 | jóvenes | 775 | 91 | dólares | 816 | 86 | añadió |
| | | | 776 | 91 | estuviera | 817 | 86 | conocía |
| 733 | 97 | minutos | 777 | 91 | marta | | | |
| | | | 778 | 91 | | 818 | 86 | conozco |
| 734 | 97 | | | | movimiento | | | |
| | | necesidad | 779 | 91 | objeto | 819 | 86 | dormir |
| 735 | 97 | oscuridad | 780 | 91 | siglos | 820 | 86 | empecé |
| 736 | 97 | salvo | 781 | 91 | tren | | | |
| 737 | 97 | tema | 782 | 90 | borde | 821 | 86 | hojas |
| 738 | 96 | encontré | 783 | 90 | cuánto | 822 | 86 | jo |
| | | | 784 | 90 | derecho | 823 | 86 | levantó |
| | | | | | | 824 | 86 | natural |
| 739 | 96 | franco | | | | 825 | 86 | pena |
| 740 | 96 | fácil | 785 | 90 | gris | 826 | 86 | santa |
| 741 | 96 | gracia | 786 | 90 | gusto | 827 | 85 | echó |
| 742 | 96 | matty | 787 | 90 | llamar | 828 | 85 | estación |
| 743 | 96 | oscuro | 788 | 90 | luna | | | |
| 744 | 96 | pasaba | 789 | 90 | mucha | | | |
| 745 | 96 | queda | 790 | 90 | sentó | 829 | 85 | estilo |
| 746 | 96 | venga | 791 | 90 | siendo | 830 | 85 | hacían |
| 747 | 95 | aquello | 792 | 90 | veinte | 831 | 85 | |
| 748 | 95 | jardín | 793 | 90 | ése | | | inmediatamente |
| 749 | 95 | juntos | 794 | 89 | busca | 832 | 85 | libre |

| | | | | | | | | |
|-----|----|------------|-----|----|---------------|-----|----|------------|
| 833 | 85 | negra | 871 | 81 | calogero | 907 | 79 | exclamó |
| 834 | 85 | nombres | | | | | | |
| | | | 872 | 81 | campo | 908 | 79 | extraño |
| 835 | 85 | penso | 873 | 81 | | | | |
| 836 | 85 | quizás | | | comprender | 909 | 79 | golpe |
| 837 | 85 | relato | 874 | 81 | decirlo | 910 | 79 | gritos |
| 838 | 84 | belleza | 875 | 81 | empieza | 911 | 79 | gustaba |
| 839 | 84 | calor | | | | | | |
| 840 | 84 | camisa | 876 | 81 | guerra | 912 | 79 | serio |
| 841 | 84 | contrario | 877 | 81 | habéis | 913 | 79 | |
| 842 | 84 | dulcino | 878 | 81 | horrible | | | totalmente |
| | | | | | | 914 | 79 | verdadera |
| 843 | 84 | especial | 879 | 81 | hospital | 915 | 78 | an |
| | | | | | | 916 | 78 | escaleras |
| 844 | 84 | esperanza | 880 | 81 | jodido | 917 | 78 | hablado |
| 845 | 84 | hechos | 881 | 81 | mitad | | | |
| 846 | 84 | llega | 882 | 81 | noches | 918 | 78 | hago |
| 847 | 84 | paredes | 883 | 81 | ocho | 919 | 78 | hermosa |
| | | | 884 | 81 | quedado | | | |
| 848 | 84 | pirrone | | | | 920 | 78 | periodico |
| 849 | 84 | tendría | 885 | 81 | regla | 921 | 78 | pone |
| 850 | 83 | basta | 886 | 80 | café | 922 | 78 | siete |
| 851 | 83 | | 887 | 80 | colega | 923 | 78 | tantos |
| | | excelencia | 888 | 80 | | 924 | 78 | varios |
| 852 | 83 | ideas | | | curiosidad | 925 | 78 | viejos |
| 853 | 83 | juan | 889 | 80 | distancia | 926 | 77 | bar |
| 854 | 83 | jugar | 890 | 80 | distinto | 927 | 77 | buscando |
| 855 | 83 | llamado | | | | 928 | 77 | conocer |
| | | | 891 | 80 | escritor | | | |
| 856 | 83 | lágrimas | | | | 929 | 77 | |
| | | | 892 | 80 | fe | | | dormitorio |
| 857 | 83 | página | 893 | 80 | lago | 930 | 77 | lámpara |
| 858 | 82 | | 894 | 80 | odio | | | |
| | | imposible | 895 | 80 | | 931 | 77 | stradlater |
| 859 | 82 | | | | pantalones | 932 | 77 | taxi |
| | | impresión | 896 | 80 | podemos | 933 | 76 | detuvo |
| 860 | 82 | in | 897 | 80 | recordar | 934 | 76 | hermanas |
| 861 | 82 | mes | | | | 935 | 76 | poner |
| 862 | 82 | monje | 898 | 80 | salido | 936 | 76 | sentada |
| 863 | 82 | murió | 899 | 80 | signos | | | |
| 864 | 82 | país | 900 | 80 | silla | 937 | 76 | sonrió |
| 865 | 82 | pocas | 901 | 80 | suya | 938 | 76 | subir |
| 866 | 82 | razones | 902 | 80 | texto | 939 | 76 | tratando |
| | | | 903 | 79 | | | | |
| 867 | 82 | seda | | | completamente | 940 | 76 | venir |
| 868 | 82 | situación | 904 | 79 | debes | 941 | 75 | chicas |
| 869 | 82 | verano | 905 | 79 | dedo | 942 | 75 | cuándo |
| 870 | 81 | animales | 906 | 79 | doce | | | |

| | | | | | | | | |
|-----|----|---------------|------|----|-------------|------|----|---------------|
| 943 | 75 | | 981 | 72 | mark | 1022 | 69 | cardenal |
| | | diferentes | 982 | 72 | paz | | | |
| 944 | 75 | entrada | 983 | 72 | ponía | 1023 | 69 | dianne |
| 945 | 75 | formas | 984 | 72 | respuesta | 1024 | 69 | |
| 946 | 75 | haría | 985 | 72 | solía | | | emperador |
| 947 | 75 | iban | 986 | 72 | suave | 1025 | 69 | francisco |
| 948 | 75 | larga | 987 | 72 | últimos | 1026 | 69 | llevó |
| 949 | 75 | ocurrió | | | | 1027 | 69 | novelas |
| 950 | 75 | pequeños | 988 | 71 | abrió | | | |
| 951 | 75 | río | 989 | 71 | arte | 1028 | 69 | |
| 952 | 75 | vuestra | 990 | 71 | espero | | | perfectamente |
| 953 | 74 | espacio | 991 | 71 | evitar | 1029 | 69 | piensa |
| | | | 992 | 71 | | 1030 | 69 | recuerdos |
| 954 | 74 | hicieron | | | momentos | 1031 | 69 | |
| | | | 993 | 71 | pobreza | | | relaciones |
| 955 | 74 | imágenes | | | | 1032 | 69 | tenga |
| 956 | 74 | | 994 | 71 | prisa | 1033 | 69 | árboles |
| | | matrimonio | 995 | 71 | sentir | 1034 | 68 | capullos |
| 957 | 74 | misimos | 996 | 71 | visita | | | |
| | | | 997 | 70 | alguno | 1035 | 68 | encuentra |
| 958 | 74 | nota | 998 | 70 | común | 1036 | 68 | hermoso |
| 959 | 74 | pub | 999 | 70 | descubrir | | | |
| 960 | 74 | s | 1000 | 70 | destino | 1037 | 68 | leyendo |
| 961 | 74 | serie | 1001 | 70 | diría | | | |
| 962 | 73 | | 1002 | 70 | esté | 1038 | 68 | llamó |
| | | bibliotecario | 1003 | 70 | evidente | 1039 | 68 | llevado |
| 963 | 73 | dejo | | | | 1040 | 68 | llorar |
| 964 | 73 | existe | 1004 | 70 | feliz | 1041 | 68 | partir |
| 965 | 73 | frase | 1005 | 70 | figura | 1042 | 68 | pobres |
| 966 | 73 | morir | 1006 | 70 | habló | 1043 | 68 | rodillas |
| 967 | 73 | partido | 1007 | 70 | harold | | | |
| 968 | 73 | patio | 1008 | 70 | metido | 1044 | 68 | siglo |
| 969 | 73 | significa | 1009 | 70 | miré | 1045 | 68 | total |
| | | | 1010 | 70 | podían | 1046 | 68 | traje |
| 970 | 73 | | 1011 | 70 | preguntas | 1047 | 68 | tuviera |
| | | simplemente | 1012 | 70 | profesor | 1048 | 68 | ustedes |
| 971 | 73 | simples | | | | 1049 | 68 | venido |
| | | | 1013 | 70 | quiso | 1050 | 67 | alegría |
| 972 | 73 | sueños | 1014 | 70 | | 1051 | 67 | completo |
| 973 | 73 | tranquilo | | | rápidamente | 1052 | 67 | cuarenta |
| 974 | 72 | ayer | 1015 | 70 | señal | | | |
| 975 | 72 | chicos | 1016 | 70 | sonido | 1053 | 67 | dando |
| 976 | 72 | dices | 1017 | 70 | volvía | 1054 | 67 | |
| 977 | 72 | diré | 1018 | 70 | ésa | | | donna fugata |
| 978 | 72 | dispuesto | 1019 | 69 | acto | 1055 | 67 | enorme |
| 979 | 72 | entrado | 1020 | 69 | antonio | | | |
| | | | | | | 1056 | 67 | escuchar |
| 980 | 72 | llegué | 1021 | 69 | asunto | | | |

| | | | | | | | | |
|------|----|-----------------|------|----|------------|------|----|--------------|
| 1057 | 67 | | 1091 | 65 | dejando | 1127 | 64 | primo |
| | | finalmente | | | | 1128 | 64 | prueba |
| 1058 | 67 | infierno | 1092 | 65 | diga | 1129 | 64 | semanas |
| | | | 1093 | 65 | | | | |
| 1059 | 67 | llegada | | | empezaba | 1130 | 64 | valor |
| 1060 | 67 | medida | 1094 | 65 | escritorio | 1131 | 64 | ven |
| 1061 | 67 | orgullo | 1095 | 65 | herejes | 1132 | 63 | animal |
| 1062 | 67 | palacio | 1096 | 65 | inútil | 1133 | 63 | cadáver |
| 1063 | 67 | quizá | 1097 | 65 | ira | | | |
| 1064 | 67 | respecto | 1098 | 65 | oficio | 1134 | 63 | casas |
| | | | 1099 | 65 | papá | 1135 | 63 | cola |
| 1065 | 67 | retrato | 1100 | 65 | pecado | 1136 | 63 | crees |
| 1066 | 67 | sucedió | 1101 | 65 | película | 1137 | 63 | daño |
| | | | | | | 1138 | 63 | |
| 1067 | 67 | temor | 1102 | 65 | quedaba | | | experiencia |
| 1068 | 67 | tercera | | | | 1139 | 63 | gato |
| 1069 | 67 | york | 1103 | 65 | reír | 1140 | 63 | hombro |
| 1070 | 67 | zapatos | 1104 | 65 | ruido | | | |
| | | | 1105 | 65 | | 1141 | 63 | jaco |
| 1071 | 66 | bajar | | | suavemente | 1142 | 63 | joder |
| 1072 | 66 | bolsa | 1106 | 65 | tantas | 1143 | 63 | |
| 1073 | 66 | gustaría | 1107 | 65 | tenía | | | movimientos |
| | | | 1108 | 65 | triste | 1144 | 63 | muertos |
| 1074 | 66 | hierbas | 1109 | 65 | varias | | | |
| 1075 | 66 | historias | 1110 | 65 | veras | 1145 | 63 | obras |
| | | | 1111 | 65 | voluntad | 1146 | 63 | |
| 1076 | 66 | nicola | | | | | | pensamiento |
| 1077 | 66 | niñas | 1112 | 64 | acabó | 1147 | 63 | seguido |
| 1078 | 66 | perder | 1113 | 64 | agradable | | | |
| 1079 | 66 | quieren | 1114 | 64 | barra | 1148 | 63 | servicio |
| | | | 1115 | 64 | cuento | | | |
| 1080 | 66 | ropa | 1116 | 64 | dan | 1149 | 63 | tanta |
| 1081 | 66 | saben | 1117 | 64 | darme | 1150 | 63 | |
| 1082 | 66 | sucedido | 1118 | 64 | | | | universidad |
| 1083 | 66 | voces | | | diferencia | 1151 | 62 | abrir |
| 1084 | 65 | | 1119 | 64 | entender | 1152 | 62 | cambiar |
| | | acontecimientos | | | | | | |
| 1085 | 65 | apareció | 1120 | 64 | mayoría | 1153 | 62 | conejos |
| | | | | | | | | |
| 1086 | 65 | blancos | 1121 | 64 | ministro | 1154 | 6 | conoce |
| | | | | | | 1155 | 62 | correr |
| 1087 | 65 | colores | 1122 | 64 | | 1156 | 62 | debido |
| 1088 | 65 | consigo | | | muchísimo | 1157 | 62 | dia |
| | | | 1123 | 64 | na | 1158 | 62 | dirigió |
| 1089 | 65 | | 1124 | 64 | necesario | 1159 | 62 | |
| | | continuación | 1125 | 64 | parecido | | | espirituales |
| 1090 | 65 | coro | | | | 1160 | 62 | fiesta |
| | | | 1126 | 64 | peligro | 1161 | 62 | lista |

| | | | | | | | | |
|------|----|--------------|------|----|--------------|------|----|----------------|
| 1162 | 62 | parecen | 1199 | 60 | estuviese | 1237 | 58 | sale |
| | | | 1200 | 60 | hacerle | 1238 | 58 | señores |
| 1163 | 62 | parecer | 1201 | 60 | italia | | | |
| 1164 | 62 | pensado | 1202 | 60 | ley | 1239 | 58 | stevie |
| | | | 1203 | 60 | negros | 1240 | 58 | the |
| 1165 | 62 | pensó | 1204 | 60 | nueve | 1241 | 58 | trabajar |
| 1166 | 62 | premio | 1205 | 60 | pidió | | | |
| 1167 | 62 | semejante | 1206 | 60 | propios | 1242 | 57 | circunstancias |
| 1168 | 62 | venía | 1207 | 60 | secretos | 1243 | 57 | contar |
| 1169 | 62 | verdadero | | | | 1244 | 57 | convertido |
| 1170 | 61 | abierto | 1208 | 60 | sorpresa | 1245 | 57 | costumbre |
| 1171 | 61 | actitud | 1209 | 60 | tienda | | | deseaba |
| 1172 | 61 | arena | 1210 | 60 | venters | 1246 | 57 | interesaba |
| 1173 | 61 | bebé | 1211 | 60 | volumen | | | |
| 1174 | 61 | brillante | | | | 1247 | 57 | interesa |
| | | | 1212 | 59 | adonde | | | |
| 1175 | 61 | caja | 1213 | 59 | contacto | 1248 | 57 | invierno |
| 1176 | 61 | cien | | | | | | |
| 1177 | 61 | confianza | 1214 | 59 | doy | 1249 | 57 | llegaba |
| 1178 | 61 | cuerpos | 1215 | 59 | entra | 1250 | 57 | matar |
| | | | 1216 | 59 | inglés | 1251 | 57 | negras |
| 1179 | 61 | dejé | 1217 | 59 | llamada | 1252 | 57 | pedido |
| 1180 | 61 | doble | | | | 1253 | 57 | posibilidad |
| 1181 | 61 | encontró | 1218 | 59 | londres | | | puertas |
| | | | 1219 | 59 | marana | 1254 | 57 | regresar |
| 1182 | 61 | frases | 1220 | 59 | naturalmente | 1255 | 57 | terreno |
| 1183 | 61 | futuro | | | | | | vaso |
| 1184 | 61 | johnny | 1221 | 59 | nuevos | 1256 | 57 | visión |
| 1185 | 61 | nínfula | 1222 | 59 | partida | 1257 | 57 | vivía |
| 1186 | 61 | olvidado | 1223 | 59 | ponerse | 1258 | 57 | aviñón |
| | | | | | | 1259 | 57 | banco |
| 1187 | 61 | perros | 1224 | 59 | primeros | 1260 | 56 | demonio |
| 1188 | 61 | redaccion | 1225 | 59 | recordaba | 1261 | 56 | esfuerzo |
| 1189 | 61 | saco | 1226 | 59 | signo | 1262 | 56 | extraña |
| 1190 | 61 | sombrero | 1227 | 59 | techo | | | fría |
| 1191 | 61 | tocar | 1228 | 58 | asiento | 1263 | 56 | hijas |
| 1192 | 60 | ackley | 1229 | 58 | camarero | | | huang |
| 1193 | 60 | autobús | 1230 | 58 | compañía | 1264 | 56 | lee |
| | | | 1231 | 58 | deben | 1265 | 56 | llevo |
| 1194 | 60 | breve | 1232 | 58 | decidido | 1266 | 56 | motivo |
| 1195 | 60 | cogió | | | | 1267 | 56 | música |
| 1196 | 60 | colegas | 1233 | 58 | flores | 1268 | 56 | |
| | | | 1234 | 58 | herejía | 1269 | 56 | |
| 1197 | 60 | dejaba | 1235 | 58 | inocente | 1270 | 56 | |
| 1198 | 60 | desaparecido | 1236 | 58 | presa | 1271 | 56 | |

| | | | | | | | | |
|------|----|---------------|------|----|------------|------|----|------------|
| 1272 | 56 | oscuras | 1310 | 55 | propósito | 1348 | 53 | creer |
| 1273 | 56 | pasillo | 1311 | 55 | salón | 1349 | 53 | curso |
| 1274 | 56 | princesa | 1312 | 54 | abiertos | 1350 | 53 | espaldas |
| | | | 1313 | 54 | ay | 1351 | 53 | estaría |
| 1275 | 56 | reconocer | 1314 | 54 | blancas | 1352 | 53 | jesús |
| 1276 | 56 | rápido | | | | 1353 | 53 | obstante |
| 1277 | 56 | sabido | 1315 | 54 | cabrones | | | |
| 1278 | 56 | siente | 1316 | 54 | cayó | 1354 | 53 | orejas |
| 1279 | 56 | tambien | 1317 | 54 | cocinero | 1355 | 53 | quedé |
| | | | | | | 1356 | 53 | respeto |
| 1280 | 56 | ted | 1318 | 54 | contigo | 1357 | 53 | sally |
| 1281 | 56 | tipos | | | | 1358 | 53 | salí |
| 1282 | 56 | tomado | 1319 | 54 | crooks | 1359 | 53 | sintió |
| | | | 1320 | 54 | dulce | 1360 | 53 | sombras |
| 1283 | 56 | vete | 1321 | 54 | escena | | | |
| 1284 | 56 | vos | 1322 | 54 | francia | 1361 | 53 | |
| 1285 | 55 | ama | 1323 | 54 | jane | | | superficie |
| 1286 | 55 | asesino | 1324 | 54 | lluvia | 1362 | 53 | torreón |
| 1287 | 55 | ayuda | 1325 | 54 | manana | 1363 | 53 | título |
| 1288 | 55 | bella | | | | 1364 | 53 | |
| 1289 | 55 | club | 1326 | 54 | | | | vacacione |
| 1290 | 55 | contestó | | | manuscrito | 1365 | 52 | ala |
| | | | 1327 | 54 | mediante | 1366 | 52 | anteojos |
| 1291 | 55 | decían | 1328 | 54 | mejores | | | |
| 1292 | 55 | error | | | | 1367 | 52 | beardsley |
| 1293 | 55 | est | 1329 | 54 | máquina | 1368 | 52 | caballos |
| 1294 | 55 | estados | | | | | | |
| 1295 | 55 | | 1330 | 54 | pasta | 1369 | 52 | |
| | | evidentemente | 1331 | 54 | recién | | | casualidad |
| 1296 | 55 | folios | 1332 | 54 | reloj | 1370 | 52 | ciertas |
| 1297 | 55 | haberse | 1333 | 54 | rich | 1371 | 52 | clara |
| | | | 1334 | 54 | roja | 1372 | 52 | cojones |
| 1298 | 55 | intento | 1335 | 54 | rosa | | | |
| 1299 | 55 | lugares | 1336 | 54 | sabemos | 1373 | 52 | |
| 1300 | 55 | línea | | | | | | conciencia |
| 1301 | 55 | necesita | 1337 | 54 | sean | 1374 | 52 | corriente |
| | | | 1338 | 54 | seguridad | 1375 | 52 | cárcel |
| 1302 | 55 | ninguno | 1339 | 54 | sofá | 1376 | 52 | decidí |
| | | | 1340 | 54 | | 1377 | 52 | extremo |
| 1303 | 55 | notaba | | | solamente | | | |
| 1304 | 55 | ocurre | 1341 | 54 | sucede | 1378 | 52 | hiciera |
| 1305 | 55 | | 1342 | 54 | tales | 1379 | 52 | humor |
| | | oportunidad | 1343 | 54 | tomó | 1380 | 52 | largas |
| 1306 | 55 | oscura | 1344 | 54 | éstos | 1381 | 52 | líneas |
| 1307 | 55 | pasión | 1345 | 53 | acaba | 1382 | 52 | mismas |
| 1308 | 55 | pequeñas | 1346 | 53 | central | | | |
| 1309 | 55 | playa | 1347 | 53 | cine | | | |

| | | | | | | | | |
|------|----|-----------|------|----|-----------|------|----|-----------|
| 1383 | 52 | posición | 1400 | 51 | laberinto | 1420 | 50 | crimen |
| | | | 1401 | 51 | maleta | 1421 | 50 | eché |
| 1384 | 52 | sexo | 1402 | 51 | miro | 1422 | 50 | esquina |
| 1385 | 52 | verla | 1403 | 51 | | | | |
| 1386 | 52 | vientre | | | muchachos | 1423 | 50 | fila |
| 1387 | 51 | acabado | 1404 | 51 | norte | 1424 | 50 | griego |
| | | | 1405 | 51 | ocasiones | 1425 | 50 | libertad |
| 1388 | 51 | altura | 1406 | 51 | ocurrido | | | |
| 1389 | 51 | ambas | | | | 1426 | 50 | llegamos |
| 1390 | 51 | capítulo | 1407 | 51 | plan | 1427 | 50 | locura |
| | | | 1408 | 51 | puntos | 1428 | 50 | maletas |
| 1391 | 51 | carácter | 1409 | 51 | revista | | | |
| | | | 1410 | 51 | señaló | 1429 | 50 | nieve |
| 1392 | 51 | claridad | 1411 | 51 | teníamos | 1430 | 50 | nina |
| | | | 1412 | 51 | | 1431 | 50 | pagar |
| 1393 | 51 | crímenes | | | terminado | 1432 | 50 | parque |
| 1394 | 51 | distintas | 1413 | 51 | tratar | 1433 | 50 | quise |
| | | | 1414 | 51 | vale | 1434 | 50 | sillón |
| 1395 | 51 | drogas | 1415 | 51 | | 1435 | 50 | sur |
| 1396 | 51 | habrían | | | vergüenza | 1436 | 50 | té |
| | | | 1416 | 50 | andar | 1437 | 50 | uso |
| 1397 | 51 | haga | 1417 | 50 | asi | 1438 | 50 | vestíbulo |
| 1398 | 51 | humo | 1418 | 50 | calma | 1439 | 50 | ánimo |
| 1399 | 51 | intención | 1419 | 50 | ciencia | | | |

ANEXO 4

PALABRAS CON FRECUENCIA MAYOR O IGUAL A 50 DEL CORPUS LIT_EN_ES

| | | | | | | | | |
|----|------|----------|----|-----|---------|-----|-----|-----------|
| 1 | 8945 | en | 39 | 510 | mí | 80 | 328 | decir |
| 2 | 5240 | me | 40 | 499 | uno | 81 | 326 | sé |
| 3 | 4766 | se | 41 | 497 | dije | 82 | 322 | durante |
| 4 | 3806 | lo | 42 | 487 | otro | | | |
| 5 | 2499 | le | 43 | 484 | otra | 83 | 317 | donde |
| 6 | 1905 | era | 44 | 483 | algo | 84 | 313 | padre |
| 7 | 1730 | más | 45 | 473 | poco | 85 | 310 | esta |
| 8 | 1541 | cuando | 46 | 465 | cómo | 86 | 310 | noche |
| 9 | 1511 | si | 47 | 455 | nunca | 87 | 309 | cosa |
| 10 | 1122 | estaba | 48 | 448 | george | 88 | 308 | hecho |
| 11 | 1110 | te | 49 | 448 | verdad | 89 | 307 | bueno |
| 12 | 1094 | todo | 50 | 446 | cabeza | 90 | 305 | aún |
| 13 | 993 | dijo | 51 | 440 | hacer | 91 | 302 | menos |
| 14 | 930 | madre | 52 | 435 | años | 92 | 301 | allí |
| 15 | 928 | muy | 53 | 431 | podía | 93 | 300 | |
| 16 | 890 | eso | 54 | 425 | luego | | | demasiado |
| 17 | 866 | vez | 55 | 425 | ser | 94 | 299 | veces |
| 18 | 854 | sobre | 56 | 423 | mano | 95 | 294 | hace |
| 19 | 812 | tenía | 57 | 419 | ojos | 96 | 291 | tengo |
| 20 | 802 | ya | 58 | 414 | antes | 97 | 284 | este |
| 21 | 716 | sólo | 59 | 413 | tío | 98 | 280 | esos |
| 22 | 707 | nada | 60 | 408 | tiene | 99 | 280 | voz |
| 23 | 672 | así | 61 | 406 | ver | 100 | 278 | mejor |
| 24 | 653 | dos | 62 | 405 | fue | 101 | 278 | va |
| 25 | 630 | tan | 63 | 400 | dice | 102 | 277 | puerta |
| 26 | 629 | ese | 64 | 388 | cosas | 103 | 274 | lolita |
| 27 | 614 | después | 65 | 378 | tal | 104 | 271 | todas |
| | | | 66 | 376 | día | 105 | 269 | cama |
| 28 | 586 | ahora | 67 | 374 | mucho | 106 | 267 | cara |
| 29 | 571 | esa | 68 | 373 | entre | 107 | 266 | nadie |
| 30 | 570 | todos | 69 | 370 | hay | 108 | 265 | hombre |
| 31 | 567 | nos | 70 | 369 | aquí | | | |
| 32 | 565 | bien | 71 | 358 | mismo | 109 | 264 | boy |
| 33 | 565 | mientras | 72 | 355 | tiempo | 110 | 261 | eran |
| | | | 73 | 354 | les | 111 | 261 | parte |
| 34 | 537 | ni | 74 | 352 | también | 112 | 261 | tú |
| 35 | 527 | sí | | | | 113 | 259 | casi |
| 36 | 523 | entonces | 75 | 346 | lennie | 114 | 259 | toda |
| | | | 76 | 346 | vida | 115 | 258 | quería |
| 37 | 523 | siempre | 77 | 333 | lado | 116 | 257 | esto |
| | | | 78 | 329 | cada | 117 | 257 | pues |
| 38 | 513 | casa | 79 | 329 | mujer | 118 | 254 | ir |

| | | | | | | | | |
|-----|-----|----------|-----|-----|------------|-----|-----|-----------|
| 119 | 253 | fuera | 164 | 198 | voy | 208 | 167 | nuestra |
| 120 | 252 | sabía | 165 | 197 | eh | 209 | 166 | dónde |
| 121 | 250 | iba | 166 | 197 | hizo | 210 | 165 | mal |
| 122 | 248 | sabes | 167 | 197 | modo | 211 | 164 | ahí |
| 123 | 244 | puede | 168 | 196 | aquel | 212 | 164 | da |
| 124 | 241 | hacía | 169 | 196 | dicho | 213 | 163 | calle |
| 125 | 238 | sick | 170 | 196 | señor | 214 | 162 | dios |
| 126 | 237 | algún | 171 | 195 | spud | 215 | 161 | mesa |
| 127 | 234 | fin | 172 | 194 | aquella | 216 | 158 | encima |
| 128 | 233 | cuarto | 173 | 193 | mundo | 217 | 158 | puto |
| 129 | 232 | momento | 174 | 193 | vi | 218 | 157 | claro |
| 130 | 231 | capullo | 175 | 193 | volvió | 219 | 157 | nosotros |
| 131 | 231 | nuevo | 176 | 190 | quién | | | |
| 132 | 231 | parece | 177 | 189 | buena | 220 | 155 | días |
| 133 | 231 | tres | 178 | 189 | decía | 221 | 155 | tras |
| 134 | 229 | habría | 179 | 187 | agua | 222 | 154 | boca |
| 135 | 229 | puedo | 180 | 186 | estar | 223 | 154 | chica |
| 136 | 229 | sido | 181 | 185 | junto | 224 | 153 | cualquier |
| 137 | 228 | bajo | 182 | 185 | manos | 225 | 153 | manera |
| 138 | 223 | creo | 183 | 185 | usted | 226 | 153 | otros |
| 139 | 223 | tanto | 184 | 184 | hija | 227 | 153 | pie |
| 140 | 221 | hablar | 185 | 180 | tener | 228 | 153 | preguntó |
| 141 | 219 | renton | 186 | 179 | contra | 229 | 151 | cual |
| 142 | 217 | bastante | 187 | 179 | | 230 | 151 | pequeña |
| | | | | | habitación | | | |
| 143 | 216 | gran | 188 | 179 | podría | 231 | 151 | recuerdo |
| 144 | 216 | habían | 189 | 179 | quizá | | | |
| 145 | 216 | quiero | 190 | 178 | visto | 232 | 149 | demás |
| 146 | 216 | tía | 191 | 175 | tarde | 233 | 149 | dentro |
| 147 | 215 | cabrón | 192 | 175 | vamos | 234 | 149 | familia |
| 148 | 215 | esas | 193 | 174 | tienes | 235 | 149 | quieres |
| 149 | 215 | viejo | 194 | 173 | alguien | 236 | 149 | señora |
| 150 | 214 | estaban | 195 | 173 | cierto | 237 | 148 | haze |
| | | | 196 | 172 | dinero | 238 | 148 | par |
| 151 | 209 | begbie | 197 | 172 | miró | 239 | 147 | empezó |
| 152 | 209 | e | 198 | 172 | punto | | | |
| 153 | 205 | cuenta | 199 | 172 | soy | 240 | 147 | llevaba |
| 154 | 205 | hubiera | 200 | 171 | lugar | 241 | 147 | medio |
| | | | 201 | 170 | estado | 242 | 146 | negro |
| 155 | 205 | niña | 202 | 170 | primera | 243 | 146 | sino |
| 156 | 204 | parecía | | | | 244 | 145 | cuanto |
| 157 | 204 | puta | 203 | 169 | mañana | 245 | 145 | joven |
| 158 | 203 | misma | | | | 246 | 144 | atrás |
| 159 | 203 | pronto | 204 | 169 | mierda | 247 | 144 | dio |
| 160 | 203 | tipo | 205 | 168 | siquiera | 248 | 143 | alguna |
| 161 | 201 | digo | | | | 249 | 143 | hermano |
| 162 | 201 | gente | 206 | 167 | cuerpo | | | |
| 163 | 198 | pelo | 207 | 167 | haber | 250 | 142 | hijo |

| | | | | | | | | |
|-----|-----|-----------|-----|-----|-----------|-----|-----|---------|
| 251 | 139 | suelo | 290 | 119 | algunos | 327 | 110 | mujeres |
| 252 | 139 | través | | | | | | |
| 253 | 138 | embargo | 291 | 119 | cuatro | 328 | 109 | ido |
| | | | 292 | 119 | otras | 329 | 109 | luz |
| 254 | 138 | humbert | 293 | 119 | rostro | 330 | 109 | ningún |
| | | | 294 | 119 | veía | 331 | 109 | piernas |
| 255 | 137 | ante | 295 | 118 | además | 332 | 108 | labios |
| 256 | 137 | realidad | | | | 333 | 108 | puesto |
| | | | 296 | 118 | conmigo | 334 | 106 | brazo |
| 257 | 136 | cinco | | | | 335 | 106 | detrás |
| 258 | 135 | puedes | 297 | 118 | esposa | 336 | 106 | dolly |
| 259 | 135 | sabe | 298 | 118 | hora | 337 | 106 | fui |
| 260 | 133 | china | 299 | 117 | daba | 338 | 106 | libro |
| 261 | 132 | clase | 300 | 117 | siento | 339 | 105 | dar |
| 262 | 132 | pensé | 301 | 117 | ventana | 340 | 105 | escuela |
| 263 | 130 | buen | | | | 341 | 105 | nombre |
| 264 | 130 | mira | 302 | 116 | camino | | | |
| 265 | 130 | todavía | 303 | 116 | charlotte | 342 | 105 | piso |
| 266 | 129 | slim | | | | 343 | 105 | veo |
| 267 | 128 | nueva | 304 | 116 | jamás | 344 | 104 | aspecto |
| 268 | 128 | palabras | 305 | 116 | padres | | | |
| | | | 306 | 116 | saber | 345 | 104 | diez |
| 269 | 128 | sea | 307 | 116 | salir | 346 | 104 | gusta |
| 270 | 127 | digamos | 308 | 115 | haciendo | 347 | 104 | largo |
| | | | 309 | 115 | idea | 348 | 104 | niños |
| 271 | 127 | final | 310 | 115 | mirando | 349 | 104 | pies |
| 272 | 127 | oh | | | | 350 | 104 | quiere |
| 273 | 127 | puso | 311 | 115 | ninguna | 351 | 104 | seguro |
| 274 | 126 | pregunté | | | | 352 | 104 | volver |
| | | | 312 | 115 | rato | 353 | 103 | candy |
| 275 | 126 | sería | 313 | 115 | segundo | 354 | 103 | mayor |
| 276 | 125 | año | | | | 355 | 103 | pequeño |
| 277 | 125 | expresión | 314 | 115 | sol | | | |
| 278 | 125 | pobre | 315 | 114 | blanco | 356 | 103 | sentado |
| 279 | 124 | curley | 316 | 113 | caso | | | |
| 280 | 124 | pasado | 317 | 113 | grandes | 357 | 102 | brazos |
| 281 | 122 | quien | | | | 358 | 102 | cocina |
| 282 | 122 | rents | 318 | 112 | frente | 359 | 102 | forma |
| 283 | 121 | | 319 | 112 | pensar | 360 | 102 | muerte |
| | | automóvil | 320 | 112 | tienen | 361 | 101 | incluso |
| 284 | 121 | seguía | 321 | 112 | trabajo | 362 | 100 | aire |
| 285 | 121 | sentía | 322 | 112 | único | 363 | 100 | amor |
| 286 | 120 | mamá | 323 | 111 | dejar | 364 | 100 | billy |
| 287 | 120 | persona | 324 | 111 | tommy | 365 | 100 | cerca |
| | | | 325 | 110 | color | 366 | 100 | oye |
| 288 | 120 | phoebe | 326 | 110 | muchas | 367 | 100 | pasa |
| 289 | 120 | tomar | | | | 368 | 100 | sentido |
| | | | | | | 369 | 99 | debía |

| | | | | | | | | |
|-----|----|-----------|-----|----|---------------|-----|----|------------|
| 370 | 99 | dolor | 413 | 89 | culo | 450 | 81 | muerto |
| 371 | 99 | llegar | 414 | 89 | juego | 451 | 81 | terrible |
| 372 | 99 | marido | 415 | 89 | mei | 452 | 81 | van |
| 373 | 99 | vista | 416 | 89 | mía | 453 | 80 | debería |
| 374 | 98 | ciudad | 417 | 89 | | 454 | 80 | dedos |
| 375 | 98 | corazón | | | probablemente | 455 | 80 | pesar |
| | | | 418 | 89 | seguida | 456 | 80 | pudiera |
| 376 | 98 | mío | | | | | | |
| 377 | 97 | amigo | 419 | 89 | semana | 457 | 80 | sentí |
| 378 | 97 | hice | | | | 458 | 80 | supongo |
| 379 | 97 | hombres | 420 | 88 | pude | | | |
| | | | 421 | 88 | ti | 459 | 80 | vivir |
| 380 | 97 | importa | 422 | 87 | fuerte | 460 | 79 | amigos |
| | | | 423 | 87 | gracias | 461 | 79 | comer |
| 381 | 97 | perro | 424 | 87 | mente | 462 | 79 | empecé |
| 382 | 97 | sangre | 425 | 87 | piel | | | |
| 383 | 97 | siguiente | 426 | 86 | diciendo | 463 | 79 | historia |
| 384 | 97 | vas | | | | | | |
| 385 | 97 | vieja | 427 | 86 | franco | 464 | 79 | nariz |
| 386 | 96 | ah | 428 | 86 | jo | 465 | 79 | niño |
| 387 | 96 | alrededor | 429 | 86 | lejos | 466 | 79 | pared |
| 388 | 96 | matty | 430 | 86 | peor | 467 | 79 | siguió |
| 389 | 96 | pasar | 431 | 86 | pregunta | 468 | 79 | suerte |
| 390 | 96 | primer | | | | 469 | 79 | volví |
| 391 | 96 | puse | 432 | 86 | propia | 470 | 78 | arriba |
| 392 | 95 | edad | 433 | 85 | algunas | 471 | 78 | dejó |
| 393 | 95 | muchos | | | | 472 | 78 | grupo |
| | | | 434 | 85 | favor | 473 | 78 | libros |
| 394 | 95 | palabra | 435 | 85 | llamaba | 474 | 78 | principio |
| 395 | 94 | apenas | | | | 475 | 78 | última |
| 396 | 94 | colegio | 436 | 84 | sensación | 476 | 77 | acuerdo |
| 397 | 94 | hotel | 437 | 84 | silencio | | | |
| 398 | 94 | tíos | | | | 477 | 77 | hombros |
| 399 | 92 | abajo | 438 | 83 | acerca | | | |
| 400 | 92 | estos | 439 | 83 | montón | 478 | 77 | |
| 401 | 92 | tuve | | | | | | lentamente |
| 402 | 91 | chico | 440 | 83 | solo | 479 | 77 | llegó |
| 403 | 91 | di | 441 | 82 | allá | 480 | 77 | loco |
| 404 | 91 | dólares | 442 | 82 | papel | 481 | 77 | meses |
| 405 | 91 | hablando | 443 | 82 | | 482 | 77 | oí |
| 406 | 90 | mirada | | | realmente | 483 | 77 | sitio |
| 407 | 90 | pensaba | 444 | 82 | tenían | 484 | 77 | stradlater |
| | | | 445 | 81 | azul | 485 | 77 | vestido |
| 408 | 90 | sonrisa | 446 | 81 | debe | 486 | 77 | vuelta |
| 409 | 90 | vaya | 447 | 81 | dolores | 487 | 77 | último |
| 410 | 89 | baño | 448 | 81 | ejemplo | 488 | 76 | colega |
| 411 | 89 | chino | | | | 489 | 76 | dado |
| 412 | 89 | cuello | 449 | 81 | jodido | 490 | 76 | gritó |

| | | | | | | | | |
|-----|----|-----------|-----|----|-----------|-----|----|------------|
| 491 | 76 | lleva | 530 | 70 | poder | 572 | 64 | llena |
| 492 | 76 | modos | 531 | 70 | polvo | 573 | 64 | lágrimas |
| 493 | 76 | paso | 532 | 70 | sentó | | | |
| 494 | 76 | tenemos | 533 | 70 | ve | 574 | 64 | |
| | | | 534 | 69 | debajo | | | muchacha |
| 495 | 76 | tono | 535 | 69 | delante | 575 | 64 | na |
| 496 | 75 | comida | 536 | 69 | dianne | 576 | 64 | nuestras |
| 497 | 75 | duda | 537 | 69 | dirección | | | |
| 498 | 75 | fuerza | 538 | 69 | especie | 577 | 64 | perdido |
| 499 | 75 | instante | 539 | 69 | frío | | | |
| | | | 540 | 69 | hacerlo | 578 | 64 | seguir |
| 500 | 75 | venga | 541 | 69 | señorita | 579 | 64 | única |
| 501 | 75 | ése | | | | 580 | 63 | aquello |
| 502 | 74 | an | 542 | 68 | alto | 581 | 63 | caballo |
| 503 | 74 | cierta | 543 | 68 | hablaba | 582 | 63 | dejado |
| 504 | 74 | jugar | | | | 583 | 63 | echó |
| 505 | 73 | bolsillo | 544 | 68 | media | 584 | 63 | jaco |
| | | | 545 | 68 | miedo | 585 | 63 | joder |
| 506 | 73 | estas | 546 | 68 | pensando | 586 | 63 | luna |
| 507 | 73 | horrible | 547 | 68 | seis | 587 | 63 | minutos |
| | | | 548 | 67 | aun | | | |
| 508 | 73 | lago | 549 | 67 | capullos | 588 | 63 | ocurrió |
| 509 | 73 | pub | | | | 589 | 63 | oído |
| 510 | 73 | supe | 550 | 67 | ganas | 590 | 63 | razón |
| 511 | 72 | espalda | 551 | 67 | gris | 591 | 63 | rojo |
| | | | 552 | 67 | malo | 592 | 62 | cielo |
| 512 | 72 | | 553 | 67 | personas | 593 | 62 | conejos |
| | | estábamos | | | | | | |
| 513 | 72 | horas | 554 | 67 | problema | 594 | 62 | empieza |
| 514 | 72 | mark | 555 | 67 | éste | | | |
| 515 | 72 | pasó | 556 | 66 | encontrar | 595 | 62 | hacen |
| 516 | 72 | primero | 557 | 66 | segunda | 596 | 62 | |
| | | | | | | | | pantalones |
| 517 | 71 | dientes | 558 | 66 | sigue | 597 | 62 | película |
| 518 | 71 | mala | 559 | 66 | sola | | | |
| 519 | 71 | miraba | 560 | 66 | sueño | 598 | 62 | pienso |
| 520 | 71 | posible | 561 | 66 | verde | 599 | 62 | pocos |
| 521 | 71 | sala | 562 | 65 | acababa | 600 | 62 | |
| 522 | 71 | según | | | | | | suavemente |
| 523 | 71 | será | 563 | 65 | ello | 601 | 62 | viaje |
| 524 | 71 | tenido | 564 | 65 | estuviera | 602 | 61 | entrar |
| 525 | 70 | adelante | 565 | 65 | grande | 603 | 61 | gustaba |
| | | | 566 | 65 | menor | | | |
| 526 | 70 | decirle | 567 | 65 | niñas | 604 | 61 | johnny |
| 527 | 70 | harold | 568 | 65 | pecho | 605 | 61 | llama |
| 528 | 70 | lleno | 569 | 65 | solía | 606 | 61 | nínfula |
| 529 | 70 | número | 570 | 64 | acercó | 607 | 61 | odio |
| | | | 571 | 64 | interior | 608 | 61 | oír |

| | | | | | | | | |
|-----|----|----------|-----|----|---------------|-----|----|-------------|
| 609 | 61 | resulta | 655 | 58 | viento | 695 | 55 | |
| 610 | 61 | teléfono | 656 | 57 | baja | | | muchísimo |
| | | | 657 | 57 | bar | 696 | 54 | brillante |
| 611 | 61 | tema | 658 | 57 | blanca | | | |
| 612 | 61 | tuvo | 659 | 57 | busca | 697 | 54 | buenas |
| 613 | 61 | vuelve | 660 | 57 | carta | 698 | 54 | cabrones |
| 614 | 61 | york | 661 | 57 | chicas | 699 | 54 | crooks |
| 615 | 60 | ackley | 662 | 57 | conocido | 700 | 54 | habla |
| 616 | 60 | culpa | 663 | 57 | encontré | 701 | 54 | inglés |
| 617 | 60 | cuánto | | | | 702 | 54 | jane |
| 618 | 60 | darle | 664 | 57 | escaleras | 703 | 54 | mar |
| 619 | 60 | estilo | 665 | 57 | | 704 | 54 | pueden |
| 620 | 60 | haya | | | esperando | 705 | 54 | rich |
| 621 | 60 | leer | 666 | 57 | falta | 706 | 54 | rodillas |
| 622 | 60 | llorar | 667 | 57 | haría | | | |
| 623 | 60 | pareció | 668 | 57 | hijos | 707 | 54 | salió |
| 624 | 60 | suave | 669 | 57 | premio | 708 | 54 | |
| 625 | 60 | venters | 670 | 56 | abierta | | | simplemente |
| 626 | 60 | ésa | 671 | 56 | autobús | 709 | 54 | sombra |
| 627 | 59 | barra | | | | 710 | 54 | sonido |
| 628 | 59 | bebé | 672 | 56 | carne | 711 | 53 | ambos |
| 629 | 59 | gracia | 673 | 56 | chaqueta | 712 | 53 | club |
| 630 | 59 | gritos | 674 | 56 | | 713 | 53 | coche |
| 631 | 59 | hermanas | | | conversación | 714 | 53 | destino |
| 632 | 59 | hoy | 675 | 56 | huang | 715 | 53 | londres |
| 633 | 59 | llamar | 676 | 56 | juntos | 716 | 53 | |
| 634 | 59 | llegado | 677 | 56 | llevó | | | matrimonio |
| 635 | 59 | llegué | 678 | 56 | mil | 717 | 53 | mes |
| 636 | 59 | madera | 679 | 56 | miré | 718 | 53 | pena |
| 637 | 59 | mirar | 680 | 56 | resto | 719 | 53 | pone |
| 638 | 59 | oscuro | 681 | 56 | sentada | 720 | 53 | ropa |
| 639 | 59 | papá | | | | 721 | 53 | sally |
| 640 | 59 | pasaba | 682 | 56 | suficiente | 722 | 53 | serio |
| 641 | 59 | veinte | 683 | 56 | ted | 723 | 52 | beardsley |
| 642 | 58 | caer | 684 | 56 | verano | 724 | 52 | cabo |
| 643 | 58 | capaz | 685 | 55 | borde | 725 | 52 | chicos |
| 644 | 58 | cartas | 686 | 55 | cincuenta | 726 | 52 | deja |
| 645 | 58 | dormir | 687 | 55 | | 727 | 52 | |
| 646 | 58 | entró | | | completamente | | | dormitorio |
| 647 | 58 | general | 688 | 55 | d | 728 | 52 | especial |
| 648 | 58 | os | 689 | 55 | deseo | | | |
| 649 | 58 | quedó | 690 | 55 | doctor | 729 | 52 | nota |
| 650 | 58 | silla | 691 | 55 | fondo | 730 | 52 | notaba |
| 651 | 58 | stevie | 692 | 55 | fácil | 731 | 52 | partido |
| 652 | 58 | tampoco | 693 | 55 | lector | 732 | 52 | ponía |
| | | | 694 | 55 | | 733 | 51 | adonde |
| 653 | 58 | the | | | muchacho | 734 | 51 | atención |
| 654 | 58 | viejos | | | | | | |

| | | | | | | | | |
|-----|----|------------|-----|----|---------|-----|----|-------------|
| 735 | 51 | cojones | 743 | 51 | s | 752 | 50 | esperaba |
| | | | 744 | 51 | semanas | | | |
| 736 | 51 | colegas | | | | 753 | 50 | |
| | | | 745 | 51 | taxi | | | exactamente |
| 737 | 51 | cuidado | 746 | 51 | vino | 754 | 50 | fuese |
| | | | 747 | 50 | café | 755 | 50 | nina |
| 738 | 51 | daño | 748 | 50 | cambio | 756 | 50 | |
| 739 | 51 | dejé | 749 | 50 | conocía | | | oportunidad |
| 740 | 51 | drogas | | | | 757 | 50 | propio |
| 741 | 51 | | 750 | 50 | difícil | 758 | 50 | tuviera |
| | | importante | 751 | 50 | espera | 759 | 50 | varios |
| 742 | 51 | mucha | | | | 760 | 50 | vio |

ANEXO 5

PALABRAS CON FRECUENCIA MAYOR O IGUAL A 50 DEL CORPUS LIT_IT_ES

| | | | | | | | | |
|----|------|-----------|----|-----|---------|-----|-----|------------|
| 1 | 5829 | se | 44 | 444 | algo | 87 | 317 | gran |
| 2 | 3430 | lo | 45 | 443 | poco | 88 | 316 | libros |
| 3 | 2644 | me | 46 | 433 | otra | 89 | 313 | cómo |
| 4 | 1624 | le | 47 | 431 | otro | 90 | 310 | esto |
| 5 | 1572 | si | 48 | 429 | sido | 91 | 309 | menos |
| 6 | 1535 | más | 49 | 427 | ni | 92 | 307 | haber |
| 7 | 1462 | era | 50 | 427 | sí | 93 | 306 | mejor |
| 8 | 1220 | ya | 51 | 422 | otros | 94 | 306 | modo |
| 9 | 1184 | dijo | 52 | 421 | aquella | 95 | 305 | tiene |
| 10 | 1074 | sobre | 53 | 420 | así | 96 | 302 | hace |
| 11 | 984 | cuando | 54 | 415 | decir | 97 | 301 | tenía |
| 12 | 942 | pereira | 55 | 408 | siempre | 98 | 289 | quizá |
| 13 | 928 | estaba | 56 | 407 | aquí | 99 | 288 | casa |
| 14 | 889 | también | 57 | 406 | mismo | 100 | 288 | sea |
| 15 | 822 | todo | 58 | 397 | uno | 101 | 287 | mundo |
| 16 | 748 | entre | 59 | 396 | parte | 102 | 283 | ver |
| 17 | 746 | guillermo | 60 | 394 | ese | 103 | 282 | todas |
| 18 | 741 | todos | 61 | 392 | antes | 104 | 278 | nunca |
| 19 | 705 | ahora | 62 | 386 | cada | 105 | 273 | parece |
| 20 | 701 | sólo | 63 | 386 | tan | 106 | 271 | tú |
| 21 | 659 | después | 64 | 379 | tanto | 107 | 268 | casi |
| 22 | 652 | dos | 65 | 378 | hecho | 108 | 267 | palabras |
| 23 | 629 | bien | 66 | 373 | tiempo | 109 | 263 | durante |
| 24 | 627 | esta | 67 | 360 | hacer | 110 | 262 | esa |
| 25 | 626 | te | 68 | 358 | sino | 111 | 262 | luego |
| 26 | 618 | nos | 69 | 355 | dicho | 112 | 261 | mano |
| 27 | 584 | este | 70 | 353 | ojos | 113 | 260 | bajo |
| 28 | 566 | usted | 71 | 353 | quien | 114 | 260 | iglesia |
| 29 | 548 | muy | 72 | 350 | puede | 115 | 259 | allí |
| 30 | 541 | vez | 73 | 348 | noche | 116 | 259 | incluso |
| 31 | 503 | libro | 74 | 344 | e | 117 | 258 | mucho |
| 32 | 498 | nada | 75 | 342 | día | 118 | 256 | sabe |
| 33 | 498 | padre | 76 | 339 | don | 119 | 256 | tal |
| 34 | 488 | mientras | 77 | 336 | mí | 120 | 255 | señor |
| 35 | 479 | hay | 78 | 334 | vida | 121 | 254 | estaban |
| 36 | 470 | donde | 79 | 331 | aún | 122 | 254 | quién |
| 37 | 467 | momento | 80 | 326 | años | 123 | 251 | ante |
| 38 | 464 | fue | 81 | 326 | hombre | 124 | 249 | príncipe |
| 39 | 464 | ser | 82 | 323 | sé | 125 | 247 | |
| 40 | 463 | entonces | 83 | 322 | dios | | | biblioteca |
| 41 | 462 | aquel | 84 | 322 | verdad | 126 | 247 | otras |
| 42 | 461 | eso | 85 | 319 | eran | 127 | 243 | tres |
| 43 | 458 | cosas | 86 | 318 | abad | 128 | 242 | cabeza |

| | | | | | | | | |
|-----|-----|-----------|-----|-----|------------|-----|-----|--------------|
| 129 | 242 | saber | 173 | 186 | mal | 214 | 157 | manera |
| 130 | 240 | podía | 174 | 185 | estos | 215 | 156 | sabía |
| 131 | 239 | toda | 175 | 184 | aquellos | 216 | 156 | éste |
| 132 | 236 | voz | 176 | 184 | cambio | 217 | 154 | maestro |
| 133 | 234 | muchos | 177 | 184 | estas | 218 | 154 | novela |
| 134 | 233 | habia | 178 | 184 | mujer | 219 | 154 | través |
| 135 | 231 | tancredi | 179 | 184 | poder | 220 | 152 | dentro |
| 136 | 229 | | 180 | 184 | rostro | 221 | 152 | estar |
| | | demasiado | 181 | 183 | caso | 222 | 151 | tierra |
| 137 | 227 | alguien | 182 | 183 | nuestra | 223 | 150 | hubiera |
| 138 | 223 | contra | 183 | 183 | sostiene | 224 | 150 | leer |
| 139 | 221 | dice | 184 | 182 | creo | 225 | 150 | paso |
| 140 | 221 | pues | 185 | 181 | demás | 226 | 150 | persona |
| 141 | 221 | rossi | 186 | 179 | doctor | 227 | 150 | podría |
| 142 | 219 | cuenta | 187 | 179 | habría | 228 | 150 | tener |
| 143 | 219 | fuera | 188 | 174 | embargo | 229 | 149 | debía |
| 144 | 218 | visto | | | | 230 | 149 | muchas |
| 145 | 217 | manos | 189 | 172 | calquier | | | |
| 146 | 216 | nadie | 190 | 172 | nombre | 231 | 148 | alguna |
| 147 | 215 | mesa | 191 | 171 | admás | 232 | 148 | dije |
| 148 | 215 | orden | 192 | 171 | hubiese | 233 | 148 | palabra |
| 149 | 215 | tarde | 193 | 171 | papa | 234 | 148 | pronto |
| 150 | 213 | da | 194 | 170 | cierto | 235 | 147 | acaso |
| 151 | 210 | punto | 195 | 170 | días | 236 | 147 | joven |
| 152 | 209 | monteiro | 196 | 170 | quiere | 237 | 147 | |
| 153 | 207 | veces | 197 | 168 | les | | | precisamente |
| 154 | 205 | cosa | 198 | 168 | | 238 | 146 | forma |
| 155 | 203 | gaetano | | | malaquías | 239 | 144 | vi |
| | | | 199 | 167 | os | 240 | 143 | cuatro |
| 156 | 203 | monjes | 200 | 166 | estado | 241 | 143 | hora |
| 157 | 202 | cual | 201 | 165 | historia | 242 | 143 | jorge |
| 158 | 199 | hablar | | | | 243 | 143 | pasado |
| 159 | 199 | quería | 202 | 165 | lado | 244 | 142 | puedo |
| 160 | 199 | senor | 203 | 164 | | 245 | 142 | puesto |
| 161 | 198 | hacía | | | habitación | 246 | 142 | |
| 162 | 197 | hombres | 204 | 164 | primera | | | respondio |
| | | | | | | 247 | 141 | algunos |
| 163 | 194 | duda | 205 | 163 | luz | 248 | 141 | amor |
| 164 | 194 | hizo | 206 | 163 | solo | 249 | 141 | cardoso |
| 165 | 194 | todavía | 207 | 162 | fabrizio | 250 | 140 | misma |
| 166 | 193 | cuerpo | | | | 251 | 139 | cuanto |
| 167 | 193 | mañana | 208 | 161 | lugar | 252 | 139 | idea |
| | | | 209 | 161 | pregunto | 253 | 139 | ubertino |
| 168 | 192 | nuevo | 210 | 160 | | 254 | 138 | sentido |
| 169 | 188 | parecía | | | berengario | 255 | 138 | último |
| 170 | 188 | puerta | 211 | 159 | junto | 256 | 137 | querido |
| 171 | 187 | abadía | 212 | 159 | tengo | 257 | 136 | decía |
| 172 | 187 | muerte | 213 | 157 | debe | 258 | 136 | sitio |

| | | | | | | | | |
|-----|-----|-----------|-----|-----|-----------|-----|----|--------------|
| 259 | 135 | bernardo | 306 | 116 | ir | 352 | 99 | adelmo |
| 260 | 135 | dado | 307 | 116 | scalambri | 353 | 99 | |
| 261 | 135 | personas | 308 | 115 | escribir | | | encontrado |
| 262 | 134 | aire | 309 | 115 | sangre | 354 | 99 | fuese |
| 263 | 134 | cillerero | 310 | 114 | ciudad | 355 | 99 | pareció |
| 264 | 134 | director | 311 | 114 | posible | 356 | 98 | aquellas |
| 265 | 134 | severino | 312 | 114 | salina | | | |
| 266 | 133 | gente | 313 | 114 | según | 357 | 98 | llegar |
| 267 | 133 | hoy | 314 | 113 | esas | 358 | 98 | pies |
| 268 | 132 | angelica | 315 | 112 | bencio | 359 | 98 | propia |
| 269 | 132 | tampoco | 316 | 112 | edificio | 360 | 98 | pueblo |
| 270 | 130 | muerto | | | | 361 | 98 | trabajo |
| 271 | 129 | claro | 317 | 112 | pie | 362 | 97 | abajo |
| 272 | 128 | cristo | 318 | 111 | adso | 363 | 97 | cara |
| 273 | 127 | fondo | 319 | 111 | digo | 364 | 97 | |
| 274 | 127 | llegado | 320 | 111 | efecto | | | franciscanos |
| 275 | 127 | ludmilla | 321 | 110 | horas | 365 | 97 | imagen |
| 276 | 127 | puso | 322 | 110 | pobre | 366 | 97 | tienen |
| 277 | 126 | agua | 323 | 109 | bastante | 367 | 97 | ve |
| 278 | 126 | algún | 324 | 109 | grandes | 368 | 96 | detrás |
| 279 | 126 | cocina | 325 | 109 | | 369 | 96 | entrar |
| 280 | 126 | et | | | muchacha | 370 | 95 | ello |
| 281 | 126 | haya | 326 | 109 | oído | 371 | 95 | iba |
| 282 | 126 | primero | 327 | 108 | michele | 372 | 95 | largo |
| | | | 328 | 108 | siquiera | 373 | 95 | medio |
| 283 | 125 | lengua | 329 | 107 | amigo | 374 | 95 | quisiera |
| 284 | 125 | venancio | 330 | 107 | ei | 375 | 94 | escrito |
| 285 | 124 | alma | 331 | 107 | pensar | 376 | 94 | gracias |
| 286 | 124 | dio | 332 | 107 | primer | 377 | 94 | pregunté |
| 287 | 124 | razón | 333 | 107 | seguro | 378 | 93 | cerca |
| 288 | 124 | seguida | 334 | 106 | sol | 379 | 93 | |
| 289 | 123 | esos | 335 | 106 | ti | | | encuentro |
| 290 | 123 | mirada | 336 | 105 | cinco | 380 | 93 | |
| 291 | 122 | dónde | 337 | 104 | conchetta | | | realmente |
| 292 | 122 | san | 338 | 104 | santo | 381 | 92 | carne |
| 293 | 122 | tenido | 339 | 103 | buen | 382 | 92 | fray |
| 294 | 122 | viejo | 340 | 103 | cuya | 383 | 92 | fuego |
| 295 | 121 | lectura | 341 | 103 | encima | 384 | 92 | mayor |
| 296 | 120 | buena | 342 | 103 | podido | 385 | 92 | tono |
| 297 | 119 | páginas | 343 | 103 | silencio | 386 | 91 | ejemplo |
| 298 | 119 | quiero | 344 | 102 | bueno | 387 | 91 | |
| 299 | 118 | sería | 345 | 102 | quienes | | | habitaciones |
| 300 | 117 | alrededor | 346 | 101 | lector | 388 | 91 | hotel |
| 301 | 117 | despues | 347 | 101 | mujeres | 389 | 91 | miedo |
| 302 | 117 | fin | 348 | 101 | tras | 390 | 91 | pueden |
| 303 | 117 | preguntó | 349 | 100 | cielo | 391 | 91 | vino |
| 304 | 117 | sabes | 350 | 100 | frente | 392 | 90 | marta |
| 305 | 116 | boca | 351 | 100 | salvatore | 393 | 90 | mente |

| | | | | | | | | |
|-----|----|-------------|-----|----|---------------|-----|----|--------------|
| 394 | 90 | tuvo | 438 | 82 | vio | 482 | 72 | cierta |
| 395 | 89 | | 439 | 82 | ésta | 483 | 72 | igual |
| | | comisario | 440 | 81 | calogero | 484 | 72 | justo |
| 396 | 89 | dar | 441 | 81 | debería | 485 | 72 | |
| 397 | 89 | fueron | 442 | 81 | delante | | | necesidad |
| 398 | 89 | fuerza | 443 | 81 | monje | 486 | 71 | dejado |
| 399 | 89 | lisboa | 444 | 81 | tiempos | 487 | 71 | habla |
| 400 | 89 | | 445 | 81 | volvió | 488 | 71 | realidad |
| | | naturaleza | 446 | 80 | color | 489 | 71 | sonrisa |
| 401 | 89 | | 447 | 80 | | 490 | 70 | blanco |
| | | scriptorium | | | conversación | 491 | 70 | hechos |
| 402 | 88 | acuerdo | 448 | 80 | general | 492 | 70 | lámpara |
| 403 | 88 | autor | 449 | 80 | partes | 493 | 70 | menudo |
| 404 | 88 | hablaba | 450 | 80 | será | 494 | 70 | problema |
| 405 | 88 | llama | 451 | 79 | caballo | 495 | 70 | sensación |
| 406 | 88 | pasar | 452 | 79 | escalera | 496 | 70 | signos |
| 407 | 88 | propio | 453 | 79 | negro | 497 | 70 | suelo |
| 408 | 88 | siguiente | 454 | 79 | ninguna | 498 | 70 | viento |
| 409 | 88 | vista | 455 | 78 | | 499 | 69 | diablo |
| 410 | 87 | algunas | | | comprendí | 500 | 69 | ido |
| 411 | 87 | recuerdo | 456 | 78 | periodico | 501 | 69 | regla |
| 412 | 87 | torno | 457 | 78 | salir | 502 | 69 | última |
| 413 | 86 | camino | 458 | 77 | allá | 503 | 68 | añadió |
| 414 | 86 | especie | 459 | 77 | leído | 504 | 68 | cardenal |
| 415 | 86 | final | 460 | 77 | ventanas | 505 | 68 | cuyo |
| 416 | 86 | lejos | 461 | 76 | dejar | 506 | 68 | hijo |
| 417 | 85 | corazón | 462 | 76 | tipo | 507 | 68 | obra |
| 418 | 85 | daba | 463 | 75 | dinero | 508 | 68 | pobreza |
| 419 | 85 | oh | 464 | 75 | encontrar | 509 | 68 | risa |
| 420 | 85 | penso | 465 | 75 | hacerlo | 510 | 67 | basta |
| 421 | 85 | sombra | 466 | 75 | texto | 511 | 67 | carta |
| 422 | 85 | va | 467 | 75 | tienes | 512 | 67 | |
| 423 | 84 | dulcino | 468 | 75 | vuestra | | | donna fugata |
| 424 | 84 | pirrone | 469 | 75 | único | 513 | 67 | pensaba |
| 425 | 84 | santa | 470 | 74 | dedos | 514 | 67 | pueda |
| 426 | 84 | segundo | 471 | 74 | miró | 515 | 67 | quizá |
| 427 | 84 | época | 472 | 74 | objeto | 516 | 67 | quizás |
| 428 | 83 | capaz | 473 | 74 | presencia | 517 | 67 | razones |
| 429 | 83 | cuales | 474 | 74 | puedes | 518 | 67 | veo |
| 430 | 83 | derecha | 475 | 74 | | 519 | 66 | aspecto |
| 431 | 83 | deseo | | | respondió | 520 | 66 | hacen |
| 432 | 83 | oro | 476 | 74 | sala | 521 | 66 | interior |
| 433 | 82 | apenas | 477 | 74 | tenemos | 522 | 66 | llevaba |
| 434 | 82 | debo | 478 | 73 | | 523 | 66 | nueva |
| 435 | 82 | | | | bibliotecario | 524 | 66 | página |
| | | excelencia | 479 | 73 | buenos | 525 | 66 | sentía |
| 436 | 82 | juan | 480 | 73 | instante | 526 | 66 | simples |
| 437 | 82 | rey | 481 | 72 | brazo | 527 | 65 | diez |

| | | | | | | | | |
|-----|----|--------------|-----|----|-----------------|-----|----|----------------|
| 528 | 65 | escritor | 573 | 61 | hierbas | 618 | 56 | amigos |
| 529 | 65 | gesto | 574 | 61 | labios | 619 | 56 | aviñón |
| 530 | 65 | herejes | 575 | 61 | mío | 620 | 56 | buenas |
| 531 | 65 | memoria | 576 | 61 | ningún | 621 | 56 | comer |
| 532 | 65 | pared | 577 | 61 | novelas | 622 | 56 | coro |
| 533 | 65 | siglos | 578 | 61 | ocasión | 623 | 56 | dicen |
| 534 | 65 | tenia | 579 | 61 | piedra | 624 | 56 | encuentra |
| 535 | 65 | trata | 580 | 61 | redaccion | 625 | 56 | figura |
| 536 | 64 | antonio | 581 | 61 | retrato | 626 | 56 | hermanos |
| 537 | 64 | | 582 | 61 | secreto | 627 | 56 | imágenes |
| | | curiosidad | 583 | 61 | vacío | 628 | 56 | oficio |
| 538 | 64 | haciendo | 584 | 61 | viaje | 629 | 56 | pelo |
| 539 | 64 | hija | 585 | 60 | anterior | 630 | 56 | perdido |
| 540 | 64 | mira | 586 | 60 | atención | 631 | 56 | pocos |
| 541 | 64 | seguir | 587 | 60 | | 632 | 56 | seis |
| 542 | 64 | sueño | | | encontraba | 633 | 56 | señorita |
| 543 | 64 | tenían | 588 | 60 | espíritu | 634 | 56 | sola |
| 544 | 64 | trataba | 589 | 60 | falta | 635 | 56 | tambien |
| 545 | 63 | alto | 590 | 60 | fe | 636 | 55 | animales |
| 546 | 63 | cena | 591 | 60 | habrá | 637 | 55 | brazos |
| 547 | 63 | | 592 | 60 | olor | 638 | 55 | centro |
| | | emperador | 593 | 60 | pecado | 639 | 55 | conocido |
| 548 | 63 | habéis | 594 | 59 | año | 640 | 55 | |
| 549 | 63 | llegó | 595 | 59 | dia | | | cualquiera |
| 550 | 63 | mar | 596 | 59 | grande | 641 | 55 | di |
| 551 | 63 | meses | 597 | 59 | marana | 642 | 55 | folios |
| 552 | 63 | parecían | 598 | 59 | palacio | 643 | 55 | frases |
| 553 | 63 | paredes | 599 | 59 | sentado | 644 | 55 | |
| 554 | 63 | pasos | 600 | 59 | siento | | | inmediatamente |
| 555 | 63 | pregunta | 601 | 58 | ayer | 645 | 55 | italia |
| 556 | 63 | ventana | 602 | 58 | buscar | 646 | 55 | llega |
| 557 | 63 | ves | 603 | 58 | esperaba | 647 | 55 | nuestras |
| 558 | 63 | volver | 604 | 58 | grupo | 648 | 55 | oscuridad |
| 559 | 62 | diciendo | 605 | 58 | herejía | 649 | 55 | quieres |
| 560 | 62 | | 606 | 58 | madre | 650 | 55 | relato |
| | | espirituales | 607 | 58 | principio | 651 | 55 | siendo |
| 561 | 62 | existe | 608 | 57 | | 652 | 55 | sucedido |
| 562 | 62 | frase | | | acontecimientos | 653 | 55 | sucedió |
| 563 | 62 | ideas | 609 | 57 | entrado | 654 | 55 | vos |
| 564 | 62 | mil | 610 | 57 | espejo | 655 | 55 | única |
| 565 | 62 | ministro | 611 | 57 | expresión | 656 | 54 | arriba |
| 566 | 62 | policía | 612 | 57 | hablado | 657 | 54 | comida |
| 567 | 62 | relación | 613 | 57 | importa | 658 | 54 | contrario |
| 568 | 62 | sigue | 614 | 57 | | 659 | 54 | manana |
| 569 | 62 | veía | | | muchacho | 660 | 54 | nicola |
| 570 | 61 | cuello | 615 | 57 | piel | 661 | 54 | número |
| 571 | 61 | difícil | 616 | 57 | placer | 662 | 54 | pensé |
| 572 | 61 | espera | 617 | 57 | tantos | 663 | 54 | vamos |

| | | | | | | | | |
|-----|----|--------------|-----|----|------------|-----|----|------------|
| 664 | 53 | atrás | 682 | 52 | juego | 700 | 50 | |
| 665 | 53 | cocinero | 683 | 52 | leyendo | | | diferentes |
| 666 | 53 | descubrir | 684 | 52 | | 701 | 50 | espacio |
| 667 | 53 | estuviese | | | manuscrito | 702 | 50 | esperar |
| 668 | 53 | in | 685 | 52 | pecho | 703 | 50 | francisco |
| 669 | 53 | izquierda | 686 | 52 | quedado | 704 | 50 | hermosa |
| 670 | 53 | | 687 | 52 | salvo | 705 | 50 | hombros |
| | | movimiento | 688 | 51 | ah | 706 | 50 | infierno |
| 671 | 53 | obras | 689 | 51 | | 707 | 50 | jóvenes |
| 672 | 53 | | | | comprender | 708 | 50 | llegada |
| | | pensamientos | 690 | 51 | cree | 709 | 50 | muertos |
| 673 | 53 | pesar | 691 | 51 | | | | |
| 674 | 53 | siglo | | | momentos | 710 | 50 | par |
| 675 | 53 | torreón | 692 | 51 | necesario | 711 | 50 | princesa |
| 676 | 53 | últimos | 693 | 51 | orgullo | 712 | 50 | queda |
| 677 | 52 | cadáver | 694 | 51 | signo | 713 | 50 | |
| 678 | 52 | cuarto | 695 | 50 | asi | | | solamente |
| 679 | 52 | dejo | 696 | 50 | cabo | 714 | 50 | vuelto |
| 680 | 52 | derecho | 697 | 50 | causa | | | |
| 681 | 52 | | 698 | 50 | conmigo | | | |
| | | impresión | 699 | 50 | cuál | | | |

ANEXO 6

PRIMERAS 100 PALABRAS DE LA LISTA DE PALABRAS CLAVE DEL CORPUS DE LITERATURA TRADUCIDA LIT_ES

| rank | freq | keyness | keyword | rank | freq | keyness | keyword |
|------|------|-----------|------------|------|-------|-----------|--------------|
| 1 | 7884 | 7.854.219 | me | 38 | 139 | 1.409.346 | humbert |
| 2 | 942 | 6.602.906 | pereira | 39 | 139 | 1.409.346 | ubertino |
| 3 | 4156 | 4.472.773 | mi | 40 | 134 | 1.358.650 | cillerero |
| 4 | 346 | 3.508.156 | lennie | 41 | 134 | 1.358.650 | severino |
| 5 | 318 | 3.224.259 | abad | 42 | 132 | 1.338.372 | angelica |
| 6 | 746 | 3.103.348 | guillermo | 43 | 129 | 1.307.954 | slim |
| 7 | 274 | 2.778.135 | lolita | 44 | 127 | 1.287.676 | ludmilla |
| 8 | 3828 | 2.773.862 | había | 45 | 125 | 1.267.397 | venancio |
| 9 | 265 | 2.686.883 | boy | 46 | 124 | 1.257.258 | curley |
| 10 | 238 | 2.413.125 | sick | 47 | 122 | 1.236.980 | rents |
| 11 | 233 | 2.362.429 | habia | 48 | 120 | 1.216.702 | phoebe |
| 12 | 232 | 2.352.290 | capullo | 49 | 117 | 1.186.284 | despues |
| 13 | 231 | 2.342.151 | tancredi | 50 | 116 | 1.176.145 | charlotte |
| 14 | 221 | 2.240.759 | rossi | 51 | 116 | 1.176.145 | scalambri |
| 15 | 219 | 2.220.480 | renton | 52 | 114 | 1.155.866 | salina |
| 16 | 217 | 2.200.202 | cabrón | 53 | 112 | 1.135.588 | bencio |
| 17 | 2050 | 2.153.912 | estaba | 54 | 111 | 1.125.449 | adso |
| 18 | 209 | 2.119.089 | begbie | 55 | 111 | 1.125.449 | tommy |
| 19 | 209 | 2.119.089 | monteiro | 56 | 108 | 1.095.031 | michele |
| 20 | 205 | 2.078.532 | monjes | 57 | 107 | 1.084.892 | ei |
| 21 | 203 | 2.058.253 | gaetano | 58 | 988 | 1.082.582 | madre |
| 22 | 2177 | 2.025.613 | dijo | 59 | 106 | 1.074.753 | dolly |
| 23 | 199 | 2.017.697 | senor | 60 | 104 | 1.054.475 | concetta |
| 24 | 3367 | 1.983.202 | era | 61 | 103 | 1.044.336 | candy |
| 25 | 195 | 1.977.140 | spud | 62 | 1146 | 1.016.891 | he |
| 26 | 187 | 1.896.027 | abadía | 63 | 100 | 1.013.918 | billy |
| 27 | 449 | 1.746.616 | george | 64 | 100 | 1.013.918 | salvatore |
| 28 | 168 | 1.703.382 | malaquías | 65 | 99 | 1.003.779 | adelmo |
| 29 | 5285 | 1.645.548 | pero | 66 | 99 | 1.003.779 | tíos |
| 30 | 162 | 1.642.547 | fabrizio | 67 | 453 | 998.540 | tío |
| 31 | 160 | 1.622.269 | berengario | 68 | 97 | 983.500 | franciscanos |
| 32 | 158 | 1.601.990 | puto | 69 | 96 | 973.361 | matty |
| 33 | 2604 | 1.577.143 | yo | 70 | 940 | 968.138 | mis |
| 34 | 148 | 1.500.599 | haze | 71 | 1736 | 909.871 | Te |
| 35 | 645 | 1.479.037 | dije | 72 | 89 | 902.387 | mei |
| 36 | 142 | 1.439.764 | respondio | 73 | 89 | 902.387 | scriptorium |
| 37 | 141 | 1.429.624 | cardoso | 74 | 12295 | 888.933 | no |
| | | | | 75 | 86 | 871.969 | Jo |
| | | | | 76 | 85 | 861.830 | penso |

| | | | | | | | |
|----|------|---------|---------------|-----|-----|---------|--------------|
| 77 | 84 | 851.691 | dulcino | 90 | 70 | 709.743 | harold |
| 78 | 84 | 851.691 | pirrone | 91 | 69 | 699.603 | dianne |
| 79 | 82 | 831.413 | monje | 92 | 68 | 689.464 | capullos |
| 80 | 81 | 821.274 | calogero | 93 | 67 | 679.325 | donna fugata |
| 81 | 81 | 821.274 | jodido | 94 | 67 | 679.325 | quiza |
| 82 | 78 | 790.856 | an | 95 | 770 | 675.393 | verdad |
| 83 | 78 | 790.856 | periodico | 96 | 66 | 669.186 | nicola |
| 84 | 77 | 780.717 | stradlater | 97 | 846 | 660.151 | mí |
| 85 | 1113 | 762.145 | tenía | 98 | 65 | 659.047 | herejes |
| 86 | 2231 | 752.304 | qué | 99 | 65 | 659.047 | tenia |
| 87 | 74 | 750.299 | pub | 100 | 64 | 648.908 | na |
| 88 | 73 | 740.160 | bibliotecario | | | | |
| 89 | 4123 | 734.636 | le | | | | |

ANEXO 7

**PRIMERAS 100 PALABRAS DEL LISTADO DE PALABRAS CLAVE DEL
CORPUS DE LITERATURA TRADUCIDA DEL INGLÉS LIT_EN_ES**

| rank | freq | keyness | keyword | rank | freq | keyness | keyword |
|------|------|-----------|-----------|------|------|---------|------------|
| | | | | 38 | 1110 | 940.875 | te |
| | | | | 39 | 81 | 935.729 | jodido |
| 1 | 5240 | 7.567.603 | me | 40 | 2499 | 892.123 | le |
| 2 | 3081 | 5.258.078 | mi | 41 | 77 | 889.521 | stradlater |
| 3 | 346 | 3.997.066 | lennie | 42 | 74 | 854.864 | An |
| 4 | 274 | 3.165.307 | lolita | 43 | 73 | 843.312 | Pub |
| 5 | 264 | 3.049.785 | boy | 44 | 993 | 828.711 | Dijo |
| 6 | 238 | 2.749.427 | sick | 45 | 70 | 808.655 | Harold |
| 7 | 231 | 2.668.562 | capullo | 46 | 204 | 802.902 | Puta |
| 8 | 219 | 2.529.935 | renton | 47 | 69 | 797.103 | dianne |
| 9 | 215 | 2.483.726 | cabrón | 48 | 197 | 779.394 | Eh |
| 10 | 209 | 2.414.413 | begbie | 49 | 67 | 773.998 | capullos |
| 11 | 448 | 2.350.066 | george | 50 | 64 | 739.342 | Na |
| 12 | 195 | 2.252.682 | spud | 51 | 63 | 727.790 | Jaco |
| 13 | 930 | 1.974.593 | madre | 52 | 63 | 727.790 | Joder |
| 14 | 158 | 1.825.250 | puto | 53 | 62 | 716.237 | conejos |
| 15 | 148 | 1.709.728 | haze | 54 | 61 | 704.685 | nínfula |
| 16 | 138 | 1.594.206 | humbert | 55 | 60 | 693.133 | Ackley |
| 17 | 1666 | 1.572.967 | yo | 56 | 60 | 693.133 | venters |
| 18 | 497 | 1.541.423 | dije | 57 | 58 | 670.028 | Stevie |
| 19 | 129 | 1.490.236 | slim | 58 | 56 | 646.924 | Huang |
| 20 | 1905 | 1.453.471 | era | 59 | 56 | 646.924 | Ted |
| 21 | 124 | 1.432.475 | curley | 60 | 54 | 623.820 | cabrones |
| 22 | 122 | 1.409.370 | rents | 61 | 54 | 623.820 | crooks |
| 23 | 120 | 1.386.266 | phoebe | 62 | 54 | 623.820 | Jane |
| 24 | 413 | 1.367.001 | tío | 63 | 54 | 623.820 | Rich |
| 25 | 1122 | 1.358.159 | estaba | 64 | 169 | 616.373 | mierda |
| 26 | 116 | 1.340.057 | charlotte | 65 | 53 | 612.267 | Sally |
| 27 | 1848 | 1.302.711 | había | 66 | 52 | 600.715 | beardsley |
| 28 | 111 | 1.282.296 | tommy | 67 | 52 | 600.715 | notaba |
| 29 | 106 | 1.224.535 | dolly | 68 | 1258 | 600.604 | Qué |
| 30 | 103 | 1.189.878 | candy | 69 | 216 | 592.350 | Tía |
| 31 | 100 | 1.155.221 | billy | 70 | 51 | 589.163 | cojones |
| 32 | 96 | 1.109.013 | matty | 71 | 2405 | 588.527 | Pero |
| 33 | 94 | 1.085.908 | tíos | 72 | 50 | 577.611 | Nina |
| 34 | 89 | 1.028.147 | mei | 73 | 248 | 574.925 | Sabes |
| 35 | 812 | 994.456 | tenía | 74 | 6257 | 564.454 | No |
| 36 | 86 | 993.490 | jo | 75 | 871 | 561.596 | Ella |
| 37 | 630 | 941.715 | mis | 76 | 890 | 559.882 | Eso |

| | | | | | | | |
|----|------|---------|---------|-----|-----|---------|-------------|
| 77 | 48 | 554.506 | Eché | 91 | 43 | 496.745 | hijoputa |
| 78 | 47 | 542.954 | pencey | 92 | 43 | 496.745 | popo |
| 79 | 47 | 542.954 | Ying | 93 | 43 | 496.745 | tsing |
| 80 | 510 | 540.279 | Mí | 94 | 43 | 496.745 | waverly |
| 81 | 1541 | 535.950 | cuando | 95 | 43 | 496.745 | Wu |
| 82 | 46 | 531.402 | Lenny | 96 | 42 | 485.193 | holden |
| 83 | 46 | 531.402 | Senté | 97 | 42 | 485.193 | rollo |
| 84 | 45 | 519.850 | carlson | 98 | 269 | 485.071 | cama |
| 85 | 45 | 519.850 | leith | 99 | 41 | 473.641 | antolini |
| 86 | 44 | 508.297 | jong | 100 | 41 | 473.641 | encantadora |
| 87 | 44 | 508.297 | taitai | | | | |
| 88 | 446 | 501.915 | cabeza | | | | |
| 89 | 448 | 499.372 | verdad | | | | |
| 90 | 431 | 496.924 | podía | | | | |

ANEXO 8

PRIMERAS 100 PALABRAS DE LA LISTA DE PALABRAS CLAVE DEL CORPUS DE LITERATURA TRADUCIDA DEL ITALIANO LIT_IT_ES

| rank | freq | keyness | keyword | 39 | 85 | 976.368 | penso |
|------|------|-----------|--------------|----|------|---------|---------------|
| 1 | 942 | 7.864.689 | pereira | 40 | 84 | 964.881 | dulcino |
| 2 | 746 | 4.075.537 | guillermo | 41 | 84 | 964.881 | pirrone |
| 3 | 318 | 3.652.765 | abad | 42 | 81 | 930.421 | calogero |
| 4 | 233 | 2.676.397 | habia | 43 | 81 | 930.421 | Monje |
| 5 | 231 | 2.653.424 | tancredi | 44 | 78 | 895.961 | periodico |
| 6 | 221 | 2.538.557 | rossi | 45 | 247 | 876.902 | biblioteca |
| 7 | 209 | 2.400.717 | monteiro | 46 | 73 | 838.528 | bibliotecario |
| 8 | 203 | 2.331.797 | gaetano | 47 | 928 | 832.011 | Estaba |
| 9 | 203 | 2.331.797 | monjes | 48 | 67 | 769.608 | donnafugata |
| 10 | 199 | 2.285.850 | senor | 49 | 67 | 769.608 | Quiza |
| 11 | 187 | 2.148.010 | abadía | 50 | 503 | 765.872 | Libro |
| 12 | 168 | 1.929.763 | malaquías | 51 | 680 | 752.447 | He |
| 13 | 162 | 1.860.843 | fabrizio | 52 | 249 | 748.058 | príncipe |
| 14 | 160 | 1.837.869 | berengario | 53 | 65 | 746.634 | herejes |
| 15 | 142 | 1.631.109 | respondio | 54 | 65 | 746.634 | Tenia |
| 16 | 141 | 1.619.622 | cardoso | 55 | 61 | 700.688 | redaccion |
| 17 | 139 | 1.596.649 | ubertino | 56 | 59 | 677.714 | Dia |
| 18 | 134 | 1.539.216 | cillerero | 57 | 59 | 677.714 | marana |
| 19 | 134 | 1.539.216 | severino | 58 | 58 | 666.228 | herejía |
| 20 | 132 | 1.516.242 | angelica | 59 | 56 | 643.254 | aviñón |
| 21 | 1980 | 1.485.231 | había | 60 | 56 | 643.254 | tambien |
| 22 | 127 | 1.458.809 | ludmilla | 61 | 161 | 636.361 | pregunto |
| 23 | 125 | 1.435.835 | venancio | 62 | 55 | 631.768 | Folios |
| 24 | 117 | 1.343.942 | despues | 63 | 54 | 620.281 | manana |
| 25 | 116 | 1.332.455 | scalambri | 64 | 54 | 620.281 | Nicola |
| 26 | 114 | 1.309.482 | salina | 65 | 1462 | 611.921 | Era |
| 27 | 112 | 1.286.509 | bencio | 66 | 566 | 610.124 | Usted |
| 28 | 2644 | 1.275.816 | me | 67 | 53 | 608.794 | cocinero |
| 29 | 111 | 1.275.022 | adso | 68 | 53 | 608.794 | torreón |
| 30 | 108 | 1.240.562 | michele | 69 | 52 | 597.308 | manuscrito |
| 31 | 107 | 1.229.075 | ei | 70 | 316 | 577.985 | Libros |
| 32 | 1184 | 1.217.924 | dijo | 71 | 50 | 574.334 | Asi |
| 33 | 104 | 1.194.615 | concetta | 72 | 183 | 562.392 | sostiene |
| 34 | 100 | 1.148.668 | salvatore | 73 | 48 | 551.361 | apóstoles |
| 35 | 99 | 1.137.182 | adelmo | 74 | 48 | 551.361 | pagina |
| 36 | 97 | 1.114.208 | franciscanos | 75 | 48 | 551.361 | refectorio |
| 37 | 2880 | 1.091.888 | pero | 76 | 48 | 551.361 | remigio |
| 38 | 89 | 1.022.315 | scriptorium | 77 | 47 | 539.874 | alinardo |

| | | | | | | | |
|----|-----|---------|----------|-----|-----|---------|------------|
| 78 | 47 | 539.874 | Frailles | 91 | 42 | 482.441 | aqui |
| 79 | 46 | 528.387 | palermo | 92 | 42 | 482.441 | limonada |
| 80 | 462 | 522.661 | Aquel | 93 | 135 | 474.920 | bernardo |
| 81 | 498 | 518.235 | Padre | 94 | 41 | 470.954 | articulo |
| 82 | 45 | 516.901 | lectora | 95 | 41 | 470.954 | sabéis |
| 83 | 45 | 516.901 | Sicilia | 96 | 40 | 459.467 | entro |
| 84 | 421 | 514.366 | aquella | 97 | 40 | 459.467 | finis |
| 85 | 44 | 505.414 | ademas | 98 | 40 | 459.467 | fraticelli |
| 86 | 44 | 505.414 | flannery | 99 | 40 | 459.467 | hereje |
| 87 | 44 | 505.414 | paolo | 100 | 40 | 459.467 | policia |
| 88 | 43 | 493.927 | cafe | | | | |
| 89 | 43 | 493.927 | ciccio | | | | |
| 90 | 43 | 493.927 | dias | | | | |

ANEXO 9

PRIMERAS 100 PALABRAS DE LA LISTA DE PALABRAS CLAVE DEL CORPUS DE OBRAS ORIGINALES EN LENGUA INGLÉS

| rank | freq | keyness | keyword |
|------|-------|------------|---------|
| 1 | 5272 | 72.261.431 | was |
| 2 | 2765 | 37.898.872 | had |
| 3 | 2016 | 27.632.596 | said |
| 4 | 1776 | 24.343.001 | is |
| 5 | 1514 | 20.751.860 | tae |
| 6 | 4410 | 20.666.939 | s |
| 7 | 2361 | 19.848.362 | ah |
| 8 | 1116 | 15.296.615 | were |
| 9 | 965 | 13.226.912 | fuckin |
| 10 | 903 | 12.377.100 | ay |
| 11 | 11442 | 10.836.294 | i |
| 12 | 731 | 10.019.557 | did |
| 13 | 674 | 9.238.279 | been |
| 14 | 648 | 8.881.906 | got |
| 15 | 590 | 8.086.920 | cunt |
| 16 | 584 | 8.004.680 | didn |
| 17 | 548 | 7.511.241 | ye |
| 18 | 3996 | 7.305.885 | my |
| 19 | 457 | 6.263.937 | wis |
| 20 | 452 | 6.195.403 | wi |
| 21 | 449 | 6.154.283 | went |
| 22 | 444 | 6.085.750 | looked |
| 23 | 407 | 5.578.604 | ve |
| 24 | 397 | 5.441.538 | oan |
| 25 | 744 | 5.269.255 | ma |
| 26 | 367 | 5.030.339 | eyes |
| 27 | 367 | 5.030.339 | lennie |
| 28 | 363 | 4.975.512 | made |
| 29 | 4173 | 4.959.522 | she |
| 30 | 356 | 4.879.566 | came |
| 31 | 3938 | 4.803.855 | her |
| 32 | 347 | 4.756.206 | told |
| 33 | 345 | 4.728.792 | knew |
| 34 | 340 | 4.660.259 | aw |
| 35 | 726 | 4.648.837 | are |
| 36 | 2781 | 4.594.785 | me |
| 37 | 316 | 4.331.300 | going |
| 38 | 310 | 4.249.060 | has |

| | | | |
|----|-----|-----------|---------|
| 39 | 296 | 4.057.167 | asked |
| 40 | 296 | 4.057.167 | saw |
| 41 | 293 | 4.016.047 | started |
| 42 | 287 | 3.933.807 | things |
| 43 | 282 | 3.865.274 | looking |
| 44 | 274 | 3.755.621 | fir |
| 45 | 659 | 3.710.479 | don |
| 46 | 266 | 3.645.967 | years |
| 47 | 255 | 3.495.194 | likesay |
| 48 | 255 | 3.495.194 | took |
| 49 | 252 | 3.454.074 | am |
| 50 | 249 | 3.412.955 | sais |
| 51 | 247 | 3.385.541 | oot |
| 52 | 245 | 3.358.128 | goddam |
| 53 | 237 | 3.248.475 | felt |
| 54 | 236 | 3.234.768 | lolita |
| 55 | 233 | 3.193.648 | used |
| 56 | 231 | 3.166.235 | wanted |
| 57 | 224 | 3.070.288 | renton |
| 58 | 218 | 2.988.049 | about |
| 59 | 216 | 2.960.635 | turned |
| 60 | 212 | 2.905.809 | kept |
| 61 | 209 | 2.864.689 | jist |
| 62 | 207 | 2.837.275 | guys |
| 63 | 205 | 2.809.862 | begbie |
| 64 | 203 | 2.782.449 | says |
| 65 | 198 | 2.713.916 | git |
| 66 | 197 | 2.700.209 | seen |
| 67 | 194 | 2.659.089 | spud |
| 68 | 194 | 2.659.089 | ya |
| 69 | 193 | 2.645.382 | wasn |
| 70 | 188 | 2.576.849 | hands |
| 71 | 743 | 2.515.713 | n |
| 72 | 177 | 2.426.076 | gave |
| 73 | 170 | 2.330.130 | called |
| 74 | 169 | 2.316.423 | done |
| 75 | 165 | 2.261.596 | yir |
| 76 | 163 | 2.234.183 | seemed |
| 77 | 163 | 2.234.183 | thit |
| 78 | 161 | 2.206.770 | getting |
| 79 | 159 | 2.179.357 | auntie |
| 80 | 158 | 2.165.650 | intae |
| 81 | 157 | 2.151.943 | fae |
| 82 | 354 | 2.099.169 | people |

| | | | |
|-----|-----|-----------|---------|
| 83 | 150 | 2.055.997 | heard |
| 84 | 149 | 2.042.290 | haze |
| 85 | 148 | 2.028.583 | doon |
| 86 | 147 | 2.014.877 | days |
| 87 | 144 | 1.973.757 | looks |
| 88 | 141 | 1.932.637 | stood |
| 89 | 140 | 1.918.930 | cunts |
| 90 | 138 | 1.891.517 | times |
| 91 | 862 | 1.859.868 | mother |
| 92 | 135 | 1.850.397 | curley |
| 93 | 135 | 1.850.397 | trying |
| 94 | 135 | 1.850.397 | walked |
| 95 | 133 | 1.822.984 | does |
| 96 | 132 | 1.809.277 | shite |
| 97 | 130 | 1.781.864 | hud |
| 98 | 129 | 1.768.157 | auld |
| 99 | 129 | 1.768.157 | gone |
| 100 | 127 | 1.740.744 | humbert |

ANEXO 10

PRIMERAS 100 PALABRAS DE LA LISTA DE PALABRAS CLAVE EN EL CORPUS DE OBRAS ORIGINALES EN LENGUA ITALIANA

| rank | freq | keyness | keyword | 40 | 231 | 240.932 | chiese |
|------|------|-----------|------------|----|-----|---------|---------------|
| 1 | 1991 | 3.969.530 | - | 41 | 746 | 239.136 | lui |
| 2 | 937 | 1.868.131 | pereira | 42 | 669 | 237.446 | Erano |
| 3 | 1178 | 1.477.178 | disse | 43 | 224 | 234.770 | rispose |
| 4 | 755 | 1.423.798 | guglielmo | 44 | 227 | 232.111 | biblioteca |
| 5 | 2581 | 1.038.789 | era | 45 | 115 | 229.280 | Bencio |
| 6 | 525 | 785.959 | don | 46 | 115 | 229.280 | Salina |
| 7 | 305 | 608.090 | c'è | 47 | 304 | 228.048 | Libri |
| 8 | 1169 | 600.073 | aveva | 48 | 111 | 221.305 | Adso |
| 9 | 331 | 566.936 | dottor | 49 | 111 | 221.305 | scalambri |
| 10 | 2237 | 473.655 | mi | 50 | 589 | 217.976 | Forse |
| 11 | 229 | 456.566 | tancredi | 51 | 109 | 217.317 | un'altra |
| 12 | 227 | 452.578 | l'abate | 52 | 107 | 213.330 | d'una |
| 13 | 310 | 427.478 | piú | 53 | 105 | 209.342 | conchetta |
| 14 | 212 | 422.672 | monteiro | 54 | 156 | 208.291 | Dissi |
| 15 | 192 | 382.797 | monaci | 55 | 183 | 204.995 | sostiene |
| 16 | 1471 | 373.036 | perché | 56 | 101 | 201.367 | adelmo |
| 17 | 246 | 371.213 | principe | 57 | 126 | 197.972 | Vidi |
| 18 | 734 | 370.195 | lei | 58 | 471 | 194.637 | Tu |
| 19 | 181 | 360.866 | po' | 59 | 888 | 190.154 | Ne |
| 20 | 171 | 340.929 | malachia | 60 | 250 | 186.609 | Fece |
| 21 | 199 | 330.912 | gaetano | 61 | 93 | 185.418 | Mo |
| 22 | 234 | 321.437 | rossi | 62 | 91 | 181.430 | Marta |
| 23 | 150 | 299.061 | jorge | 63 | 87 | 173.455 | Frate |
| 24 | 148 | 295.073 | berengario | 64 | 87 | 173.455 | scriptorium |
| 25 | 771 | 286.439 | quel | 65 | 86 | 171.461 | dolcino |
| 26 | 141 | 281.117 | cardoso | 66 | 84 | 167.474 | pirrone |
| 27 | 497 | 280.439 | libro | 67 | 198 | 166.226 | voleva |
| 28 | 138 | 275.136 | ubertino | 68 | 83 | 165.480 | l'altro |
| 29 | 137 | 273.142 | c'era | 69 | 81 | 161.493 | calogero |
| 30 | 136 | 271.148 | bernardo | 70 | 368 | 157.788 | So |
| 31 | 134 | 267.161 | cellario | 71 | 148 | 148.685 | romanzo |
| 32 | 130 | 259.186 | angelica | 72 | 436 | 147.069 | Fosse |
| 33 | 130 | 259.186 | severino | 73 | 349 | 145.114 | Occhi |
| 34 | 128 | 255.198 | d'un | 74 | 297 | 144.175 | Padre |
| 35 | 306 | 254.606 | stava | 75 | 71 | 141.555 | bibliotecario |
| 36 | 3113 | 253.895 | ma | 76 | 70 | 139.562 | l'ho |
| 37 | 126 | 251.211 | ludmilla | 77 | 150 | 138.157 | maestro |
| 38 | 161 | 244.855 | fabrizio | 78 | 108 | 137.923 | Dì |
| 39 | 122 | 243.236 | venanzio | 79 | 67 | 133.580 | signorina |

| | | | |
|-----|-------|---------|-------------|
| 80 | 1020 | 132.346 | Io |
| 81 | 66 | 131.587 | donnafugata |
| 82 | 157 | 128.279 | Volto |
| 83 | 88 | 127.744 | Parve |
| 84 | 172 | 127.714 | papa |
| 85 | 63 | 125.605 | eretici |
| 86 | 102 | 121.411 | salvatore |
| 87 | 1868 | 119.861 | lo |
| 88 | 60 | 119.624 | apostoli |
| 89 | 60 | 119.624 | zione |
| 90 | 91 | 118.342 | cosí |
| 91 | 59 | 117.630 | erbe |
| 92 | 57 | 113.643 | chiesi |
| 93 | 57 | 113.643 | delitti |
| 94 | 318 | 113.159 | avrebbe |
| 95 | 11490 | 110.875 | che |
| 96 | 641 | 110.490 | me |
| 97 | 55 | 109.656 | d'oro |
| 98 | 98 | 109.079 | pensò |
| 99 | 79 | 107.706 | monaco |
| 100 | 54 | 107.662 | lisboa |

ÍNDICE DE LAS TABLAS

| | |
|--|-----|
| Tabla 1: clasificación de culturemas | 33 |
| Tabla 2: clasificación de técnicas de traducción de culturemas | 52 |
| Tabla 3: comparación lingüística de corpus y lingüística racionalista | 58 |
| Tabla 4: presentación corpus LIT_ENIT_ES | 87 |
| Tabla 5: correspondencias corpus paralelo inglés-español | 89 |
| Tabla 6: correspondencias corpus paralelo italiano-español | 89 |
| Tabla 7: presentación corpus LIT_ES | 90 |
| Tabla 8: presentación SUBcorpus LIT_EN_ES | 92 |
| Tabla 9: presentación subcorpus LIT_IT_ES | 93 |
| Tabla 10: clasificación culturemas LIT_ES | 106 |
| Tabla 11: categoría nombres propios | 109 |
| Tabla 12: culturemas más utilizados en LIT_ES | 109 |
| Tabla 13: subcategorías más utilizadas en LIT_ES | 110 |
| Tabla 14: culturemas más utilizados en LIT_EN_ES | 111 |
| Tabla 15: culturemas más utilizados en LIT_IT_ES | 111 |
| Tabla 16: subcategorías más utilizadas en LIT_EN_ES | 112 |
| Tabla 17: subcategorías más utilizadas en LIT_IT_ES | 113 |
| Tabla 18: culturema más frecuentes en LIT_ES | 116 |
| Tabla 19: categorías culturemas más frecuentes LIT_ES..... | 130 |
| Tabla 20: Gráfico visual culturemas más frecuentes en el corpus d literatura traducida LIT_ES | 132 |
| Tabla 21: cambio de categoría culturemas en LIT_ES | 133 |

| | |
|---|-----|
| Tabla 22: cultuemas más frecuentes en LIT_EN_ES | 137 |
| Tabla 23: categorías de cultuemas más frecuentes en LIT_EN_ES | 144 |
| Tabla 24: Gráfico visual cultuemas más frecuentes en el corpus d literatura traducida LIT_EN_ES | 146 |
| Tabla 25: cambio de categoría cultuemas LIT_EN_ES | 147 |
| Tabla 26: cultuemas más frecuentes en LIT_IT_ES | 149 |
| Tabla 27: categorías de cultuemas más frecuentes en LIT_IT_ES | 153 |
| Tabla 28: gráfico visual cultuemas más frecuentes cultuemas literatura traducida del italiano al español | 154 |
| Tabla 29: cambio de categoría cultuemas LIT_IT_ES | 156 |
| Tabla 30: técnicas de traducción literatura traducida LIT_ES, LIT_EN_ES y LIT_IT_ES | 162 |
| Tabla 31: técnica de traducción inequivalencia en los corpus de literatura traducida LIT_ES, LIT_EN_ES y LIT_IT_ES | 164 |
| Tabla 32: técnicas de traducción más utilizadas en los subcorpus LIT_EN_ES y LIT_IT_ES | 165 |
| Tabla 33: cultuemas clave en LIT_ES | 166 |
| Tabla 34: cultuemas clave en LIT_EN_ES | 168 |
| Tabla 35: cultuemas claves en LIT_IT_ES | 170 |
| Tabla 36: cultuemas clave obras originales en lengua inglés | 172 |
| Tabla 37: : cultuemas clave obras originales en lengua italiana | 173 |
| Tabla 38: comparación cultuemas clave literatura traducida y obras originales | 176 |

ÍNDICE DE LOS ANEXOS

| | |
|---|-----|
| Anexos 1: <i>Stop list</i> utilizada para eliminar las palabras gramaticales de las listas de palabras analizadas | 186 |
| Anexo 2: Elementos culturales eliminados después del análisis de los cultuemas en su contexto..... | 187 |
| Anexo 3: Palabras con frecuencia mayor o igual a 50 del corpus LIT_ES..... | 189 |
| Anexo 4: Palabras con frecuencia mayor o igual a 50 del corpus LIT_EN_ES..... | 202 |
| Anexo 5: Palabras con frecuencia mayor o igual a 50 del corpus LIT_IT_ES..... | 209 |
| Anexo 6: Primeras 100 palabras de la lista de palabras clave del corpus de literatura traducida LIT_ES..... | 215 |
| Anexo 7: Primeras 100 palabras de la lista de palabras clave del corpus de literatura traducida del inglés LIT_EN_ES..... | 217 |
| Anexo 8: Primeras 100 palabras de la lista de palabras clave del corpus de literatura traducida del italiano LIT_EN_ES..... | 219 |
| Anexo 9: Primeras 100 palabras de la lista de palabras clave del corpus de obras originales en lengua inglés..... | 221 |
| Anexo 10: Primeras 100 palabras de la lista de palabras clave del corpus de obras originales en lengua italiana..... | 224 |

BIBLIOGRAFÍA

ALCINA CAUDET, A. Y A. ESTELLES PALANCA, (2006) «Etiquetado del corpus txtceram orientado a la extracción de información conceptual» en *Actas del XXIV congreso Aprendizaje de lenguas, uso del lenguaje y modelación cognitiva: perspectivas aplicadas entre disciplinas*, UNED, Madrid.

BAKER, M., (1995): «Corpora in Translation Studies: An Overview and Some Suggestions for Future Research», *Target* 7-2.

BEEBY, A. Y OTROS (2000): *Investigating Translation: Selected papers from the 4th International Congress on Translation, Barcelona, 1998*, John Benjamins Publishing, Amsterdam.

BERBER SARDINHA, T. (2000): «Lingüística de corpus: histórico y problemática», *D.E.L.T.A.*, 16-2, 2000

BOWKER, L. (2002): *Computer-aided translation technology : a practical introduction*, University of Ottawa Press, Ottawa.

BOWKER, L. Y J. PEARSON, (2002): *Working with Specialized Language: A Practical Guide to Using Corpora*, Routledge, London.

CHESSA, S. (2011): *Il concetto di culturema: traduzione di HIJOS DEL MEDIODÍA DE EVA DÍAZ PÉREZ*, facoltà di Lingue e Letterature Straniere, Università degli Studi di Sassari en <http://issuu.com/110elode/docs/sara.chessa> (última visita el día 30/10/2014).

CORPAS PASTOR, G. (2001): «Compilación de un corpus ad hoc para la enseñanza de la traducción inversa especializada», *Trans: Revista de traductología*, 5.

COSERIU, E. (1981): *Lecciones de lingüística general*, Gredos, Madrid.

GANDIN, S. (2009): «Linguistica dei corpora e traduzione: definizioni, criteri di compilazione e implicazioni di ricerca dei corpora paralleli», *AnnalSS*, 5, 2005 (2009).

GARCÍA IZQUIERDO, I. (2005): «Corpus electrónico, género textual y traducción: metodología, concepto y ámbito de la Enciclopedia electrónica para traductores GENTT», *Meta : journal des traducteurs / Meta: Translators' Journal*, 50-4, 2005. En <http://id.erudit.org/iderudit/019918ar> (última visita en 06/08/14) (última visita el día 06/08/14).

HATIM, B. Y I. MASON (1995): *Teoría de la traducción: una aproximación al discurso*, Ariel, Barcelona.

KATAN, D. (1999): *Translating culture: an introduction for translators, interpreters and mediators*, St Jerome, Manchester.

LAVIOSA, S. (2002): *Corpus-based Translation Studies: Theory, Findings, Applications*, Rodopi, Amsterdam, New York.

LEECH, G. (1992): «Corpora and theories of linguistic performance», en J. SVARTVIK (org.) *Directions in Corpus Linguistics. Proceedings of Nobel Symposium 82, Stockholm, 4-8 August 1991*. De Gruyter, Berlin, New York.

LÉPINETTE, B. *La historia de la traducción. Metodología. Apuntes bibliográficos*. *HISTAL* 2004, en <http://www.google.it/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0CB8QFjAA&url=http%3A%2F%2Fwww.histal.ca%2Fwp->

[content%2Fuploads%2F2011%2F08%2FLa-historia-de-la-traduccion-metodologia-
apuntes-
bibliograficos.pdf&ei=H_QbVOCXDenmywOt7YKADA&usg=AFQjCNGJTgp8WJA
Wz0M3i_vey1WPG22WVQ&bvm=bv.75774317,d.bGQ](#) (última visita el día
19/09/2014).

LÓPEZ-MEZQUITA MOLINA, M. T. (2007): *La evaluación de la competencia léxica: Test de vocabulario. Su fiabilidad y validez*, Ministerio de Educación y Ciencia, Madrid.

MALMKJAER, K. (1998) « Love thy Neighbour : Will Parallel Corpora Endear Linguists to Translators? » *Meta : journal des traducteurs / Meta: Translators' Journal*, 43-4, 1998.

MANGIRON I HEVIA, C. (2006): *El tractament dels referents culturals a les traduccions de la novel·la Botxan: la interacció entre els elements textuais i extratextuais*, tesis doctoral, Universitat Autònoma de Barcelona, Departamento de Traducción e Interpretación, Barcelona. En <http://hdl.handle.net/10803/5270> (última visita el día 30/10/2014).

MAYORAL ASENSIO, R. (1994): «La explicitación de la información en la traducción intercultural», en Hurtado A. (ed.) *Estudis sobre la traducció*, Publicacions de la Universitat Jaume I, Castellón de la Plana.

MERCADO PERCIA, H. (2008): «Fundamentos de la lingüística de corpus» en [http://aprendeenlinea.udea.edu.co/lms/moodle/mod/resource/view.php?inpopup=true&i
d=44650](http://aprendeenlinea.udea.edu.co/lms/moodle/mod/resource/view.php?inpopup=true&i
d=44650) (última visita en fecha 06/08/14)

MOLINA MARTÍNEZ, L. (2006): *El otoño del pingüino*, Publicacions de la Universitat Jaume I, Castellón de la Plana.

MOLINA MARTÍNEZ, L., (2001): *Análisis descriptivo de la traducción de los culturemas árabe-español*, Universitat Autònoma de Barcelona, Departamento de Traducción e Interpretación, Barcelona. En <http://hdl.handle.net/10803/5263> (última visita el día 30/10/2014)

MONZÓ I NEBOT, E. (2003): «Las socializaciones del traductor especializado: el papel de los géneros», *Revista de la Facultad de Lenguas Modernas*, 6, Universidad Ricardo Palma, Lima.

NADAL, L. L. (2009): «Los culturemas: ¿unidades lingüísticas, ideológicas o culturales?», *Language Design*, 11.

NEWMARK, P. (1988): *A textbook of translation*, Prentice Hall, New York.

NIDA, A. (1945): «Linguistics and Ethnology in Translation Problems», *Word*, 1.

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

NORD, C. (1997): *Translating as a purposeful activity*, St Jerome, Manchester

PROCHÁZKOVÁ, P. (2006): «Fundamentos de la lingüística de corpus. Concepción de los corpus y métodos de investigación con Corpus», en [www.prochazkova.de/fundamentos de la ling%C3%BC%C3%ADstica de corpus.pdf](http://www.prochazkova.de/fundamentos_de_la_ling%C3%BC%C3%ADstica_de_corpus.pdf) (última visita el día 06/08/14)

ROBINSON, G. L. (1988): *Crosscultural Understanding*, Prentice Hall, New York.

ROJO, G., (2002): Sobre la Lingüística basada en el análisis de corpus, en UZEI *Hizkuntza-corporak. Oraina eta geroa* (2002-10-24/25), en <http://uzei.com/Modulos/UsuariosFtp/Conexion/archivos54A.pdf> (última visita el día 06/08/14)

RUIZ PI, S. (2013): *Los referentes culturales y su tratamiento en la traducción del manga - Los casos de La espada del inmortal, GALS!, Vidas étlicas y Ranma ½*, Universitat Autònoma de Barcelona, Departamento de Traducción e Interpretación, Barcelona. En <http://hdl.handle.net/2072/227488> (última visita el día 30/10/14)

SAMOVAR, L. y otros (2014 [1997]): *Intercultural Communication: A Reader (14th ed.)*, Cengage Learning, Boston.

SANCHEZ, A. (1995): «Definición e historia de los corpus», en SANCHEZ, A. y otros (org.). *CUMBRE – Corpus Linguístico de Español Contemporáneo*, SGEL (Sociedad General Española de Librería), Madrid

SANTAMARIA 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 30/10/2014).

SINCLAIR, J. (1991): *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

SINCLAIR, J. (2003): *Reading Concordances: An Introduction*, Pearson/Longman, London.

SORIA ORTI, M. (2012): *Análisis traductológico de referentes culturales en La Testa Perduta di Damasceno Monteiro de Antonio Tabucchi*, Universitat Autònoma de Barcelona, Departamento de Traducción e Interpretación, Barcelona.

TANQUEIRO, H. (2001): *Autotradução: autoridade, privilegio e modelo*, tesis doctoral, Universitat Autònoma de Barcelona, Departamento de Traducción e Interpretación, Barcelona. En www.tdx.cat/bitstream/10803/5259/.../ht1de3.pdf (última visita el día 30/10/2014).

TOGNINI –BONELLI, E. (2001): *Corpus Linguistics at Work*, John Benjamins, Amsterdam y Philadelphia.

TRIVEDI, H. (2005): «Translation culture vs. cultural translation», the International Writing Program, University of Iowa on-line journal *91st Meridian*, vol. 4-1 (mayo 2005) en <http://iwp.uiowa.edu/91st/vol4-num1/translating-culture-vs-cultural-translation> (última visita el día 27/10/2014).

TRUJILLO, V.C. (2012): «Una aportación al tratamiento lingüístico de los elementos culturales: el signo lingüístico cultural», *Çédille revista de estudios franceses*, 8. En <http://www.google.it/url?sa=t&rct=j&q=&esrc=s&source=web&cd=10&ved=0CGAQFjAJ&url=http%3A%2F%2Fcedille.webs.ull.es%2F8%2F18trujillo.pdf&ei=3LZUVJuDJ o7gatjwgMgF&usg=AFQjCNHNqanwrpahqAaSNiMiX6pGUnsJNw&bvm=bv.78677474,d.d2s&cad=rja> (última visita el día 31/10/2014).

TYLOR, E. B. (1977): *Cultura primitiva*, Ayuso, Madrid.

VÁZQUEZ-AYORA, G. (1977): *Introducción a la traductología: curso básico de traducción*, Georgetown University Press, Washington D.C.

VINAY, J.P. y J. DARBELNET (1995): *Comparative stylistics of French and English: a methodology for translation* John Benjamins, Amsterdam.

WHORF, B. L. (1971): *Lenguaje, pensamiento, realidad*, Barral, Barcelona.

WU, S. H. (2013): *Traducción y recepción de la subtitulación chino-español. Análisis de la cultura lingüística como referente cultural*, tesis doctoral, Universitat Autònoma de Barcelona, Departamento de Traducción e Interpretación en <http://hdl.handle.net/10803/120546> (última visita el día 30/10/2014).

XIAO, R. Y Y. MING (2009): «Using corpora in translation studies: the state of art», en

BAKER, P. (2012): *Contemporary Corpus linguistics* A&C Black, London.