



Universidad de Valladolid

FACULTAD DE TRADUCCIÓN E INTERPRETACIÓN

Grado en Traducción e Interpretación

TRABAJO FIN DE GRADO

La traducción español-inglés de las fichas descriptivas de chorizo: un estudio de caso basado en el contraste de un corpus paralelo y de un corpus comparable

Presentado por Dña. Ana López Tirado

Tutelado por la Dra. M^a Teresa Ortego-Antón

Soria, junio de 2019

Atiendi, Asturias

*Atiendi, Asturias, atiendi
lo que falo na to fala
anque nos oyíos tengas
muncha llingua castellana.*

*Sé que sofristi abundu
que fusti más nueche qu'alba
que fixisti too dafechu
lo que senties pel alma.*

*Sélo bien, pero'l dolor
golviósete una mordaza
dexándote a fuerza sombra
una llaceria na fala.*

*Esa llaceria esparidióse
como una postiella gafa
fendió to llingua llariega
fasta torgate falala.*

*Y la hestoria que pulsiasti
cola pallabra bien alta
quedando fo callandina
como ensin acordanza.*

*Y ensin acordanza un pueblu
nun tien futuru a la llarga.*

*Atiendi, Asturias, atiendi
lo que falo na to fala.*

Manuel Asur

ÍNDICE

RESUMEN	3
ABSTRACT	3
1. INTRODUCCIÓN	4
1.2. COMPETENCIAS DESARROLLADAS.....	4
1.3. OBJETIVOS DEL ESTUDIO	5
2. MARCO TEÓRICO.....	6
2.1. LA TRADUCCIÓN EN EL SECTOR AGROALIMENTARIO	6
2.2. LOS CORPUS Y LA TRADUCCIÓN.....	7
2.2.1. DEFINICIÓN DE CORPUS	7
2.2.2. TIPOLOGÍA DE CORPUS	8
2.2.3. UTILIDADES DE LOS CORPUS COMO HERRAMIENTAS DE TRADUCCIÓN	10
2.2.4. CARACTERÍSTICAS DE LOS CORPUS <i>AD HOC</i> COMO HERRAMIENTAS DE TRADUCCIÓN	12
3. METODOLOGÍA	14
3.1. INTRODUCCIÓN.....	14
3.2. REVISIÓN Y ALIMENTACIÓN DE P-GEFEM	14
3.3. LA MUESTRA DE ANÁLISIS	16
3.4. LOS EQUIVALENTES Y SU FRASEOLOGÍA	17
3.5. EXPLOTACIÓN DE C-GEFEM Y COMPARACIÓN LINGÜÍSTICA.....	18
4. ANÁLISIS Y RESULTADOS	20
4.1. RESULTADOS EN P-GEFEM	20
4.1.1. Grasa.....	20
4.1.2. Azúcar	20
4.1.3. Producto.....	21
4.1.4. Información.....	21
4.1.5. Extra	21
4.1.6. Fresco y conservar	21
4.1.7. Saturado	21
4.1.8. Ingrediente y loncha.....	21
4.2. CONTRASTE DE RESULTADOS EN C-GEFEM	22
4.2.1. Proteína.....	22
4.2.2. Grasa.....	23
4.2.3. Hidrato de carbono	23
4.2.4. Información.....	23
4.2.5. Extra	23

4.2.6. Saturado	24
4.2.7. Azúcar y azúcares.....	24
4.2.8. Peso.....	25
4.2.9. Leche.....	25
4.2.10. Loncha	25
4.2.11. Fresco	25
4.2.12. Conservar	26
4.2.13. Valor medio	26
4.2.14. Kilocoría, gramo, kilogramo	26
4.2.15. Producto, peso neto, ingrediente.....	26
4.3. LA BASE DE DATOS TERMINOLÓGICA	26
5. CONCLUSIONES	28
6. BIBLIOGRAFÍA.....	30

ÍNDICE DE FIGURAS

Figura 1. Hoja de cálculo con información sobre cada archivo .txt y los productos que se muestran en las páginas en español.....	15
Figura 2. Hoja de cálculo con información sobre cada archivo .txt y los productos que se muestran en las páginas en inglés	15
Figura 3. Listado de unidades léxicas más frecuentes en el subcorpus en español de P-GEFEM extraídas con la función Word List de AntConc 3.4.4. (Anthony, 2018)	16
Figura 4. Ejemplo de concordancia de una unidad léxica y su traducción en español y en inglés con ParaConc.....	17
Figura 5. Fragmento de la hoja de cálculo con información terminológica y fraseológica extraída de P-GEFEM	18
Figura 6. Fragmento de la hoja de cálculo utilizada para la recogida de datos para el contraste de resultados.....	19
Figura 7. Ficha terminológica del término «extra».....	27
Figura 8. Ficha terminológica del término «información»	27

ÍNDICE DE TABLAS

Tabla 1. Tamaño de P-GEFEM.....	16
Tabla 2. Comparación de la traducción español-inglés del término milk protein y su uso en los textos originales en inglés	22
Tabla 3. Comparativa de texto original en español con su traducción al inglés y con texto original en inglés	24

RESUMEN

El sector agroalimentario tiene gran presencia en España y es uno de los que más beneficios aporta a la economía de nuestro país. Sin embargo, apenas existen estudios relacionados con la traducción de los textos agroalimentarios (Rivas Carmona y Veron González, 2018; Ortego-Antón, 2019). Por ello, el objetivo de este trabajo es comprobar si la terminología recogida en las fichas descriptivas del chorizo de las principales empresas del sector cárnico se trasvasa al inglés utilizando los mismos equivalentes que se emplean cuando se redactan originalmente en lengua inglesa. Para ello, llevaremos a cabo un análisis lingüístico basado en los datos de un corpus paralelo y de un corpus comparable con una muestra de 20 términos relacionados con el chorizo. Con los resultados que se desprendan de nuestro trabajo, podremos obtener una panorámica del trasvase interlingüístico, así como detectar patrones de comportamiento típicos de las traducciones y de los textos redactados originalmente en lengua inglesa que podrán ser aplicados a la generación de aplicaciones basadas en el procesamiento del lenguaje natural.

Palabras clave: traducción, inglés, español, corpus, terminología, sector agroalimentario.

ABSTRACT

The agrifood sector has a strong presence in Spain and it is one of the most beneficial sectors for the Spanish economy. However, the number of studies related to the translation of agrifood texts is scarce (Rivas Carmona y Veron González, 2018; Ortego-Antón, 2019). For this reason, the purpose of this paper is to determine whether the terminology displayed on the chorizo product sheets of the main companies in the meat industry is transferred into English by using the same equivalents that are used when the texts are written originally in English. We will conduct a linguistic analysis based on the data of parallel and a comparable corpus with a sample of 20 terms related to the chorizo. The findings of our study will allow us to get a panoramic view of the interlinguistic transfer and to detect patterns of behaviour which are typical of the translations or of the texts originally written in English. These patterns could be applied to the creation and development of natural language processing-based applications.

Keywords: translation, English, Spanish, corpora, terminology, agrifood sector.

1. INTRODUCCIÓN

El presente estudio, denominado *La traducción español-inglés de las fichas descriptivas de chorizo: un estudio de caso basado en el contraste de un corpus paralelo y de un corpus comparable*, se encuadra dentro de un trabajo de introducción a la investigación en el marco de una beca de colaboración en tareas de investigación en departamentos L.O.U. otorgada por el Consejo Social de la UVa¹ y que estamos desarrollando en el Departamento de Lengua Española bajo la tutela de la Dra. M^a Teresa Ortego-Antón.

En el marco de la mencionada beca, los objetivos que nos propusimos alcanzar antes del comienzo de dicho proyecto de investigación fueron, entre otros, ser capaz de trabajar dentro de un equipo de investigación, iniciarnos en tareas de investigación y profundizar en la traducción especializada de los términos relativos a la industria cárnica en las lenguas inglesa y española.

Asimismo, la finalidad de estas becas es proporcionar una formación investigadora inicial a los alumnos universitarios para facilitar su futura orientación profesional o investigadora, mediante su participación o colaboración en tareas de investigación, vinculadas con los estudios que están cursando, en un Departamento universitario o Instituto L.O.U. Esta finalidad se pretende cumplir en régimen de compatibilidad con sus estudios, permitiéndose a la vez realizar el Proyecto Fin de Carrera o Trabajo de Fin de Grado (Convocatoria becas Consejo Social de colaboración en tareas de investigación en departamentos e institutos L.O.U. para el curso académico 2018-19).

Por otro lado, uno de los requisitos de dicha beca es que exista una cierta conexión entre la labor a desarrollar y el historial científico-técnico de los candidatos y, en su caso, con el Trabajo de Fin de Grado o Fin de Máster que estén realizando en esta Universidad, de ahí que este trabajo encuentre su base en el estudio que estamos desarrollando en el contexto de la mencionada beca de colaboración.

A partir de lo establecido en la convocatoria, hemos centrado el presente estudio en el análisis de la terminología y fraseología del sector cárnico (concretamente del chorizo) para comprobar si las traducciones que se recogen en las páginas web de algunas empresas españolas se adecúan al contexto anglosajón. Para ello, las compararemos con textos del mismo género y tipología redactados originalmente en inglés con el fin de destacar la importancia de un buen trasvase interlingüístico cuando se ofrece la mencionada información en las páginas web tanto en inglés como en español. Además, este tema nos resulta de gran interés debido a que procedemos de una región cárnica como es Asturias donde, dentro del sector de las chacinas destaca, por encima de todo, un producto típico como el chorizo. Por lo tanto, esta será la temática de este estudio, que se desarrollará en los capítulos siguientes.

1.2. COMPETENCIAS DESARROLLADAS

Con este trabajo pretendemos desarrollar una serie de competencias, que se exponen en el Proyecto docente de la asignatura Trabajo Fin de Grado de la UVa durante el curso 2018/2019², a saber, las competencias generales G1 y G3 y las competencias específicas E5, E8, E17, E18, E19, E35, E36, E37, E39, E40, E47, E50, E51 y E52.

¹ <http://consejosocial.uva.es/>

² https://alojamientos.uva.es/guia_docente/uploads/2018/423/41142/1/Documento.pdf

1.3. OBJETIVOS DEL ESTUDIO

El objetivo general de este estudio es comprobar si la terminología recogida en las fichas descriptivas del chorizo de las principales empresas del sector cárnico se trasvasa al inglés utilizando las mismas denominaciones que cuando se redactan originalmente en lengua inglesa.

Por otro lado, este estudio también presenta unos objetivos más específicos, que se enumeran a continuación:

- Verificar si la traducción de una muestra de términos se corresponde con la denominación empleada al redactar textos originalmente en inglés o, por el contrario, solamente se emplean en las traducciones del español al inglés.
- Poner en práctica los conocimientos adquiridos durante la carrera en asignaturas como Documentación para traductores, Informática aplicada la traducción o Terminología, entre otras, con el fin de desarrollar una metodología basada en corpus que aplicaremos para analizar la muestra de análisis.
- Ser capaz de examinar y diferenciar si un término es producto de la traducción o si está bien empleado en un contexto determinado.
- Aplicar los conocimientos y competencias adquiridos para compilar, explotar y analizar un corpus paralelo y un corpus comparable.
- Analizar el comportamiento léxico y fraseológico de los términos de la muestra de análisis.
- Alimentar una base de datos terminológica con entradas en español y en inglés que recoja los términos de la muestra.
- Exponer los resultados y conclusiones del estudio de manera clara y concisa.

2. MARCO TEÓRICO

En este capítulo presentamos una breve introducción de la situación de la traducción en el sector agroalimentario en España, destacando la importancia de las lenguas como factor esencial para que las empresas extiendan su internacionalización en el mundo globalizado actual. Asimismo, proporcionamos una definición del concepto de «corpus» para comprender uno de los elementos principales en los que hemos basado nuestro estudio y ofrecemos una clasificación de los distintos tipos de corpus atendiendo a sus características fundamentándonos en autores como Torruella y Llisterri (1999) o Corpas Pastor (2001). También describimos los beneficios del uso de los corpus como herramientas de traducción y señalamos las principales características que estos deben cumplir para satisfacer el mencionado uso. La definición y delimitación de estos conceptos serán los pilares en los que se asiente nuestro trabajo.

2.1. LA TRADUCCIÓN EN EL SECTOR AGROALIMENTARIO

El sector agroalimentario abarca tanto el sector primario como el secundario y el terciario, lo que conlleva la producción de un gran volumen de textos, con distintos grados de especialización y pertenecientes a diversos dominios y géneros textuales. Además, este sector cada vez requiere de un mayor número de traductores especializados como consecuencia de la demanda y la consiguiente exportación de productos españoles a cualquier parte del mundo. De hecho, este sector es de gran relevancia en nuestro país, ya que constituye una de las principales actividades impulsoras de la economía española.

Sin embargo, la documentación en este campo del saber es más bien escasa y, por lo que respecta a los estudios sobre la traducción de los textos de este ámbito, el número es todavía menor. En consecuencia, con este trabajo pretendemos introducirnos en la traducción agroalimentaria a través de un enfoque terminológico que pueda ayudar a las empresas de este sector a optimizar la comercialización de sus productos, ya que estas están adquiriendo cada vez más relevancia en el panorama internacional debido al proceso de globalización. Entre los servicios y competencias de las mencionadas empresas destaca la presencia de las mismas en Internet, consecuencia de la creciente digitalización a la que se ven sometidas la mayoría de las actividades hoy en día. Esta digitalización es clave en el proceso de expansión comercial de las empresas. Por ello, vemos necesario que las páginas web con las que las empresas agroalimentarias se muestran a potenciales clientes de cualquier parte del mundo estén disponibles en varias lenguas con el fin de extender el alcance de sus productos y, por supuesto, que las traducciones del español a otras lenguas sean adecuadas al contexto y a la cultura de la lengua de llegada para asegurar que las empresas españolas puedan competir en el mercado global.

Por otro lado, el tema que abordamos será de gran utilidad para dar respuesta a los problemas terminológicos que pueden surgir al traducir textos de este ámbito ya que, en ocasiones, resulta complejo el trasvase de ciertos términos de una lengua a otra y, sobre todo, de una cultura a otra, puesto que estamos ante productos típicos de una determinada área geográfica que pueden resultar desconocidos más allá de nuestras fronteras, como ocurre con el pimentón de La Vera, el chorizo o la morcilla. Ante este tipo de situaciones, como indica Epstein (2009), es necesario conocer bien las distintas estrategias de traducción que se pueden emplear y saber cuál es la más adecuada dependiendo del tipo de dificultad a la que nos enfrentemos con el fin de solventarla.

Por estos motivos, en este trabajo pretendemos abordar el trasvase interlingüístico (del español al inglés) de una muestra de términos agroalimentarios que se utilizan en la redacción de fichas descriptivas de chorizo y de la fraseología de dichos términos, esto es, los equivalentes en lengua inglesa de los términos empleados en español, a partir de una metodología de análisis basada en corpus con el fin último de observar los principales patrones de comportamiento terminológico durante el trasvase interlingüístico del español al inglés.

2.2. LOS CORPUS Y LA TRADUCCIÓN

2.2.1. DEFINICIÓN DE CORPUS

Como ya hemos apuntado, la industria agroalimentaria, aun siendo una de las actividades primordiales en España —en especial en una serie de regiones en las que este sector representa una parte importante de la economía—, no ha merecido especial atención por parte de la investigación de los Estudios de Traducción (Ortego Antón, 2019: 24). Por esta razón y con el objetivo de extender la competencia y el alcance de las empresas dedicadas a la producción de chacinas y, más concretamente, de chorizo, hemos llevado a cabo el presente estudio.

Dado que vamos a basar la metodología en corpus, conviene señalar que no existe una univocidad con respecto a la designación de dicho concepto por parte de los expertos en traducción, lo que nos lleva a encontrar distintas denominaciones para hacer referencia a los corpus *ad hoc*³, es decir, a los corpus diseñados para cubrir unas necesidades concretas y preestablecidas. No obstante, de aquí en adelante, el término que emplearemos será el de corpus *ad hoc*, siguiendo la propuesta de Aston (1999).

Ante definiciones del concepto de corpus poco entusiastas y simples como la ofrecida por Leech (1992: 106): «*On the face of it, a computer corpus is an unexciting phenomenon: a helluva lot of text, stored on a computer*», nos encontramos con otras que ofrecen un mayor número de características. Por ejemplo, Alvar Ezquerro *et al.* (1994: 9-10) proporcionan la siguiente definición de corpus: «Un corpus es un conjunto homogéneo de documentos de cualquier tipo [...] que se toman como modelo de un estado o nivel de la lengua predeterminado, al cual representan o se pretende que representen».

Por otro lado, Sinclair (1996: 4) define corpus de una forma similar: «*A collection of pieces of language that are selected and ordered according to explicit linguistic criteria in order to be used as a sample of the language*».

En esta línea está también la definición ofrecida posteriormente por Sinclair (2005: 1), quien considera que los corpus son conjuntos de textos en formato electrónico seleccionados a partir de una serie de parámetros que representan, en la medida de lo posible, una lengua como fuente de datos para su análisis.

No obstante, estas definiciones corresponden al concepto de corpus en su sentido más general. Los corpus, a su vez, pueden clasificarse en función de una serie de características como el número de lenguas incluidas, la modalidad de la lengua de los

³ Este tipo de corpus (los corpus *ad hoc*) también ha recibido las siguientes denominaciones: corpus virtual (Corpas Pastor, 2012), *disposable corpus*, *do-it-yourself (DIY) corpus* (Zanettin, 2002), *domain-specific corpus* (Corpas Pastor, 2004: 226), *web corpus* (Fletcher, 2004), *electronic corpus* (Corpas Pastor, 2001; Varantola, 2003), *ephemeral corpus* (Corpas Pastor, 2004: 226), *precision corpus* (Varantola, 1997) y *special purpose corpus* (Pearson, 1998; Sánchez-Gijón, 2003a, 2003b), entre otras.

textos o la distribución de los mismos, entre otras, que procedemos a describir en el siguiente epígrafe.

2.2.2. TIPOLOGÍA DE CORPUS

Una de las primeras tipologías de corpus apareció en el informe de Sinclair para EAGLES en 1994. Desde entonces, son múltiples las taxonomías propuestas para establecer una clasificación de los tipos de corpus en función de sus características (Laviosa, 1997; Granger, 2003; Zanettin, 2012; Faya Ornia, 2014, etc.). Ante la multitud de clasificaciones, en este trabajo vamos a basarnos en la que ofrece Corpas Pastor (2001), que recoge en gran parte la propuesta de Torruella y Llisterri (1999).

En primer lugar, es necesario establecer una diferencia entre los corpus escritos y los orales, distinción que viene marcada por la modalidad de la lengua de los textos. Como su nombre indica, los primeros reúnen muestras de lengua escrita y los segundos, de lengua oral, como pueden ser vídeos, grabaciones o transcripciones de grabaciones. Además, existen también los corpus mixtos, que combinan ambas modalidades.

Por otro lado, atendiendo al número de lenguas incluidas en el corpus, existen corpus monolingües, bilingües y multilingües, que estarán compuestos por textos en una, dos o varias lenguas, respectivamente. A su vez, dentro de los corpus bilingües y multilingües, destacan varias categorías, dependiendo de la relación existente entre los textos representativos de las lenguas recogidas. Estos tipos son corpus paralelos y corpus comparables. Danielsson y Ridings (1996: 1) definen corpus paralelo como «*a text in one language that has been translated into one or more other languages. All of these taken together are called "parallel texts", since they, ideally, contain the same information in parallel with each other*». Así pues, si las lenguas de trabajo fuesen español e inglés, este estaría compuesto por textos originales en español y por sus traducciones en inglés (de ahí que también se denomine corpus de traducciones). Según Olohan (2004: 24), los corpus paralelos ofrecen a su vez una subdivisión entre corpus monodireccionales y bidireccionales. Los primeros (monodireccionales) estarán constituidos por los originales en la lengua A y sus traducciones en la lengua B y, en el caso de los segundos (bidireccionales), estos estarán compuestos también por textos originales en la lengua A y sus traducciones en la lengua B pero, además, incluirán también textos originales en la lengua B y sus traducciones en la lengua A. Por lo tanto, si el corpus que se va a compilar es paralelo, debe estar claro si será del primer tipo o del segundo (en este último caso, estaría compuesto por más textos). En cambio, un corpus comparable no incluye traducciones, dado que está constituido por textos originales tanto en español como en inglés, compartiendo estos las mismas características textuales.

En cuanto a la proporción y la distribución de los textos, podemos distinguir entre corpus grandes, equilibrados, piramidales, léxicos, textuales y de referencia. Los corpus grandes tienen una extensión ilimitada; en este tipo de corpus no existe la necesidad de que sea equilibrado, a diferencia de los corpus equilibrados propiamente dichos, que «*contienen diversas variedades de lengua en porcentajes similares*» (Corpas Pastor, 2001: 158). Los corpus piramidales, por su parte, organizan sus textos en niveles y van aumentando de manera progresiva la complejidad de la temática de los mismos a medida que el número de textos va disminuyendo. Los corpus léxicos recogen fragmentos muy cortos de textos y su finalidad es principalmente léxica; los textuales están integrados por textos completos representativos de la lengua común. Finalmente,

los corpus de referencia incluyen solo fragmentos de documentos que representen la variedad de la lengua que se pretende estudiar.

Por otra parte, en función del límite de palabras establecido, existen los corpus cerrados y los abiertos. Para la compilación de los primeros, se establece previamente un límite de palabras y, cuando se alcanza este límite, se da por finalizada la compilación. Por el contrario, los corpus abiertos, también denominados corpus monitores, son corpus dinámicos, que se actualizan constantemente con elementos nuevos, a la vez que se elimina contenido antiguo con el fin de mantener un volumen de textos constante.

Además, el nivel de especificidad de los textos nos permite distinguir entre corpus generales, especializados, genéricos y canónicos. Los primeros son representativos de la lengua común y están integrados por textos que se desarrollan en situaciones comunicativas cotidianas. Los corpus especializados, por el contrario, incluyen textos de variedades especializadas de la lengua, y los corpus canónicos están formados por las obras completas de un determinado autor.

Dependiendo del periodo temporal en el que se desarrollan los textos del corpus, estos pueden ser periódicos o cronológicos, diacrónicos o históricos y sincrónicos. Los periódicos incluyen textos pertenecientes a un determinado periodo de tiempo y los diacrónicos recogen textos de etapas cronológicas sucesivas. Los corpus sincrónicos, por su parte, estudian variedades de la lengua de textos que se desarrollan en el presente para analizar un determinado comportamiento en la actualidad.

Por último, distinguimos, por una parte, corpus anotados o codificados y no anotados o simples, en función de si se han añadido etiquetas o algún tipo de información lingüística a los documentos que lo componen o, por el contrario, no se ha proporcionado ningún dato adicional. Por otra parte, existen los corpus documentados y los no documentados, clasificación que depende de si a cada texto se ha incorporado una explicación con la procedencia del mismo o no.

Ante tal variedad de corpus, en nuestro estudio vamos a utilizar un corpus *ad hoc* (Laviosa, 2010; Zanettin, 2012; Bowker y Pearson, 2002; McEnery y Hardie, 2012) que, como hemos mencionado previamente, corresponde a aquel que se compila a partir de textos extraídos de Internet con el objetivo de resolver necesidades puntuales y concretas.

Además, Zanettin (2002) señala que los corpus *ad hoc* pueden ser tanto comparables como monolingües, y los define como «*a collection of Internet documents or more precisely web pages in HTML format created ad hoc as a response to a specific text to be translated*».

Varantola (2003), que los denomina *disposable corpora*, los define como «*small specialized corpora assembled ad hoc for the purpose of a specific translation*».

Por otra parte, a pesar de que Toruella y Llisterri (1999) no incluyen los corpus *ad hoc* en su clasificación, Corpas Pastor (2004), los define de la siguiente manera: «[un corpus *ad hoc* es] un corpus virtual que se compila puntualmente para la realización de un determinado encargo de traducción en cualquier dirección (directa, inversa o indirecta)».

Por último, Seghiri (2011) amplía la definición precedente:

Un corpus virtual⁴ puede definirse como aquel que se crea exclusivamente a partir de recursos electrónicos disponibles en la red Internet para la realización puntual de un determinado encargo de traducción o documentación de un bloque textual, y cuyo principal objetivo es satisfacer las principales necesidades documentales del traductor de forma económica [...], rápida [...] y fiable.

Tal y como se puede extraer de los párrafos anteriores, los corpus *ad hoc* se adaptan a la definición de corpus pero, además, se caracterizan por haberse compilado para responder a unas cuestiones específicas.

Habiendo desarrollado la definición y una vez descritas las distintas tipologías, vamos a justificar por qué los corpus se han convertido en una herramienta indispensable dentro de los Estudios de Traducción.

2.2.3. UTILIDADES DE LOS CORPUS COMO HERRAMIENTAS DE TRADUCCIÓN

Numerosos autores (Bowker, 1998; Zanettin, 1998; Varantola, 2000; Seghiri, 2011, entre otros) han defendido las ventajas de la utilización de corpus como herramientas de traducción, apoyándose sobre todo en el argumento de que estos son de gran ayuda a la hora de resolver problemas documentales y dificultades de traducción. Por ejemplo, Aston (1999) considera que un corpus «*is compiled 'on the fly' by the translator in order to investigate a specific problem encountered during a particular translation*».

Corpas Pastor (2002: 201), por su parte, se pronuncia en una línea muy similar con respecto al objetivo de los corpus en traducción: «reunir toda la documentación disponible sobre un tema en muy poco tiempo, ya se trate de documentar un único texto o bien de preparar todo un bloque textual». Como vemos, subraya la economía que supone la compilación de corpus en cuanto al tiempo invertido (lo que potenciará la productividad del traductor, un factor esencial en la calidad de su labor) y, a diferencia de Aston, destaca la utilidad de estos no solo para resolver problemas puntuales de un encargo, sino también colectivos.

Sánchez Trigo (2005: 138), vuelve a incidir en la utilidad de los corpus para resolver problemas de traducción e insiste en el ahorro de tiempo que supone su uso:

[Los corpus] constituyen una herramienta interesante para solucionar problemas de diferente naturaleza (temáticos, terminológicos, textuales, estilísticos, etc.) [...] ya que permiten compilar una documentación fiable y específica de manera económica (en tiempo y coste) y muy eficaz.

También son múltiples los investigadores —como Bowker (1998; 1999b) o Zanettin (1998; 2002), entre otros— que argumentan que el uso de corpus favorece el desarrollo de las competencias requeridas en un traductor (competencia traductora, lingüística y textual en las lenguas de origen y de destino, cultural y documental) (Corpas, 2012: 7), así como de la labor docente en el caso de los profesionales académicos (Seghiri, 2017: 45). En este sentido, Corpas Pastor y Seghiri (2009: 77) enuncian varias de las ventajas del uso de corpus:

Some of the principal advantages of using them [corpora] are their objectivity, their reusability and multiple usage as a single resource. In addition, they are user-friendly and allow access and management to huge quantities of information in almost no time (Corpas Pastor y Seghiri, 2009: 77).

⁴ Esta autora opta por la denominación «corpus virtual» para referirse al concepto de corpus *ad hoc*. Como hemos mencionado anteriormente, existen diversos términos para hacer referencia al mismo concepto.

De hecho, los corpus pueden tener múltiples aplicaciones en el ámbito de la traducción y la interpretación, tal y como señala Corpas Pastor (2001: 159-164), a saber, la traducción automática, las bases de datos para la traducción asistida por ordenador, la comparación interlingüística y caracterización de la lengua meta y las aplicaciones de los corpus en la docencia.

Por lo que respecta a la traducción automática, Forcada (2010: 218-220) distingue dos tipos de traducción automática: *rule-based* y *corpus-based*, categoría esta última que, a su vez, se divide en *example-based* y *statistical machine translation* y, actualmente, *neural machine translation*. Su funcionamiento consiste en analizar muestras reales de la lengua y de sus traducciones para proporcionar el resultado más aproximado posible. Por ello, una de las utilidades principales de los corpus en traducción es la evaluación de la calidad del *output* de los sistemas de traducción automática ya que, como hemos mencionado, la compilación de corpus paralelos es fundamental en el desarrollo de estos sistemas y de bases de datos textuales como son las memorias de traducción. Estas últimas son componentes de la traducción asistida por ordenador, por lo que proporcionan asistencia al traductor mediante el rastreo en corpus paralelos alineados de segmentos similares traducidos previamente y, en caso de que no se encuentren coincidencias, el motor de traducción automática le proporciona una posible traducción basándose en búsquedas en corpus comparables y paralelos (Corpas Pastor, 2012: 5). De forma similar trabajan la mayoría de las herramientas que forman parte de las tecnologías de la traducción, como los programas de concordancias o los extractores y gestores de terminología, normalmente basados en corpus.

Por otro lado, los corpus son importantes como método de análisis e investigación, ya que permiten observar las diferencias existentes entre la producción de los textos originales y de sus traducciones, como la ambigüedad de la lengua o la distribución de los rasgos lingüísticos (Baker, 1993). Se cree que esto podría ayudar a desvelar ciertas tendencias en los traductores que podrían ser consideradas universales de traducción (Corpas Pastor, 2012: 12). Baker (1993, 1996) observa el comportamiento traductor y enumera algunos de estos universales investigados mediante el uso de corpus. Explica que, en general, los textos traducidos suelen ser más simples y menos ambiguos y que, además, estos presentan una distribución diferente de algunos de sus rasgos lingüísticos. Señala también que estos se muestran mucho más uniformes y homogéneos entre sí que los textos originales.

Hoy en día, los corpus siguen siendo de gran utilidad como materiales de comparación interlingüística, sobre todo en el caso de los corpus comparables bilingües y multilingües. El hecho de que este tipo de corpus esté constituido por textos originales en distintas lenguas permite el cotejo del comportamiento de las mismas a partir de las muestras recogidas y facilita el establecimiento de equivalentes de traducción para unos términos concretos a través de la observación de los patrones lingüísticos que se presentan en determinadas situaciones.

Además, los corpus también han llegado a las aulas de Traducción e Interpretación tras su aplicación a la enseñanza de lenguas, y se ha demostrado que su uso permite a los alumnos conseguir una mejor comprensión del texto origen a la vez que facilita el proceso de producción del texto meta. Numerosos autores (Bowker, 1999b; Linquist, 1999; o Zanettin, 1998; 2002) han defendido el uso de corpus (tanto comparables como paralelos) como materiales complementarios en la enseñanza. Esta aplicación,

innovadora en su tiempo, ofrece diversas ventajas, como el hecho de que potencia el uso de tecnologías en las aulas de Traducción e Interpretación y que ayuda a desarrollar la competencia traductora y la actividad docente.

Como podemos observar, son numerosas las utilidades de los corpus como herramientas de traducción. Además, muchas veces los corpus son esenciales para llevar a cabo el proceso de manera completa y constituyen un factor determinante en la calidad del producto final. Por ello, hoy en día, cada vez son más los profesionales (no solo de la traducción, sino de diversos campos) que recurren a estas herramientas como método de apoyo a la hora de realizar su trabajo y, de alguna manera, complementarlo. Asimismo, en el campo que nos atañe, los corpus son de gran ayuda porque optimizan la productividad, lo que explica que se hayan convertido en una herramienta indispensable para el traductor.

2.2.4. CARACTERÍSTICAS DE LOS CORPUS AD HOC COMO HERRAMIENTAS DE TRADUCCIÓN

A pesar de que los corpus *ad hoc* son una herramienta de gran utilidad para los traductores, su empleo no está exento de dificultad, como señala Seghiri (2017: 47): «Los corpus que encontramos en la red no son altamente especializados en su mayoría o, en el caso de que encontrásemos corpus especializados, es probable que no cubran todas las necesidades del género textual y temática que deseamos documentar».

Esto, sumado a la poca fiabilidad de algunas de las fuentes a las que podemos acceder a través de Internet, obliga al traductor en algunos casos a compilar su propio corpus, para lo que deberá seguir una metodología y unos criterios precisos con el fin de asegurar el equilibrio y la representatividad del mismo. Por esta razón, destacamos las principales características que se deben tener en cuenta al compilar un corpus.

McEnery y Hardie (2012) definen el equilibrio como «*a property of a corpus (or, more properly, of a corpus sampling frame)*». Explican que «*A corpus is said to be balanced if the relative sizes of each of its subsections have been chosen with the aim of adequately representing the range of language that exists in the population of texts being sampled* (McEnery y Hardie, 2012: 239)».

Además, estos autores consideran que un corpus es representativo si «*[it] is sampled in such a way that it contains all the types of text, in the correct proportions, that are needed to make the contents of the corpus an accurate reflection of the whole of the language or variety that it samples. See also "balance"*» (McEnery y Hardie, 2012: 250).

Una vez definidos esos parámetros, también es necesario determinar si el corpus será paralelo (textos originales y sus traducciones) o comparable (textos originales en ambas lenguas). Dependiendo de cuál sea nuestro propósito, compilaremos un corpus de un tipo o de otro, ya que ambos tienen distintas utilidades. Por ejemplo, en el caso de un corpus paralelo nos resultará muy sencillo encontrar en el texto meta cualquier noción expresada en el texto original (Rabadán y Fernández Nistal, 2002: 51-52). De otra forma, en un corpus comparable conoceremos exactamente los términos que se emplean en ambas lenguas para designar un concepto con la certeza de que ese es el término correcto en esa lengua, dado que son textos originales.

Otra característica que se debe tener en cuenta al compilar un corpus es su tamaño, que puede ser variable dependiendo del objetivo del trabajo.

En su diseño no prima tanto la cantidad como la calidad: por regla general, el corpus *ad hoc* no incluye un número demasiado elevado de textos, pero sí textos muy adecuados, equiparables al texto original en cuanto a la temática, el género y la variedad textual (Corpas Pastor, 2004: 236).

Asimismo, es importante que exista también una univocidad en los criterios de selección de las fuentes y de los textos incluidos. Por ejemplo, que estos se hayan publicado de forma más o menos contemporánea, lo que otorgará homogeneidad al corpus a pesar de la diversidad del mismo.

Teniendo en cuenta los criterios y características que debe tener un corpus *ad hoc*, hemos revisado y alimentado un corpus ya existente, que hemos tomado como punto de partida para llevar a cabo el estudio terminológico y fraseológico en el campo de los embutidos, cuya metodología desarrollamos en detalle en el próximo capítulo.

3. METODOLOGÍA

3.1. INTRODUCCIÓN

Como hemos explicado anteriormente, nuestro estudio se ha basado en la compilación de un corpus para, posteriormente, analizar el comportamiento de diversos elementos terminológicos y fraseológicos dentro del mismo.

A partir del corpus paralelo P-GEFEM, compuesto por fichas descriptivas de embutidos en español y sus traducciones al inglés (Ortego-Antón y Fernández Nistal, en prensa), hemos revisado los textos que componen dicho corpus, dado que algunos de ellos no cumplían con los requisitos necesarios para llevar a cabo nuestro análisis al no cumplir exactamente la misma función textual. En nuestro caso, acotamos la búsqueda de textos solamente a un determinado género textual, las fichas descriptivas de embutidos, y nos centramos en la terminología del chorizo. Por este motivo, fue necesario comprobar las características de todos los textos que componían el corpus P-GEFEM: si los textos originales y sus traducciones presentaban la misma estructura, si no había omisiones o adiciones en la traducción, si eran fichas descriptivas de producto, etc. Una vez revisado, procedimos a eliminar todos aquellos que no cumplían con las características requeridas para nuestro análisis y, a continuación, a alimentarlo con textos que cumplieran los requisitos de búsqueda. Para llevar a cabo dicho proceso, nos basamos en el protocolo de compilación expuesto por Seghiri (2017: 47), constituido por cuatro fases: búsqueda, descarga, formato y almacenamiento.

3.2. REVISIÓN Y ALIMENTACIÓN DE P-GEFEM

La primera fase, la búsqueda, consistirá en buscar en la red textos en páginas web que cumplan con los requisitos de nuestro estudio. Para ello, las páginas deberán tener disponible una versión original en español y otra traducida al inglés y mostrar fichas de producto (teniendo en cuenta que el único artículo que nos interesa es el chorizo). Llevaremos a cabo un proceso de búsqueda en páginas de diversas empresas nacionales basadas en la producción de chacinas, como Ortiz⁵, Incarlopsa⁶, Ezequiel⁷, Tello⁸, Espuña⁹, Monter¹⁰ o Emcesa¹¹, entre otras. A partir de las páginas web de estas empresas, descargaremos las fichas de producto como texto en formato plano (.txt) y sustituiremos aquellas que no son válidas en el corpus original por los nuevos textos. Para almacenarlos, crearemos una carpeta compartida en Google Drive denominada P-GEFEM a la que añadiremos una carpeta denominada GESTIÓN y otra denominada CORPUS, que se divide en dos subcarpetas (ES y EN) a las que subiremos los textos en formato .txt en ambas lenguas con una denominación preestablecida (esta constituirá la fase de almacenamiento), tal y como explican Ortego-Antón y Fernández Nistal (2019):

[Los archivos] se guardan correlativamente con el número (101, 102, 103, etc.), el género textual (DM = dried meats en lengua inglesa), el origen (WS = web), dos iniciales con la denominación de la empresa (por ejemplo, IM = Ibéricos Montellano [o ES = Espuña]), el tipo textual [en este caso, fichas de producto] (PD = product description), [...] la fecha (AAMMDD), el campo (Foodie) y la lengua (ES o EN).

⁵ <http://www.embutidosortiz.com/> (Fecha de consulta: 11/02/2019)

⁶ <https://www.incarlopsa.es/> (Fecha de consulta: 11/02/2019)

⁷ <https://www.embutidosezequiel.com/> (Fecha de consulta: 15/02/2019)

⁸ <http://www.tello.es/> (Fecha de consulta: 19/02/2019)

⁹ <https://www.espuna.es/> (Fecha de consulta: 19/02/2019)

¹⁰ <http://www.embutidosmonter.es/es/> (Fecha de consulta: 19/02/2019)

¹¹ <https://emcesa.com/> (Fecha de consulta: 22/02/2019)

Un posible ejemplo es el archivo 125DMwsESPD190121FoodieES, que hace referencia a un texto de la empresa España, descargado el 21 de enero de 2019.

Además, en la carpeta de GESTIÓN dentro de P-GEFEM también crearemos una hoja de cálculo compuesta por varias columnas con distinta información: la denominación de los archivos .txt, la URL de la página web en la que se encuentra la ficha de producto, el ID del investigador (en nuestro caso, nuestras iniciales: ALT), el nombre de la empresa, la fecha de descarga, el tipo de producto (chorizo), el país y la denominación que la empresa da al producto en cuestión (por ejemplo, chorizo ibérico de bellota).

	A	B	C	D	E	F	G	H	I
16	ID	URL	Researcher ID	Author	Company	Download date	Product	Country	Company denomination
17	115DMwsTEPD190128FoodieES	http://www.tello...	ALT		Tello	28/01/2019	chorizo	Spain	Chorizo Castellano Dulce
18	116DMwsTEPD190128FoodieES	http://www.tello...	ALT		Tello	28/01/2019	chorizo	Spain	Chorizo Extra Tradicional
19	117DMwsTEPD190128FoodieES	http://www.tello...	ALT		Tello	28/01/2019	chorizo	Spain	Chorizo Sarta Extra Picante
20	118DMwsTEPD190128FoodieES	http://www.tello...	ALT		Tello	28/01/2019	chorizo	Spain	Chorizo Sarta Extra Dulce
21	119DMwsTEPD190128FoodieES	http://www.tello...	ALT		Tello	28/01/2019	chorizo	Spain	Chorizo Pamplona Extra
22	120DMwsLQPD190206FoodieES	https://www.los...	ALT		Los Quijales	06/02/2019	chorizo	Spain	Chorizo Blanco
23	121DMwsLQPD190206FoodieES	https://www.los...	ALT		Los Quijales	06/02/2019	chorizo	Spain	Chorizo Etiqueta Negra
24	122DMwsLQPD190206FoodieES	https://www.los...	ALT		Los Quijales	06/02/2019	chorizo	Spain	Chorizo Sarta Dulce
25	123DMwsTNPD190207FoodieES	https://www.torre...	ALT		Torre de Núñez	07/02/2019	chorizo	Spain	Chorizo Primera 500 gr.
26	124DMwsTNPD190207FoodieES	https://www.torre...	ALT		Torre de Núñez	07/02/2019	chorizo	Spain	Chorizo Sarta Extra Picante
27	125DMwsESPD190121FoodieES	https://www.espu...	ALT		España	21/01/2019	chorizo	Spain	Chorizo
28	126DMwsESPD190121FoodieES	https://www.espu...	ALT		España	21/01/2019	chorizo	Spain	Chorizo Picante

Figura 1. Hoja de cálculo con información sobre cada archivo .txt y los productos que se muestran en las páginas en español

	A	B	C	D	E	F	G	H	I
16	ID	URL	Researcher ID	Author	Company	Download date	Product	Country	Company denomination
17	115DMwsTEPD190128FoodieTEN	http://www.tello...	ALT		Tello	28/01/2019	chorizo	Spain	Non Spicy String Castellano Chorizo
18	116DMwsTEPD190128FoodieTEN	http://www.tello...	ALT		Tello	28/01/2019	chorizo	Spain	Traditional Chorizo Extra
19	117DMwsTEPD190128FoodieTEN	http://www.tello...	ALT		Tello	28/01/2019	chorizo	Spain	Spicy Sarta Chorizo
20	118DMwsTEPD190128FoodieTEN	http://www.tello...	ALT		Tello	28/01/2019	chorizo	Spain	Non Spicy Sarta Chorizo
21	119DMwsTEPD190128FoodieTEN	http://www.tello...	ALT		Tello	28/01/2019	chorizo	Spain	Pamplona Chorizo Extra
22	120DMwsLHPD190206FoodieTEN	https://www.losq...	ALT		Los Quijales	06/02/2019	chorizo	Spain	Extra White Chorizo
23	121DMwsLQPD190206FoodieTEN	https://www.losq...	ALT		Los Quijales	06/02/2019	chorizo	Spain	Black Label Chorizo
24	122DMwsLQPD190206FoodieTEN	https://www.losq...	ALT		Los Quijales	06/02/2019	chorizo	Spain	Sarta Chorizo
25	123DMwsTNPD190123FoodieTEN	https://www.torre...	ALT		Torre de Núñez	07/02/2019	chorizo	Spain	Galician Chorizo Prime 500 gr.
26	124DMwsTNPD190207FoodieTEN	https://www.sabo...	ALT		Torre de Núñez	07/02/2019	chorizo	Spain	Chorizo Sarta Extra Spicy
27	125DMwsESPD190121FoodieTEN	https://www.espu...	ALT		España	21/01/2019	chorizo	Spain	Chorizo
28	126DMwsESPD190121FoodieTEN	https://www.espu...	ALT		España	21/01/2019	chorizo	Spain	Hot Chorizo

Figura 2. Hoja de cálculo con información sobre cada archivo .txt y los productos que se muestran en las páginas en inglés

En consecuencia, P-GEFEM se corresponde con un corpus *ad hoc*, dado que se ha creado para suplir unas necesidades específicas; *virtual* (o *ad hoc*), ya que está

compuesto por textos obtenidos de Internet; *paralelo*, al integrar textos originales en español y sus traducciones en inglés; *bilingüe* y *monodireccional*, dado que las traducciones solo son en una dirección (de la lengua A, español, a la lengua B, inglés); y *textual*, al estar compuesto por textos que, en nuestro caso, son fichas descriptivas de producto.

Además, el corpus P-GEFEM cumple con las características de la representatividad (puesto que está compuesto por textos del mismo género textual) y del equilibrio, ya que el corpus está compuesto por 100 textos en español y 100 en inglés obtenidos de más de 20 empresas distintas. Por lo que respecta al tamaño, en la Tabla 1 se muestra el número de palabras en total (casos) y el número de palabras distintas (tipos) de cada corpus (*tokens* y *types* en inglés respectivamente).

	Casos	Tipos
ES	10 387	1017
EN	29 119	1895

Tabla 1. Tamaño de P-GEFEM

3.3. LA MUESTRA DE ANÁLISIS

En primer lugar, debemos seleccionar la muestra de análisis, para lo que utilizaremos la herramienta AntConc 3.4.4. (Anthony, 2018), un analizador de concordancias de corpus y de análisis textual. Introduciremos todos los textos del corpus en cada una de las lenguas codificados como UTF-8 para, tras filtrar los resultados con una *stoplist* con el fin de evitar los determinantes, pronombres, artículos y conjunciones, observar los términos que aparecen con más frecuencia en los textos en español (Figura 3). Esto lo haremos mediante la lista de palabras (*Word List*) que proporciona el programa con el fin de seleccionar aquellos que constituirán la muestra de análisis (20 en total): «proteína», «grasa», «hidrato de carbono», «kilocaloría», «producto», «información», «extra», «saturado», «azúcar», «azúcares», «peso», «peso neto», «gramo», «kilogramo», «ingrediente», «leche», «loncha», «fresco», «valor medio» y «conservar».

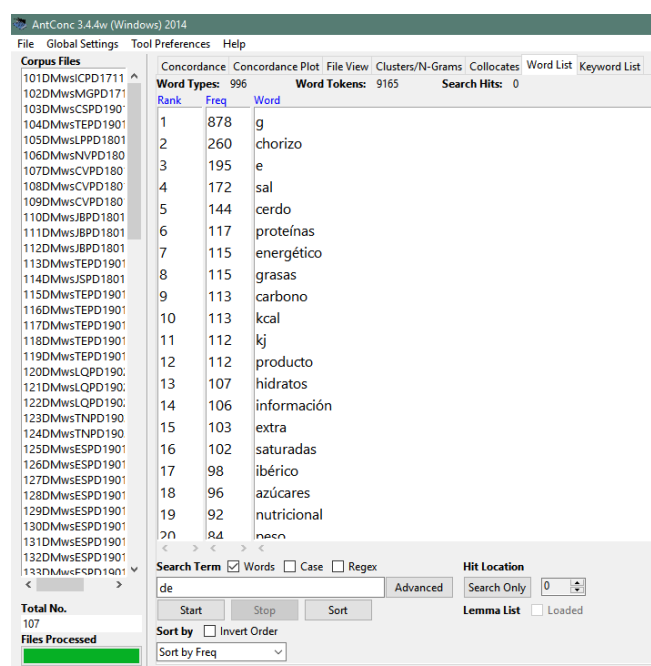


Figura 3. Listado de unidades léxicas más frecuentes en el subcorpus en español de P-GEFEM extraídas con la función Word List de AntConc 3.4.4. (Anthony, 2018)

3.4. LOS EQUIVALENTES Y SU FRASEOLOGÍA

Una vez establecida la muestra de análisis y con el fin de analizar las concordancias de P-GEFEM en profundidad, procederemos a utilizar ParaConc, una herramienta de gran interés en nuestro caso, ya que es un analizador de concordancias multilingüe de corpus paralelos. Esta aplicación alinea los textos del corpus (que, al igual que en AntConc, deben estar guardados como .txt) en una lengua y en otra, extrae las concordancias de los términos que se introduzcan en el buscador y las muestra en su contexto. Además, es multilingüe porque puede gestionar corpus compuestos de hasta cuatro subcorpus. Esto significa que se podría analizar de forma simultánea un corpus compuesto por textos originales en español y por sus traducciones hasta en cuatro lenguas. Una de las ventajas de este programa es que su interfaz ofrece al usuario una vista que permite comparar los textos en cada lengua en un vistazo, tal y como muestra la siguiente imagen.

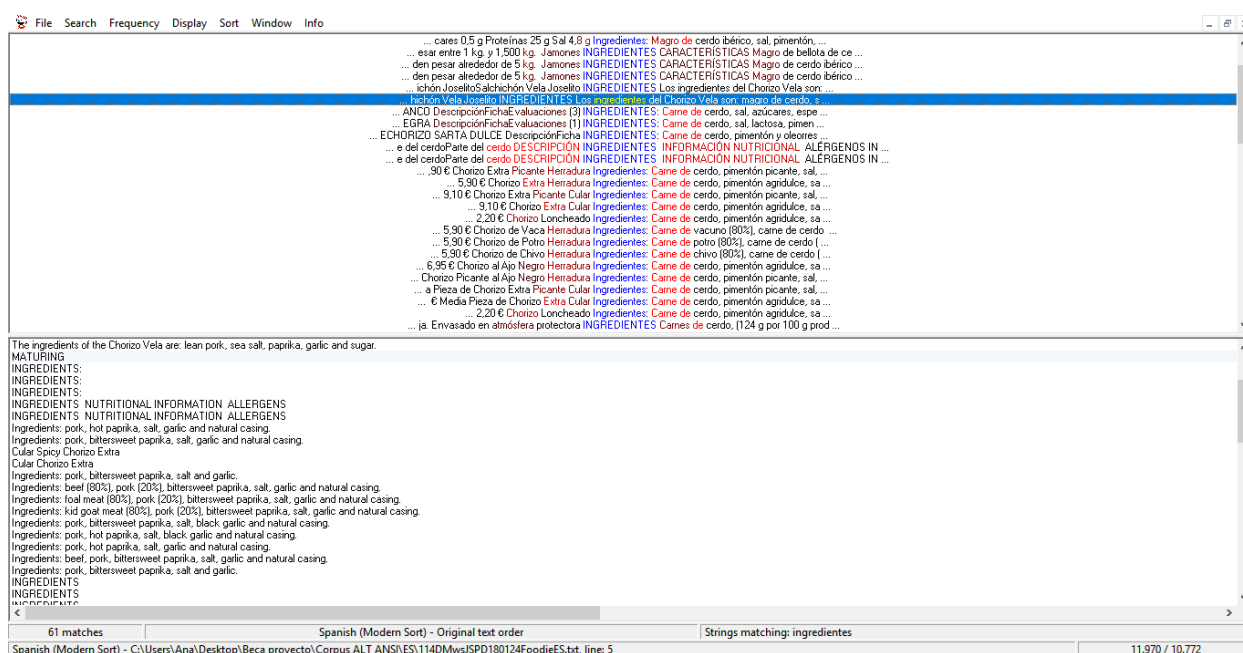


Figura 4. Ejemplo de concordancia de una unidad léxica y su traducción en español y en inglés con ParaConc

Además, ParaConc ofrece la frecuencia de aparición de las unidades léxicas (en la imagen, 61 matches, en la parte inferior izquierda de la ventana) y de sus colocaciones, que aparecen destacadas en ambas partes de la pantalla. En este trabajo, vamos a definir «colocación» como «coaparición de palabras»; por ejemplo, algunas de las colocaciones de «proteína» se corresponden con «proteína de soja», «proteína láctea» o «proteína grasa». Asimismo, ParaConc permite buscar posibles traducciones en el corpus traducido para el término que se esté estudiando y, al igual que AntConc, ofrece la opción de extraer un listado con las palabras más frecuentes dentro del corpus en cada una de las lenguas a través del comando *Frequency>Frequency Order*.

Con la información proporcionada por esta herramienta, diseñaremos una hoja de cálculo que reúna información sobre los términos de la muestra de análisis, las ocurrencias¹² de los mismos en español, la traducción al inglés en P-GEFEM, las ocurrencias del término traducido, las colocaciones de cada término en español y la

¹² Número de veces que aparece cada término en el corpus.

traducción de dichas colocaciones al inglés. Además, en otra columna, añadiremos posibles observaciones relacionadas con el uso de cada término y de su traducción.

	A	B	C	D	E	F	G
1	Term ES	Ocurrencias term ES	Traducción EN	Ocurrencias term EN	Colocaciones	Traducción colocaciones	Observaciones
2	proteína	110 proteína solo 137 proteína en general (solo y colocaciones)	protein	84	proteína de soja: 6	soy protein	
3					proteína de (la) leche: 6	milk protein	
4					proteína láctea: 3		
5					fuelle de proteínas: 6	source of protein	
6					proteína grasa: 1	protein fat	
7	grasa	19 con el significado de tocino 103 con el significado de	fat	70	grasa saturada: 7	saturated fat	Nota de uso: "grasa" en singular se emplea para hablar del tocino del cerdo. En plural, se usa para hablar de los lípidos (se traduce como "fat" en ambos casos).
8							
9							
10							
11							
12							
13							
14							
15							
16							
17							
18							
19							
20							

Figura 5. Fragmento de la hoja de cálculo con información terminológica y fraseológica extraída de P-GEFEM

Por último, compararemos las traducciones con los términos y la fraseología empleados originalmente en inglés a partir del estudio en un segundo corpus: C-GEFEM.

3.5. EXPLOTACIÓN DE C-GEFEM Y COMPARACIÓN LINGÜÍSTICA

C-GEFEM es un corpus *ad hoc*, virtual y comparable, es decir, compuesto por textos originales en inglés, cuyo diseño y compilación se llevó a cabo en el marco del proyecto I+D nacional titulado *Análisis contrastivo y traducción inglés-español: Aplicaciones III* (Ref. FFI2013-42994-R¹³). Por lo tanto, nos basaremos en un corpus comparable ya existente de la misma naturaleza que P-GEFEM para comprobar si los patrones de comportamiento lingüísticos observados en el corpus paralelo se reflejan en el corpus comparable y, por ende, si en las traducciones al inglés se emplean los equivalentes establecidos en inglés.

En esta fase, volveremos a utilizar AntConc 3.4.4. (Anthony, 2018). Siguiendo la metodología descrita por Ortego-Antón y Fernández Nistal (en prensa), comprobaremos si los términos de las traducciones de P-GEFEM se emplean en los textos originales, introduciremos cada término en inglés en el buscador de palabras del programa y observaremos sus ocurrencias y sus colocaciones. Si AntConc no ofrece resultados para el término que se busca, dicho término no se utiliza en los textos redactados originalmente en lengua inglesa. En el caso de que sí aparezcan ocurrencias, observaremos las colocaciones del término de manera eficaz gracias a la ilustrativa interfaz del programa. Con la información recopilada, añadiremos cuatro columnas a la hoja de cálculo que habíamos creado en la fase anterior con la siguiente información: término empleado en inglés, ocurrencias, colocaciones del término en inglés junto con el número de ocurrencias y posibles observaciones (Figura 7). Esto nos permitirá contrastar de manera sencilla el comportamiento de la terminología traducida al inglés

¹³ http://contraste2.unileon.es/web/es/corpus0_C-GEFEM.html (Fecha de consulta: 30/04/2019).

y comprobar si se asemeja a la recogida en los textos originalmente redactados en esta lengua.

	A	B	C	D	E	F	G	H	I	J	K
	Term ES	Ocurrencias term ES	Traducción EN	Ocurrencias term EN	Colocaciones	Traducción colocaciones	Observaciones	Término empleado en inglés	Ocurrencias	Colocaciones term EN	Observaciones
1	proteína	110 proteína solo 137 proteína en general (solo y colocaciones)	protein	84	proteína de soja: 6	soy protein		protein	285	soya protein: 10/soy protein: 4	"protein fat" no aparece Nota de uso: "milk protein" aparece en plural 11 veces del total de 15
2					proteína de (la) leche: 6	milk protein				high in protein: 9	
3					proteína láctea: 3					pork protein: 4	
4					fuelle de proteínas: 6	source of protein				meat protein: 3	
5					proteína grasa: 1	protein fat				fibre nil protein: 13	
6										protein kick: 3	
7										rich in protein: 6	
8											
9											
10											
11	grasa	19 con el significado de tocino 103 con el significado de lípido	fat	70	grasa saturada: 7	saturated fat	Nota de uso: "grasa" en singular se emplea para hablar del tocino del cerdo. En plural, se usa para hablar de los lípidos (se traduce como "fat" en ambos casos).	fat	421	saturated fat: 27/sat fat: 6	"fatty acid" no aparece
12										white fat: 13	
13										total fat: 31	
14										high fat: 5	
15										trans fat: 20	
16										reduced fat: 5	
17										fat content: 1	
18										fat level: 1	
19										lump of fat: 1	
20										duck fat: 4	
21										pork fat: 9	
				fat (curiosamente, "grasa" a veces se traduce como "fatty"						baaf fat: 2	

Figura 6. Fragmento de la hoja de cálculo utilizada para la recogida de datos para el contraste de resultados

Por último, crearemos una base de datos terminológica con la herramienta SDL MultiTerm 2019 a partir de un archivo de definición (archivo que determina el número y el tipo de los campos que están presentes en cada entrada de la base de datos terminológica) creado previamente (Ortego-Antón, 2019: 96). La utilización de este programa para completar esta última fase está establecida previamente al formar parte el estudio de un proyecto de mayor envergadura. Esta fase nos servirá para desarrollar una aplicación práctica con los resultados que se desprenden del análisis lingüístico basado en corpus, ya que crearemos una entrada para cada término. En cada entrada, en el primer nivel, se añadirá información de gestión relevante, como el autor (en nuestro caso, ALT), la fecha de creación y de revisión de la entrada (con formato *dd/mm/aa*) o el campo al que pertenecen los términos recogidos dentro de su sistema conceptual. En el segundo nivel (idioma), cada ficha incluirá el término en español y su equivalente en inglés y, en el tercer nivel (término), se añadirá distinta información para cada lengua: la categoría gramatical del término, posibles abreviaturas, un ejemplo en contexto, posibles variantes y colocaciones más frecuentes. En el caso de los términos con más de una acepción, se creará una entrada para cada significado.

Una vez expuesta la metodología, procedemos en el siguiente capítulo a presentar el análisis de los datos obtenidos.

4. ANÁLISIS Y RESULTADOS

En el presente capítulo comprobaremos cómo se traducen los términos a través del análisis en P-GEFEM y, a continuación, contrastaremos los resultados obtenidos en el corpus comparable C-GEFEM.

4.1. RESULTADOS EN P-GEFEM

En este apartado, nos centraremos solo en aquellos casos en los que hemos observado que no se trasvasa utilizando el equivalente establecido. Dedicaremos un epígrafe a cada término, en cada uno de los que comentaremos la información observada en relación a cada uno de ellos.

4.1.1. Grasa

En primer lugar, observamos que un término puede hacer referencia a dos conceptos diferentes dentro de este campo del saber. Por ejemplo, el término «grasa» se repite regularmente a lo largo de los textos; no obstante, tiene dos acepciones distintas: o bien hace referencia al tocino del cerdo, es decir, al sebo del animal, o bien se refiere a los lípidos del chorizo una vez preparado para el consumo (valor que aparece en la tabla de información nutricional). Por esta razón y como hemos explicado anteriormente, cuando alimentemos la base de datos terminológica, tendremos que crear dos entradas distintas («grasa1» y «grasa2») para recoger sus dos significados. En cuanto a la traducción de dicho término, existe solo un equivalente en el subcorpus en inglés de P-GEFEM: *fat*; es decir, en lengua inglesa, el término *fat* puede referirse tanto a la cantidad de sebo presente en una loncha de chorizo como a la proporción de grasas del producto, igual que sucede en español con su equivalente «grasa».

Además, «grasa» presenta dos traducciones distintas en inglés dentro del corpus paralelo P-GEFEM: *fat* y *fatty acid*. Sin embargo, el equivalente de este último término se corresponde con «ácido graso», que no es exactamente lo mismo que «grasa» (por explicarlo de una manera simplificada, las grasas están compuestas por ácidos grasos). No obstante, las veces que aparece el término «ácido graso» en los textos en español, su traducción al inglés es igualmente sorprendente, ya que se traduce como *fat* (grasa) y, por lo tanto, no se corresponde con el equivalente establecido: *fatty acid*. Podría decirse que nos encontramos ante una traducción cruzada. En estos casos, observamos en las traducciones una carencia de equivalencia y exactitud, ambos factores muy importantes en traducción, ya que se debe transmitir exactamente el mismo significado que en el original para garantizar la calidad de la traducción.

4.1.2. Azúcar

Un fenómeno similar al detectado con «grasa», en el que se pone de manifiesto el empleo de distintas acepciones para un mismo término, también se reproduce con el término «azúcar» que, en singular, hace referencia al ingrediente y, en plural, a la proporción de hidratos de carbono presentes en el producto (en este caso, el significado varía en función de una cuestión morfológica como es el número en el que se emplea el término). Por esta razón, también crearemos dos entradas distintas en la base de datos terminológica para este término («azúcares» y «azúcar»).

4.1.3. Producto

El término «producto» es uno de los más frecuentes en el corpus (100 ocurrencias). Observamos que, en inglés, aparece la colocación *dairy product*, mientras que en español, la colocación «producto lácteo» (que se correspondería con la traducción literal) no está presente, sino que la denominación *dairy products* se refiere a la colocación «leche y derivados». Así pues, volvemos a encontrarnos ante una falta de exactitud en la traducción de esta colocación, que aparece en los textos en ambas lenguas al presentar los alérgenos del producto.

4.1.4. Información

Observamos que el término «información» se traduce como *information* en la mayoría de los casos, pero también como *fact* en ciertas ocasiones, y que se presenta también una abreviatura, *info*. En el siguiente apartado comprobaremos si estos términos se utilizan en los textos redactados originalmente en inglés y si, por ende, las traducciones proporcionadas durante el trasvase interlingüístico del español al inglés son correctas.

4.1.5. Extra

Este término presenta tres variantes dentro del corpus paralelo: *extra*, *premium* y *premium quality*, que funcionan como adjetivos en los tres casos (Ej.: *premium quality cured Galician chorizo*¹⁴).

4.1.6. Fresco y conservar

Otros términos que presentan distintas traducciones son «fresco» y «conservar». Para el primero, observamos tanto *fresh* como *cool*. En cuanto al segundo, este se traduce utilizando las unidades léxicas *keep* y *store*.

4.1.7. Saturado

Al margen de la existencia de varias denominaciones para trasvasar un término del español al inglés, también hemos constatado variaciones respecto al uso de ciertos términos. Por ejemplo, cabe mencionar que el término «saturado» aparece siempre en femenino plural, ya que siempre acompaña a «grasas». En muchos casos, la colocación que se emplea en español es «de las cuales saturadas», que se traduce al inglés por *of which saturated*.

4.1.8. Ingrediente y loncha

El término «ingrediente» también aparece siempre en plural, ya que encabeza la lista que enumera los ingredientes del producto (Ej.: Ingredientes: Magro de cerdo ibérico, sal, pimentón, azúcar, especias, antioxidante, (E-301), aroma de humo, emulgentes (E-450i, E451i y E-452i), colorante (E-160c), conservador (E-250), tripa natural de cerdo. Puede contener trazas de leche y soja¹⁵). Asimismo, el término «loncha» aparece también la mayoría de las veces en plural.

Del mismo modo, hemos detectado errores ortotipográficos en las páginas web en español, como la omisión de la tilde en algunas unidades léxicas («*dioxido» en «dióxido de carbono»), o la adición de punto (*kg.), de inicial mayúscula (*Kg) y de-*s* de plural

¹⁴ 158DMwsTNPd190211FoodieTEN.txt

¹⁵ 107DmwsCVPD180123FoodieES.txt

(*kgs) a los símbolos «kg» o «kcal» (Fundéu, 2012)¹⁶. En general, las empresas le restan importancia a ciertos elementos sus páginas web, descuidando en muchos casos la ortografía o la redacción de los textos publicados en su propia lengua.

4.2. CONTRASTE DE RESULTADOS EN C-GEFEM

Esta fase nos ha permitido comprobar si las traducciones que se ofrecen para los términos seleccionados en el corpus paralelo realmente se emplean en lengua inglesa a partir del contraste de los resultados del corpus comparable, es decir, en los textos redactados originariamente en inglés.

En general, podemos afirmar que observamos que el número de colocaciones es notablemente superior en los textos redactados originalmente en lengua inglesa que en los redactados en lengua española. A continuación, procederemos del mismo modo que en el apartado anterior: desarrollando brevemente en distintos epígrafes los casos que nos han llamado más la atención.

4.2.1. Proteína

Por lo que respecta al término «proteína», constatamos que las colocaciones que aparecen en los textos en español y, por tanto, en las traducciones a inglés, coinciden en general con las de los textos originales en inglés. No obstante, podemos ver algunas diferencias: por ejemplo, para la colocación «proteína de soja», en las páginas traducidas de español a inglés, se proporciona solamente la traducción *soy protein* mientras que, en los textos redactados originalmente en inglés se emplea tanto *soy protein* como *soya protein*. Al comprobar el uso de ambos términos en los textos redactados originalmente en inglés, observamos que la colocación más empleada es *soya protein* (10 ocurrencias frente a 4 para *soy protein*), por lo que podemos afirmar que el equivalente que se utiliza en las traducciones no es el establecido en inglés. Las traducciones del resto de las colocaciones recogidas en los textos traducidos de español a inglés para «proteína» se corresponden con las colocaciones originales en inglés, por lo que su traducción es correcta. La única observación destacable es el hecho de que *milk protein* aparece la mayoría de las veces en plural y no en singular como se detecta en las traducciones), tal y como se muestra en la siguiente tabla.

<p>Traducción español-inglés</p>	<p><i>INGREDIENTS: Pork meat, pork belly, salt, paprika, milk protein, dextrin, dextrose, lactose, garlic, spices, stabilizers (E-450, E-451), antioxidants (E-301, E-331), flavour enhancer (E-621), preservatives (E-250, E-252) and coloring (E-120)</i>¹⁷.</p>
<p>Texto redactado originalmente en inglés</p>	<p><i>Ingredients: Pork, Salt, Lactose (Cows' Milk), Paprika (1.5 %), Dextrose, Sugar, Milk Proteins (Cows' Milk), Garlic, Paprika Extract, Acidity Regulator: Sodium citrate; Antioxidant: Sodium Erythorbate; Preservatives: Sodium Nitrite, Potassium Nitrate; Rosemary Extract, Black Pepper, Oregano</i>¹⁸.</p>

Tabla 2. Comparación de la traducción español-inglés del término *milk protein* y su uso en los textos originales en inglés

¹⁶ La norma dice que los símbolos «no van seguidos por un punto abreviativo» y que «no tienen plural» (Fundéu).

¹⁷ 172DMwsMTPD190121FoodieTEN.txt

¹⁸ 006DMwsSBS160624FoodieEN.txt

4.2.2. Grasa

La colocación más frecuente en español es «grasa saturada», que se traduce al inglés como *saturated fat*. Esta colocación también aparece en los textos en inglés, por lo que se trasvasa utilizando el equivalente establecido. También detectamos que el término *fatty acid* no se utiliza en los textos originales en inglés, por lo que la traducción de «grasa» por *fatty acid*, que hemos registrado en el subcorpus en inglés de P-GEFEM, estaría injustificada y sería un error de traducción.

4.2.3. Hidrato de carbono

El término «hidrato de carbono» se trasvasa al inglés utilizando el equivalente establecido (*carbohydrate*) y esta es la unidad léxica que se emplea en los textos redactados originalmente en lengua inglesa. No obstante, nos gustaría detenernos y explicar un detalle relacionado con su uso: en español aparece en plural («hidratos de carbono») en las 95 ocurrencias, pero se traduce al inglés en ocasiones en singular (*carbohydrate*), y esta es una decisión sin explicación aparente. Además, el sinónimo «carbohidrato», que es más similar a la forma inglesa, aparece en cinco ocasiones en español y se traduce también correctamente como *carbohydrate*.

4.2.4. Información

Con respecto al término «información», comentábamos anteriormente que se empleaban dos variantes en las traducciones: *information* y *fact*. Al acudir al corpus comparable para verificar si su uso era correcto comprobamos, en efecto, que se emplean ambas formas en los textos originales en inglés, por lo que podemos afirmar que el trasvase interlingüístico se ha realizado con éxito. El término *fact* aparece en la colocación *nutrition facts*, referente a la información nutricional de los productos. Por otra parte, la abreviatura *info.* se utiliza también en los textos originales. En consecuencia, las tres traducciones detectadas en P-GEFEM para la colocación «información nutricional» son correctas, ya que hemos comprobado que las tres se utilizan en C-GEFEM: *nutritional information*, *nutrition facts* y *nutritional info*.

La otra colocación relativa al término «información» se corresponde con «información sobre alérgenos», que se traduce en el corpus paralelo como *allergen information*. En el corpus comparable aparece reflejada dicha colocación, por lo que podemos afirmar que se emplea en inglés. No obstante, comprobamos que existe otra forma más recurrente que únicamente se emplea en los textos redactados en lengua inglesa: *allergy information*.

4.2.5. Extra

Sobre el término «extra» (referente a la calidad del chorizo) hemos descrito previamente las tres alternativas que se recogían en el corpus paralelo: *extra*, *premium* y *premium quality*. Estas tres formas aparecen reflejadas en el corpus comparable (con lo que se concluye que es correcta su elección en las traducciones) pero, además, aparece también la colocación *extra special* y lo hace un número mayor de veces que *premium* (*extra special* presenta 6 ocurrencias y *premium*, 5. Ej.: *extra special Spanish chorizo*¹⁹).

¹⁹ 013DMwsSD160629FoodieEN.txt

4.2.6. Saturado

Este término presenta un gran número de ocurrencias (en particular dentro del corpus paralelo), ya que el género textual en que basamos nuestro estudio son fichas descriptivas de producto en las que aparece la información nutricional de dicho producto. La única colocación para este término en español se corresponde con «de las cuales saturadas», que se traduce en el corpus paralelo por *of which saturated*. Sin embargo, esta colocación no aparece en el corpus comparable. Consideramos que el motivo por el que no aparece no es porque no se emplee, sino porque, en los textos originales en inglés, se omite la fórmula al exponer la información de una forma más esquemática. Por lo tanto, el hecho de que no aparezca la colocación *of which saturated* en los textos originalmente redactados en inglés no quiere decir que sea una mala traducción de «de las cuales saturadas»: simplemente indica que en los textos que conforman el corpus comparable no se da el caso de que aparezca. A continuación, presentamos un ejemplo para facilitar la comprensión.

Texto original en español ²⁰ :	Traducción al inglés ²¹ :	Texto original en inglés ²² :
Valores nutricionales (100 g) Valor energético: 323 Kcal / 1344 KJ Hidratos de carbono 1,8 g de los cuales: Azúcares: 1,8 g Proteínas: 27 g Grasas 23.1 de las cuales: Saturadas: 9,6 g Fibra alimentaria: No contiene Sodio: 1060 mg	Nutritional values (100 g) Energy value: 323 Kcal / 1344 KJ Carbohydrates 1.8 g of which: Sugars: 1.8 g Proteins: 27 g Fats 23.1 of which: Saturated: 9.6 g Fibre elementary: None Sodium: 1060 mg	Nutritional Facts Serving Size 2 oz (56g) Servings Per Container 8 Calories 130 Calories From Fat 90 Total Fat 10g Saturated Fat 3.5g Trans Fat 0g Cholesterol 40g Sodium 420mg Total Carbohydrates 2g Dietary Fiber 0g Sugars 0g Protein 8g

Tabla 3. Comparativa de texto original en español con su traducción al inglés y con texto original en inglés

Por otro lado, en los textos originales en inglés aparece también una abreviatura del término «saturado», que es *sat* (*sat fat*) y que no se encuentra en las traducciones al inglés de los textos en español.

4.2.7. Azúcar y azúcares

El siguiente término es «azúcar» que, como hemos mencionado con anterioridad, tiene un significado distinto dependiendo de si se usa en singular o en plural. Esta distinción se refleja también en inglés (*sugar* y *sugars*). Con respecto a *sugar* en singular, hemos comprobado que se trasvasa siempre por el equivalente establecido, que es el término que se emplea en los textos redactados por nativos en lengua inglesa. Por otro lado, las colocaciones de *sugars* en inglés en el corpus paralelo son *total sugars* y *of which sugars*, y ambas se recogen en el corpus comparable. En este caso, podemos observar que la construcción *of which sugars* sigue el mismo patrón que el descrito en el caso de *of which saturated* y se emplea en los textos originales en inglés, por lo que su traducción es correcta en ambos casos.

²⁰ 105DMwsLPPD180124FoodieES.txt

²¹ 105DMwsLPPD180124FoodieTEN.txt

²² 202DMwsCM190114FoodieEN.txt

4.2.8. Peso

Las colocaciones de las traducciones al inglés y de los textos originales coinciden, salvo en el caso de «peso de entrega», que se traduce al inglés por *delivery weight* y en el corpus comparable la colocación que aparece es *shipping weight*. De esto extraemos que la forma correcta sería la segunda y no la primera, por ser la que aparece en los textos originales.

4.2.9. Leche

El siguiente término es «leche», que recogía varias colocaciones. Por ejemplo, «leche en polvo», colocación que se traduce al inglés como *powdered milk* y como *milk powder*. En los textos redactados originalmente en inglés, solo se emplea la segunda, por lo que concluimos que *milk powder* es, en efecto, el equivalente establecido para la colocación «leche en polvo». Asimismo, en la primera parte del análisis de resultados, comentamos que, en las ocasiones en que en español aparece «leche y derivados», la traducción que se ofrece en inglés es *milk and dairy products* (literalmente, «leche y productos lácteos») y no *milk and derivatives*, que aparece en los textos originales y que, por lo tanto, sería el equivalente adecuado (con una estructura más próxima a la de la colocación española). Otras colocaciones como *milk protein* o *traces of milk* se reflejan en los textos originales, por lo que están bien traducidas.

4.2.10. Loncha

Observamos que la colocación *thin slices* se emplea en C-GEFEM, así que podemos afirmar que el trasvase interlingüístico se ha producido con éxito.

4.2.11. Fresco

Como hemos mencionado con anterioridad, este término presenta dos traducciones dentro del corpus paralelo: *cool* y *fresh*. Ambas formas aparecen también en el corpus comparable: la primera aparece en colocaciones como *in a cool dry place* («en lugar fresco y seco») o *keep cool* («mantener refrigerado»), que se pueden leer en las instrucciones de conservación de producto, y la segunda la podemos observar al fijarnos, por ejemplo, en la descripción de las características del chorizo (*the family Goikoa first began producing fresh chorizo sausages in 1929*²³).

Comentamos el primer caso, dado que las traducciones propuestas para la segunda denominación son adecuadas. La traducción que se le da a la colocación «en lugar fresco y seco» en los textos traducidos de español a inglés es *in a cool dry place*, por lo que es el equivalente establecido, ya que se usa en los textos originales en inglés. No obstante, en ocasiones se traduce el sintagma «en sitio fresco» como *at room temperature*, que significa en realidad «a temperatura ambiente». En estos casos, se utiliza el equivalente de otro término ya que, como acabamos de ver, el equivalente establecido para la colocación «en lugar fresco y seco» es *in a cool dry place* (en este caso, «en sitio fresco» > *in a cool place* y no *at room temperature*) ya que es, además, la que aparece en los textos originales en inglés. El hecho de traducir la colocación «en sitio fresco» como *at room temperature*, que significa «a temperatura ambiente», puede tener consecuencias, ya que la temperatura ambiente, por ejemplo, en California durante el verano no es

²³ 086DMwsCN161126FoodieEN.txt

precisamente fresca y el chorizo podría no conservarse de forma adecuada si el cliente sigue las instrucciones traducidas al inglés que se muestran en la página web.

4.2.12. Conservar

Este término tenía dos equivalentes en el corpus paralelo: *keep* y *store* que, de nuevo, tienen que ver con la conservación del producto. El análisis nos sirvió para concluir que las colocaciones más empleadas en los textos originales son *keep refrigerated* y *store in a cool dry place*, que aparecen también en las traducciones. Por lo tanto, salvo cuando se emplea «a temperatura ambiente», este término se trasvasa adecuadamente en las fichas descriptivas de producto.

4.2.13. Valor medio

También decidimos observar el comportamiento de la colocación «valor medio», que aparece al introducir la información nutricional de los productos para indicar que los valores que se muestran corresponden a 100 gramos de producto (Ej.: valor medio por 100 g de producto²⁴). En las traducciones al inglés, «valor medio» aparece traducido como *serving size*, que literalmente significa «tamaño de la ración». Esta traducción nos resultó extraña y, en efecto, al comprobar si se empleaba en los textos originales en inglés, verificamos que, la mayoría de las veces, en su lugar se utilizaba la colocación *typical values* (obtuvimos 218 resultados para *typical values* frente a apenas 24 ocasiones en las que aparecía *serving size*). Por esta razón, esa es la colocación que habría resultado más apropiado emplear en las traducciones al inglés.

4.2.14. Kilocaloría, gramo, kilogramo

En cuanto a los términos de estudio «kilocaloría», «gramo», «kilogramo» las traducciones son correctas ya que, en la mayoría de los casos, se emplea su símbolo (*kcal*, *g* y *kg*). No obstante, como hemos comentado en el apartado anterior, observamos errores tipográficos a la hora de escribir los símbolos, sobre todo dentro del corpus paralelo, es decir, en los textos en español y en las traducciones de los mismos.

4.2.15. Producto, peso neto, ingrediente

Por último, con respecto a términos como «producto», «peso neto» o «ingrediente» no hay mayores observaciones que destacar que las ya mencionadas.

4.3. LA BASE DE DATOS TERMINOLÓGICA

Una vez extraídos los equivalentes, hemos procedido a crear y alimentar la base de datos terminológica, que cuenta con 21 entradas. Como ejemplo, mostramos en las Figuras 7 y 8, en la página siguiente, dos capturas de pantalla, relativas a las fichas terminológicas de los términos «extra» e «información».

A continuación, ya expuestos y explicados los resultados de nuestro análisis, procederemos a extraer las principales conclusiones de nuestro trabajo en el siguiente capítulo.

²⁴ 140DMwsEZPD190121FoodieES.txt

The screenshot shows the SDL MultiTerm interface with the following details:

- Language:** Spanish (selected)
- Term:** extra
- ID de entrada:** 8
- FRAME1:** -
- DATE:** 15/03/2019
- AUTHOR:** ALT
- REV DATE:** 18/03/2019
- SPANISH Section:**
 - TERM REF: 162DMwsORPD190122FoodieES.txt
 - POS: adj.
 - EX: El Chorizo Extra Sarta Ortiz se elabora con la mejor carne de cerdo, sin gluten ni colorantes artificiales.
 - EX REF: 162DMwsORPD190122FoodieES.txt
 - PHRAS: chorizo ~, chorizo cular ibérico ~, chorizo ibérico ~
- ENGLISH Section:**
 - TERM REF: 089DMwsCN161130FoodieEN.txt
 - POS: adj.
 - CONT: "Extra" denomination means the highest quality being the "pimentón" or paprika the key ingredient.
 - CONT REF: 089DMwsCN161130FoodieEN.txt
 - EX: Dried thin sausage "fuet extra".
 - EX REF: 178DMwsTS190113FoodieEN.txt
 - VAR: extra special, premium, premium quality
 - VAR REF: 012DMwsSD160629FoodieEN.txt
 - PHRAS: chorizo ~, salchichon ~, ~ chorizo, ~ meat

Figura 7. Ficha terminológica del término «extra»

The screenshot shows the SDL MultiTerm interface with the following details:

- Language:** Spanish (selected)
- Term:** información
- ID de entrada:** 7
- FRAME1:** -
- DATE:** 15/03/2019
- AUTHOR:** ALT
- REV DATE:** 19/03/2019
- SPANISH Section:**
 - TERM REF: 122DmwsLQPD190206FoodieES.txt
 - ABB: info
 - ABB REF: 198DMwsANPD190207FoodieES.txt
 - POS: n. f.
 - EX: Información sobre alérgenos: Contiene derivados de la soja y de la leche. Sin gluten.
 - EX REF: 122DmwsLQPD190206FoodieES.txt
 - PHRAS: ~ sobre alérgenos, ~ nutricional
- ENGLISH Section:**
 - EX REF: 239DMwsME190115FoodieEN.txt
 - ABB: info
 - ABB REF: 136DMwsMR190113FoodieEN.txt
 - POS: n.
 - EX: You should always read the label before consuming or using the product and never rely solely on the information presented here.
 - EX REF: 239DMwsME190115FoodieEN.txt
 - PHRAS: product ~, other ~, additional ~, dietary ~, nutritional ~

Figura 8. Ficha terminológica del término «información»

5. CONCLUSIONES

El contraste interlingüístico que hemos desarrollado basándonos en los corpus paralelo y comparable nos ha permitido comprobar a partir de una muestra de análisis si ciertos elementos léxicos y fraseológicos relacionados con la terminología del chorizo presentes en las fichas descriptivas de este producto se habían trasladado adecuadamente del español al inglés. Para ello, hemos contrastado los términos y la fraseología utilizados en las traducciones con su empleo en textos redactados originalmente en lengua inglesa.

Cabe señalar que, durante la realización del presente trabajo se han cumplido los objetivos expuestos en el epígrafe del capítulo 1 dedicado a este asunto, como la puesta en práctica de los conocimientos adquiridos durante la titulación, por ejemplo, el desarrollo de la capacidad de compilación y análisis de un corpus paralelo y un corpus comparable, entre otros.

En primer lugar, para realizar el trasvase interlingüístico en este ámbito, consideramos que es un requisito imprescindible tener conocimientos de este campo en ambas lenguas, puesto que se trata de trasladar la terminología específica de un determinado campo del saber. De hecho, se trata de una tarea compleja ya que, en ocasiones, existen varias opciones de traducción para un mismo término y puede ocurrir que solo haya un equivalente adecuado en un contexto concreto, como hemos mostrado con los ejemplos de *soy protein/soya protein* o *milk powder/powdered milk* en el análisis de resultados.

En general, los textos traducidos no presentan errores muy graves que imposibiliten la comprensión pero, como hemos comentado en el capítulo anterior, hemos destacado una serie de errores de traducción en los textos traducidos al inglés que nos hacen reafirmarnos en la necesidad de contratar a profesionales en el sector cuando se desea trasladar el contenido de una página web o de cualquier otra plataforma. Un detalle que se pasa por alto a menudo es que dominar una lengua no significa ser capaz de traducir un texto, así que el resultado se corresponde con traducciones realizadas por personas que no se han formado previamente en traducción y que contienen errores por trasladar la terminología utilizando equivalentes literales (por ejemplo, traducir «peso de entrega» como *delivery weight* y no como *shipping weight*).

Asimismo, hemos visto una falta de exactitud en varias ocasiones al traducir algunos de los términos de la muestra de análisis, así como errores ortotipográficos en algunas de las páginas web en los textos redactados en español (tratándose de empresas españolas). Esto demuestra una falta de rigurosidad por parte de las empresas al exponer el contenido de sus páginas web, lo cual es de suma importancia cuando se intenta vender un producto (como en este caso el chorizo), ya que la página web de una empresa es la imagen que esta proyecta al mundo, como si fuese un escaparate en línea, y este puede ser un factor determinante a la hora de decidir si se adquiere un producto o no.

Como se puede comprobar con este trabajo, formar a traductores especializados en este campo del saber constituye una tarea esencial en los programas de formación de las facultades de Traducción e Interpretación, dado que es necesario que las empresas del sector agroalimentario cuenten con traductores profesionales y formados previamente para que la información de sus productos llegue a los potenciales clientes de forma adecuada.

Por lo tanto y a modo de resumen de las conclusiones de nuestro estudio, sostenemos que es necesario cuidar la imagen en línea de las empresas, tanto a través de la contratación de traductores profesionales que garanticen resultados de calidad como a través de la revisión y el cuidado de la redacción y la ortografía del contenido expuesto en ese «escaparate en línea».

6. BIBLIOGRAFÍA

Alvar-Ezquerro, M., Blanco-Rodríguez, M. J. & Pérez-Lagos, F. (1994). «Diseño de un corpus español en el marco de un corpus europeo». En M. Alvar-Ezquerro y J. A. Villena Ponsoda (Coords.), *Estudios para un corpus del español* (pp. 9-29). Málaga: Servicio de Publicaciones de la Universidad.

Anthony, L. (2018). *AntConc (Versión 3.4.4) [Software]*. Tokio, Japón: Universidad de Waseda. Recuperado de: <http://www.laurenceanthony.net/software> (Fecha de consulta: 26/05/2019).

Aston, G. (1999). «Corpus use and learning to translate». *Textus*, 12: 289-314.

Baker, M. (1993): «Corpus Linguistics and Translation Studies: Implications and Applications». En M. Baker, G. Francis y E. Tognini-Bonelli (Eds.), *Text and Technology: In honour of John Sinclair* (pp. 233-250). Ámsterdam/Filadelfia: John Benjamins.

Baker, M. (1996): «Corpus-based Translation Studies: The Challenges that Lie Ahead». En H. Somers (Ed.), *Terminology, LSP and Translation. Studies in Language Engineering in Honour of Juan C. Sager* (pp. 175-186). Ámsterdam/Filadelfia: John Benjamins.

Bowker, L. (1998): «Using Specialized Monolingual Native-Language Corpora as a Translation Resource: A Pilot Study». *Meta*, 43, (4), 631-651. Recuperado de: <https://doi.org/10.7202/002134ar> (Fecha de consulta: 28/04/2019).

Bowker, L. (1999a): «Using a corpus to access student translations: a pilot study». En B. Lewadowska-Tomaszczyk y P. J. Melia (Eds.), *PALC'99: Practical Applications in Language Corpora* (pp. 529-540). Bern: Peter Lang.

Bowker, L. (1999b): «The Design and Development of a Corpus-based Aid for Assessing Translations». *Teanga*, 18: 11-24.

Bowker, L. (2003): «Corpus-based applications for translator training: Exploring possibilities». En S. Granger, S. Lerot y J. Petch-Tyson (Eds.), *Corpus-based approaches to Contrastive Linguistics and Translation Studies* (pp. 169-183). Ámsterdam/Nueva York: Rodopi.

Bowker, L. y Pearson, J.: (2002): *Working with Specialized Language: A practical guide to using corpora*. Londres: Routledge.

Corpas Pastor, G. (2001): «Compilación de un corpus ad hoc para la enseñanza de la traducción inversa especializada». *Trans*, 5, 155-184. DOI: 10.24310. Disponible en: <http://www.revistas.uma.es/index.php/trans/article/view/2916/2710> (Fecha de consulta: 25/03/2019).

Corpas Pastor, G. (2002): «Traducir con corpus: de la teoría a la práctica». En J. García Palacios y M. T. Fuentes Morán (Eds.), *Texto, Terminología y Traducción* (pp. 189-226). Salamanca: Almar.

Corpas Pastor, G. (2004): «Localización de recursos y compilación de corpus vía Internet: aplicaciones para la didáctica de la traducción médica especializada». En C. Gonzalo García y V. García Yebra (Eds.), *Manual de documentación y terminología para la traducción especializada* (pp. 223-257). Madrid: Arco.

Corpas Pastor, G. (2012): «Corpus, Tecnología y Traducción». En M. Casas Gómez (Dir.) y M. García Antuña (Ed.), *XII Jornadas de Lingüística* (pp. 1-21). Cádiz: Servicio de Publicaciones de la Universidad de Cádiz.

Corpas Pastor, G. y Seghiri, M. (2009): «Virtual Corpora as Documentation Resources: Translating Travel Insurance Documents (English-Spanish)». En A. Beeby, P. Rodríguez Inés y P. Sánchez-Gijón (Eds.), *Corpus Use and Translating: Corpus Use for Learning to Translate and Learning Corpus Use to Translate* (pp. 75– 107). Ámsterdam/Filadelfia: John Benjamins.

Danielsson, D y Ridings, P. (1996): «Corpus and Terminology: Software for the Translation Program at Göteborgs Universitet, or Getting Students to do the Work». En S. Botley, J. Glass, T. McEnery y A. Wilson (Eds.), *Proceedings of Teaching and Language Corpora 1996* (pp. 57–67). Technical Paper 9. Lancaster: University Centre for Computer Corpus Research on Language.

Epstein, B. J. (2009): «What's cooking: translating food». *Translation Journal*, 13 (3). Recuperado de: <https://translationjournal.net/journal/49cooking.htm> (Fecha de consulta: 02/04/2019).

Faya Ornia, G. (2014): «Propuesta y clasificación de corpus», *Babel*, 60 (2), 234-252.

Fletcher, W. H. (2004): «Facilitating the Compilation and Dissemination of Ad-Hoc Web Corpora». En G. Aston, S. Bernardini y D. Stewart (Eds.), *Corpora and Language Learners*, (pp. 271–300). Ámsterdam: John Benjamins.

Forcada M. L. (2010): «Machine Translation Today». En Y. Gambier, L. Van Doorslaer (Eds.), *Handbook of Translation Studies*, vol. 1 (pp. 215-223). Ámsterdam/Filadelfia: John Benjamins.

Fundéu BBVA (24 de enero de 2012): «Símbolos y abreviaturas, claves de redacción» [Entrada en blog]. Fundéu BBVA. Recuperado de: <https://www.fundeu.es/recomendacion/seis-claves-para-usar-las-siglas-y-las-abreviaturas-1189/> (Fecha de consulta: 26/04/2019).

Granger, S. (2003): «The corpus approach: a common way forward for Contrastive Linguistics and Translation Studies». En S. Granger, J. Lerot y S. Petch-Tyson (Eds.), *Corpus-based Approaches to Contrastive Linguistics and Translation Studies* (pp. 17-30). Ámsterdam/Nueva York: Rodopi.

Laviosa, S. (1997), «How Comparable Can 'Comparable Corpora' Be?». *Target*, 9 (2), 289-319.

Laviosa, S. (2010): «Corpora». En Y. Gambier y L. Van Doorslaer, *Handbook of Translation Studies. Volume 1*. Ámsterdam/Filadelfia: John Benjamins.

Leech, G. (1992): «Corpora and theories of linguistic performance». En J. Startvik (Ed.), *Directions in corpus linguistics* (pp. 105-122). Berlin: Mouton de Gruyter.

Liquidst (1999): «Electronic Corpora as Tools for Translation». En G. Anderman y M. Rogers (Eds.), *Words, Text, Translation: Liber*.

McEnery, T. y Hardie, A. (2012): *Corpus Linguistics: Method, Theory and Practice*. Cambridge: Cambridge University Press.

Olohan, M. (2004): *Introducing Corpora in Translation Studies*. Londres/Nueva York: Routledge.

Ortego Antón, M. T. (2019): *La terminología del sector agroalimentario (español-inglés) en los estudios contrastivos y de traducción especializada basados en corpus: los embutidos*. Berlín: Peter Lang.

Ortego Antón, M. T. y Fernández Nistal, P. (en prensa): «Estudio contrastivo de la terminología de embutidos en inglés y en español con ParaConc y tiCorpus a partir del corpus paralelo P-GEFEM y del comparable C-GEFEM». En Míriam Seghiri (Ed.), *El uso de los corpus lingüísticos como herramienta pedagógica para la enseñanza-aprendizaje de lenguas, traducción e interpretación*. Berna: Peter Lang.

Pearson, J. (1998): *Terms in Context*. Amsterdam/Philadelphia: John Benjamins.

Rabadán, R. y Fernández Nistal, P. (2002): *La traducción inglés-español, fundamentos, herramientas, aplicaciones*. Universidad de León.

Rivas Carmona, M. M y Veroz González, M. A. (2018): *Agroalimentación: lenguajes de especialidad y traducción*. Granada: Comares.

Sánchez-Gijón, P. (2003a): «És la web pública la nova biblioteca del traductor?». *Tradumàtica: Traducció i tecnologies de la informació i la comunicació 2*. Recuperado de: <http://www.fti.uab.es/tradumatica/revista/num2/articles/07/07art.htm> (Fecha de consulta: 03/05/2019).

Sánchez-Gijón, P. (2003b): *Els documents digitals especialitzats: utilització de la lingüística del corpus com a front de recursos per a la traducció* [Tesis doctoral]. Barcelona: Universidad Autónoma de Barcelona.

Sánchez Trigo, E. (2005): «Investigación traductológica en la traducción científica y técnica». *Trans*, 9, 131-148.

Seghiri, M. (2011): «Metodología protocolizada de compilación de un corpus de seguros de viajes: aspectos de diseño y representatividad». *RLA. Revista de Lingüística Teórica y Aplicada*, 49 (2), 13-30. DOI: 10.4067. Recuperado de: https://scielo.conicyt.cl/scielo.php?script=sci_arttext&pid=S0718-48832011000200002 (Fecha de consulta: 25/04/2019).

Seghiri, M. (2017): «Metodología de elaboración de un glosario bilingüe y bidireccional (inglés-español/ español-inglés) basado en corpus para la traducción de manuales de instrucciones de televisores». *Babel*, 63 (1), 43-64. DOI: 10.1075/babel.63.1.04seg (Fecha de consulta: 10/04/2019).

Sinclair, J. (1996): *Preliminary Recommendations on Corpus Typology*. Birmingham, Reino Unido: EAGLES. Disponible en: <http://www.ilc.cnr.it/EAGLES96/corpusstyp/corpusstyp.html> (Fecha de consulta: 02/04/2019).

Sinclair, J. (2005): «Corpus and Text: Basic principles». En M. Wynne (Ed.), *Developing Linguistic Corpora: a Guide to Good Practice*. Oxford: Oxbow books. Disponible en: <http://ota.ox.ac.uk/documents/creating/dlc/chapter1.htm> (Fecha de consulta: 10/04/2019).

Torruella, J. y Llisterri, J. (1999): «Diseño de corpus textuales y orales». En J. M. Blecua; G. Clavería; C. Sánchez; J. Torruella (Eds.); *Filología e informática: nuevas tecnologías en los estudios filológicos* (pp. 45-77). Barcelona: Editorial Milenio.

Varantola, K. (1997): «Translators, Dictionaries and Text Corpora». En S. Bernardini y F. Zanettin (Eds.), *I corpora nella didattica della traduzione*. Bolonia: CLUEB, 117-133.

Varantola, K. (2003). «Translators and Disposable Corpora». En F. Zanettin, S. Bernardini y D. Stewart (Eds.), *Corpora in Translator Education* (pp. 55-70). Manchester: St Jerome.

Zanettin, F. (1998): «Bilingual Comparable Corpora and the Training of Translators». *Meta*, 43 (4), 616-630.

Zanettin, F. (2002): «Corpora in Translation Practice», *Language Resources for Translation Work and Research* (pp. 10-14).

Zanettin, F. (2012): *Translation-Driven Corpora: Corpus Resources for Descriptive and Applied Translation Studies*. Manchester: St Jerome.