

Las fichas descriptivas de embutidos en español y en inglés: Un análisis contrastivo de la estructura retórica basado en corpus

Dried-meats Descriptive Cards in Spanish and English: A Corpus Based Study of the Rhetoric Structure

María Teresa Ortego-Antón

UNIVERSIDAD DE VALLADOLID
ESPAÑA

mariateresortego@uva.es

Recibido: 04-VII-2018 / **Aceptado:** 22-IV-2019

DOI: 10.4067/S0718-09342020000100170

Resumen

La internacionalización de la industria agroalimentaria en España va aparejada de un aumento exponencial de los servicios de redacción y de traducción del español al inglés. En el marco de esta necesidad pretendemos realizar un estudio contrastivo de la estructura retórica de las fichas descriptivas de embutidos en español y en inglés a partir del diseño, la compilación y la anotación retórica de un corpus virtual comparable, C-GEFEM. Dicho estudio nos permitirá detectar los patrones de comportamiento retórico de un determinado género textual, las fichas descriptivas de embutidos, en las lenguas española e inglesa. Con los resultados que se desprendan del análisis de la estructura retórica a partir de la anotación de los movimientos y de los pasos del corpus con el Etiquetador de movimientos retóricos podremos vislumbrar similitudes y diferencias en las lenguas española e inglesa en este género textual para, en posteriores etapas, diseñar GEFEM, un generador de fichas descriptivas de embutidos basado en el procesamiento del lenguaje natural y destinado a satisfacer las necesidades de comunicación especializada del español al inglés de traductores y redactores del sector agroalimentario.

Palabras Clave: Estructura retórica, fichas descriptivas de embutidos, corpus, inglés, español.

Abstract

The internationalization of the agri-food sector in Spain has increased the need for drafting and translation services from Spanish into English. Within this socioeconomic context, our proposal is to carry out a cross-linguistic study in order to address the rhetoric structure of dried-meat descriptive cards in Spanish and English based on the design, compilation and annotation of a virtual comparable corpus (C-GEFEM). This study will allow us to identify the rhetorical patterns of a specific genre, product descriptive cards in Spanish and English. Results from the analysis of the rhetoric structure will show similarities and differences between Spanish and English, so as to model GEFEM, a tool based on natural language processing which automatically generates dried-meat descriptive cards to satisfy the needs of translators and writers who draft specialized texts about dried meats for the agri-food sector.

Key Words: Rhetoric structure, dried-meat descriptive cards, corpus, English, Spanish.

INTRODUCCIÓN

El aumento de las exportaciones de productos del sector agroalimentario español a la Unión Europea y a terceros países hace necesario, más que nunca, profesionales de la traducción y de la redacción formados en este campo para asistir a las empresas durante la adaptación de los datos de sus productos (páginas web, fichas descriptivas de producto, fichas técnicas, etiquetado, información nutricional, etc.) al mercado internacional, con el objetivo de atraer nuevos consumidores y de fidelizar a los ya existentes.

De hecho, la industria agroalimentaria es el sector económico de mayor importancia en España (Marco Estratégico para la Industria de Alimentación, MAPAMA, 2017) y, además, es el cuarto sector en Europa y el octavo en el mundo. Entre las industrias más relevantes, la industria cárnica es la primera en lo que respecta a facturación y a empleos directos. Asimismo, esta industria contribuye a potenciar, sostener y mejorar las áreas rurales, sobre todo con la creación de empleos para evitar la despoblación. Sin embargo, la mayoría de las industrias pertenecientes a este sector se corresponde con pequeñas y medianas empresas (96,5%), característica que se convierte en un obstáculo a la hora de tener acceso a la internacionalización, a la innovación y a la mejora de la productividad.

Dentro del sector cárnico, los embutidos constituyen una parcela de conocimiento que no se ha estudiado en profundidad, a pesar de la creciente necesidad de las empresas cárnicas por exportar sus productos al exterior. Por tanto, las empresas de embutidos necesitan que se desarrollen herramientas lingüísticas que les ayuden durante el proceso de internacionalización y que automaticen los servicios de redacción y de traducción del español al inglés.

En consecuencia, en este trabajo tenemos como objetivo principal establecer la estructura retórica de un determinado género textual, las fichas descriptivas de embutidos en las lenguas española e inglesa a partir del diseño, la compilación, la anotación retórica y el análisis de un corpus virtual comparable (C-GEFEM) para, en posteriores etapas, una vez establecida la estructura retórica, desarrollar una herramienta basada en el procesamiento del lenguaje natural que automatice los servicios de traducción y redacción a la lengua inglesa¹. Partimos de la hipótesis de que la estructura retórica de las fichas descriptivas de embutidos tiene convenciones diferenciadas en español y en inglés.

Para validar dicha hipótesis, en este trabajo comenzamos definiendo el género objeto de estudio, las fichas descriptivas de embutidos, y revisando los estudios previos que focalizan la atención en el análisis de la estructura retórica desde una perspectiva contrastiva (sección 1). A continuación, describimos el protocolo que hemos empleado para diseñar y compilar C-GEFEM (sección 2). En la sección 3 explicamos la metodología que utilizamos para analizar la estructura retórica de los textos que componen C-GEFEM, cuyos resultados se exponen en la sección 4. Por último, finalizamos extrayendo las conclusiones.

1. Las fichas descriptivas de embutidos

Las fichas descriptivas de embutidos se corresponden con un determinado género textual. El concepto de género textual es muy amplio y han sido múltiples los autores que han ofrecido una definición, ya sea desde la perspectiva de la lingüística sistémico-funcional (Martin, 1997; Eggins & Martin, 2000, 2003), de la lingüística textual (Heinemann, 2000), de la lingüística de corpus (el grupo ACTRES), de los lenguajes de especialidad o lenguas para fines específicos (Bhatia, 1993; Swales, 1990/2001) y de los Estudios de Traducción (García Izquierdo, 2002). Dado que el objeto de estudio de este trabajo no es adentrarnos en la definición de género, vamos a emplear la definición propuesta por Biber, Connor y Upton (2007: 9), que conciben el género como “the internal structure and organization of texts from a specific variety”. Desde la perspectiva de los géneros textuales, las fichas descriptivas de embutidos no han recibido la suficiente atención, a pesar de que, actualmente, son clave para las empresas cárnicas, puesto que la información en ellas recogida puede resultar decisiva para los consumidores a la hora de decantarse por un determinado producto. En la Figura 1 se ofrece un ejemplo de este género textual en español y en la Figura 2 se muestra una ficha descriptiva de producto en lengua inglesa.

CHORIZO EXTRA SARTA NATURAL DULCE

Categoría: Chorizos

Un gran producto elaborado con las mejores carnes del Cerdo, y con ingredientes totalmente naturales, como son Pimentón ahumado de la Vera, ajo y sal. No contiene ningún tipo de aditivos, ni conservantes ni colorantes, ni artificiales ni naturales.

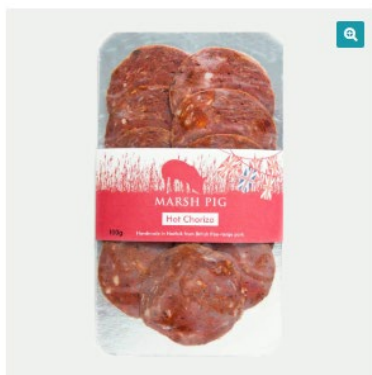
Apto para su consumo crudo, en lonchas o tacos, o cocinado con cualquier plato en que pueda intervenir chorizo, su exquisito sabor y delicado bouquet le dejará gratamente sorprendido.



INGREDIENTES Y OTROS DATOS

- **Materias Primas:**
 - Paletilla y Magro: 50%
 - Panceta: 50%
- **Condimentos y especias:** Sal, pimentón dulce de la Vera y ajo
- **Tipo de curación:** Mixta. Un primer proceso de 7 días en secadero climatizado y un segundo proceso de 14 días en secaderos naturales con carbón vegetal y aire.
- **Otros datos:** Tripa natural de cerdo. Envasado en atmósfera modificada pieza por pieza y empaquetado en cajas de cartón a 18 piezas por caja. El peso aproximado por pieza es de 330 grs. Consumir preferentemente antes de 180 días.
- **Información Nutricional por 100 gr de producto**
 - Nivel calórico: ± 419 Kcal ± 1740 KJ
 - Proteínas: ± 27.60 g
 - Hidratos de Carbono: ± 0.60 g
 - Grasa: ± 34.10 g
 - Contenido en sal: ± 3.74 g
 - Colesterol: ± 70 mg

Figura 1. Ejemplo de ficha descriptiva de embutido en español.



Hot Chorizo

£5.00

We cranked up the spice in our chorizo by using Pimenton from del la Vera to create a hot chorizo that packs a punch with every mouthful.

1

Add to cart

Categories: [All Products](#), [Sliced Salami & Chorizo](#)

Tags: [CHILLI](#), [Chorizo](#), [free range](#), [pork](#), [Spicy](#)



Nutrient	Unit	per 100g
Energy	kcal	382
Energy	kJ	1585

Figura 2. Fragmento de ficha descriptiva de embutido en inglés.

Para abordar en profundidad un género, Rabadán (2016) señala que es necesario conocer las partes retóricas que lo integran, la función comunicativa y las convenciones asociadas a cada parte. Por lo que respecta a la función comunicativa de las fichas descriptivas de embutidos, la función prioritaria de este género textual se corresponde con informar, si bien en algunos fragmentos de la ficha, por ejemplo, en

la descripción del producto, puede aparecer, de forma secundaria, la función persuasiva.

En relación con las convenciones asociadas a cada parte retórica que integra este género, la retórica contrastiva ofrece la posibilidad de estudiar las diferencias y similitudes entre distintos patrones lingüísticos y la forma de estructurar un texto en dos lenguas diferentes. Chesterman (1998: 153) apunta a que su finalidad es “to establish the nature and range of possible variants of message-expression, and determine the appropriateness conditions for their use in different cultures”. En consecuencia, el contraste en dos lenguas de la estructura retórica nos permitirá observar las similitudes y diferencias para poder realizar un trasvase lingüístico que se adecue a las convenciones utilizadas en lengua inglesa, lo que será de ayuda para traductores y redactores no nativos.

A pesar de que existen numerosos trabajos sobre el análisis de los distintos géneros, ya sea desde un enfoque monolingüe en español (Martínez Hincapié, 2015; Parodi, 2015; Parodi & Burdiles, 2015; Venegas, Zamora & Galdames, 2016, entre otros), en inglés (Swales, 2004; Kwan, 2006) o desde una perspectiva contrastiva, en portugués y español (Navarro, 2015) o en las lenguas inglesa y española, por ejemplo, Rabadán (2016) se centra en el estudio de las actas, Martín-Martín (2005) y López Arroyo, Fernández-Antolín y de Felipe-Boto (2007) en los resúmenes científicos, Méndez-Cendón (2009) en los informes de casos médicos o Williams (2010) en los artículos de investigación en medicina, no hemos encontrado trabajos que se centren en este género textual. No obstante, desde el grupo interuniversitario ACTRES (Análisis Contrastivo y Traducción Inglés-Español)² son varios los estudios realizados en el marco del sector agroalimentario: Labrador, Ramón, Alaiz-Moretón y Sanjurjo-González (2014) estudian las descripciones de queso en línea en inglés, López Arroyo y Roberts (2014) focalizan la atención en las notas de cata del vino y su fraseología y Ortego-Antón (2019) y Ortego-Antón y Fernández Nistal (2019) en la terminología de los embutidos y su fraseología.

Ante la ausencia de estudios previos, consideramos pertinente abordar contrastivamente la estructura retórica de las fichas descriptivas de embutidos en las lenguas española e inglesa. En este sentido, López Arroyo y Roberts (2014: 155) definen la estructura retórica como “the hierarchic organization of a text. It involves the various sections and subsections of a text, moves and steps”. Las fichas descriptivas de embutidos, como cualquier género textual, se caracteriza por estar formado por una serie de componentes retóricos denominados *moves* o movimientos (Biber et al., 2007: 23), que estos autores definen como:

“a section of a text that performs specific communicative function. Each move not only has its own purpose but also contributes to the overall communicative purposes of a genre”.

De hecho, algunos movimientos tienen lugar con mayor frecuencia que otros, de manera que podemos distinguir entre movimientos convencionales y movimientos opcionales. A su vez, los movimientos pueden dividirse en varios *steps* o pasos, cuya función es “to achieve the purpose of the move to which it belongs” (Biber et al., 2007: 24).

Para identificar los principales movimientos y pasos que configuran la estructura retórica de las fichas descriptivas de embutidos es necesario diseñar, compilar y anotar un corpus virtual comparable en las lenguas española e inglesa compuesto por fichas descriptivas de embutidos del que extraer los movimientos (*moves*) y los pasos (*steps*) principales y opcionales. Así, podremos establecer un prototipo de estructura en cada una de las lenguas de trabajo que asistirá a hablantes hispanos que redactan en lengua inglesa textos pertenecientes a este género textual.

2. El diseño y la compilación de C-GEFEM

2.1. El diseño de C-GEFEM

C-GEFEM es un corpus, es decir, un conjunto 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 (Sinclair, 2005). Por tanto, el empleo de una metodología basada en el análisis de los datos de un corpus nos permite fundamentar nuestra investigación en datos reales de la lengua en uso. La segunda característica de nuestro corpus se corresponde con comparable: “a comparable corpus is one which selects similar texts in more than one language or variety” (EAGLES, 1996).

Además, es un corpus virtual³, que Corpas Pastor (2008: 91) define como “a corpus in which there are not many texts but that the few texts included are suited to the field of knowledge, genre and textual variety.” Entre las ventajas de utilizar corpus comparables, Sánchez Trigo (2005: 138) señala:

“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”.

Corpas Pastor y Seghiri (2009) añaden otras ventajas tales como la objetividad, la reutilización y los usos múltiples de un único recurso. Asimismo, los corpus virtuales son fáciles de utilizar y permiten el acceso y la gestión a grandes cantidades de información en poco tiempo.

Sin embargo, los corpus que encontramos en Internet son, frecuentemente, de carácter general y no se centran en recopilar textos que cubran únicamente el género textual que deseamos documentar, por lo que no queda otra alternativa que compilar un corpus que se ajuste a las necesidades de un determinado estudio, una opción avalada previamente por autores como Seghiri (2006) o Zannettin (2012). A ello se suma el hecho de que el uso de Internet implica una serie de dificultades, como apunta Austermühl (2001: 52):

“[f]inding data on the worldwide is not problem at all. But finding reliable information is rather a difficult task. And finding the information you really need can be very time-consuming and often frustrating”.

Por consiguiente, a la hora de compilar un corpus hemos seguido un protocolo (Seghiri, 2015, 2017) para asegurar la calidad, de manera que el corpus compilado pueda caracterizarse por ser equilibrado y representativo (Bowker, 2002; McEney & Hardie, 2012; Seghiri, 2017). Un corpus puede considerarse equilibrado si el tamaño de cada una de las secciones que lo componen se escoge de manera que sea representativo de la lengua que se pretende estudiar (McEney & Hardie, 2012). Además, será representativo si contiene la proporción de textos necesaria para que el contenido del corpus sea un reflejo de la lengua que se pretende mostrar (McEney & Hardie, 2012).

Por tanto, nuestro corpus cumplirá con los parámetros de equilibrio y de representatividad, puesto que estará compuesto por 100 textos, es decir, 100 fichas descriptivas de embutidos —chorizo, lomo embuchado, salchichón y chistorra—, en cada una de las lenguas de trabajo: español e inglés. Serán comparables entre sí porque se corresponden con fichas descriptivas de producto redactadas originalmente en español o en inglés. Respecto a la procedencia, en el caso de la lengua inglesa, los textos están extraídos de empresas de Reino Unido (65%), Estados Unidos (15%), Canadá (10%), Irlanda (6%) y Australia (4%), para asegurar que todas las variedades de la lengua inglesa están representadas. Las fichas descriptivas de producto en español procederán solo de España, puesto que este trabajo se enmarca en un proyecto de mayor envergadura se corresponde con el desarrollo de una herramienta de asistencia durante la redacción de fichas descriptivas de embutidos del español al inglés y que estará destinada a profesionales españoles del sector agroalimentario. No obstante, dejamos constancia de que la muestra procederá de, al menos, 30 empresas de reconocido prestigio del sector.

2.2. Protocolo de compilación de C-GEFEM

Una vez establecidos los parámetros, procedemos a presentar los pasos seguidos para compilar el corpus basándonos en la metodología de Seghiri (2015, 2017), que consta de cuatro fases: la búsqueda, la descarga, el formato y el almacenamiento.

2.2.1. Búsqueda

En la primera fase hemos localizado los textos que constituirán nuestro corpus. Hemos tenido que realizar distintas búsquedas en *Internet* para encontrar fichas descriptivas de embutidos en inglés y en español. La búsqueda que mejores resultados ofreció fue explotar las páginas de empresas agroalimentarias de reconocido prestigio en la comercialización de embutidos como pueden ser Sainsbury's, Morrisons, Ocado, Quijote Foods, Enrique Tomás, Carrefour, etc.

2.2.2. Descarga

En la segunda etapa hemos descargado las fichas descriptivas de embutidos desde el sitio web que las alberga para guardarlos, posteriormente, en el ordenador. La descarga de textos se ha hecho de forma manual guardándolos en formato HTML.

2.2.3. Formato

Para que los textos puedan ser procesados por los programas de gestión de corpus es necesario convertirlos a formato TXT (UTF8). En consecuencia, nos hemos asegurado de que todos los archivos tuviesen el mencionado formato, así que hemos guardado los textos en formato TXT seleccionando el contenido de la página *web*, copiándolo y pegándolo en un documento de texto. En otras ocasiones, las fichas descriptivas de embutidos aparecen en formato PDF, así que las hemos descargado y convertido a TXT (UTF8) con la ayuda del programa Adobe Acrobat Professional.

2.2.4. Almacenamiento

En esta última fase hemos codificado todos los archivos que componen el corpus C-GEFEM en carpetas y subcarpetas. Hemos creado dos carpetas. En la primera de ellas, denominada 'BIBLIOTECA DIGITAL', que se divide en dos subcarpetas, 'EN' para las páginas web en inglés y 'ES' para las páginas en español, hemos guardado los archivos HTML clasificados por lenguas.

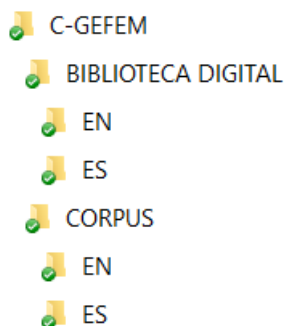


Figura 3. Estructura de C-GEFEM.

La segunda carpeta, denominada ‘CORPUS’, también se divide en dos subcarpetas, ‘EN’ y ‘ES’. En dichas carpetas hemos guardado los textos en formato TXT.

Una vez creada la estructura de almacenamiento, procedemos a denominar los archivos con un número (001, 002, etc.), ‘DM’ como abreviatura de *dried meats*, que es la temática del corpus, ‘ws’ para indicar que los textos han sido extraídos de la web, la abreviatura de la empresa de la que proceden (por ejemplo, MR para Morrisons, MS para Moreno Sáez, ET para Enrique Tomás, etc.), la fecha en la que fueron descargados (aammdd), el dominio (Foodie) y la lengua (EN o ES), de manera que los textos siguen la siguiente codificación: 001DMwsMR160624FoodieEN.txt, 001DMwsBH160625FoodieES.txt, etc. De hecho, esta codificación permitirá, en el futuro, la posible ampliación del corpus.






 001DMwsMR160624FoodieEN	31/10/2016 15:14	Documento de tex	3 KB
 002DMwsTS160624FoodieEN	31/10/2016 15:43	Documento de tex	3 KB
 003DMwsMS160624FoodieEN	03/11/2016 18:20	Documento de tex	2 KB
 004DMwsSBS160624FoodieEN	03/11/2016 18:28	Documento de tex	3 KB
 005DMwsSBS160624FoodieEN	03/11/2016 18:31	Documento de tex	3 KB

Figura 4. Ejemplo de denominación de archivos.

El resultado es un corpus virtual comparable bilingüe de fichas descriptivas de embutidos integrado por 100 textos en inglés (25425 *tokens* o casos) y 100 textos en español (14196 *tokens* o casos), que es representativo a nivel cualitativo gracias a los parámetros de diseño y protocolo de compilación seguidos. La diferencia en el número de palabras en cada lengua se debe a que, en lengua inglesa, se especifican las características del producto, el embalaje y la preparación y uso, en tanto que en español la información es más sintética, probablemente por las diferencias culturales existentes. Para concluir el proceso, solo queda comprobar la representatividad a nivel cuantitativo.

2.3. La representatividad cuantitativa

Para determinar la representatividad cuantitativa de la muestra utilizamos el programa ReCor⁴. Seghiri (2017: 50) define representatividad cuantitativa como “la terminología básica empleada en este género”. Tras analizar cada subcorpus (inglés y español), el programa genera las gráficas de representatividad (Figura 5 y Figura 6).

En la primera representación gráfica del corpus que ReCor genera (Figura 5 – Estudio gráfico A) se presenta en el eje horizontal el número de archivos del corpus, mientras que en el eje vertical se muestra el cociente tipos/casos (T_j/to). Se recogen dos funciones, una para los archivos ordenados alfabéticamente (línea roja) y otra (línea azul) para los archivos elegidos aleatoriamente, de forma que nos aseguramos,

mediante doble comprobación, de que el orden de los textos no repercute en la representatividad del corpus. Ambas funciones representan un descenso exponencial al seleccionar más textos. Sin embargo, cuando ambas funciones se estabilizan (línea roja y línea azul) podemos afirmar que el corpus es representativo y, precisamente, en ese punto, se podrá observar con qué número de textos el corpus es representativo cuantitativamente. Simultáneamente se genera otra gráfica (Estudio gráfico B) en la que se ofrece en el eje horizontal el número de casos (*tokens*) a partir del que se puede extraer, además, el número de palabras mínimo que debe incluir el corpus para ser representativo en lo relativo a la terminología básica empleada en este género.

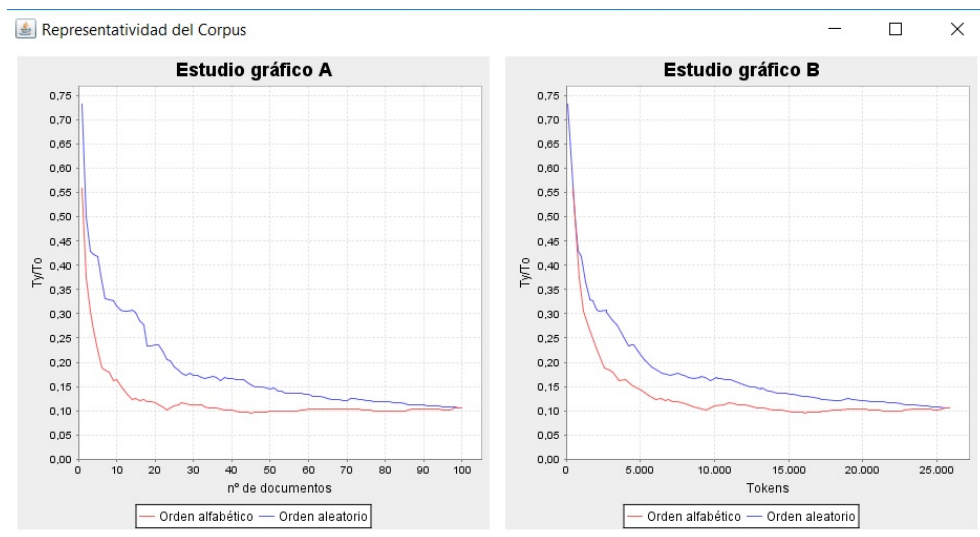


Figura 5. Representatividad cualitativa del subcorpus en inglés de C-GEFEM calculada con ReCor.

De la observación de la Figura 5 se desprende que el subcorpus de fichas descriptivas de embutidos en inglés es representativo a partir de los 90 documentos (Estudio gráfico A) y las 25000 palabras (Estudio gráfico B), puesto que en ambas gráficas las líneas se aproximan, se estabilizan y la curva apenas varía antes y después de las mencionadas cifras. Asimismo, los datos que se desprenden del análisis estadístico muestran que a partir del texto 90 el número de nuevos tipos al incorporar más textos es siempre inferior a 4.

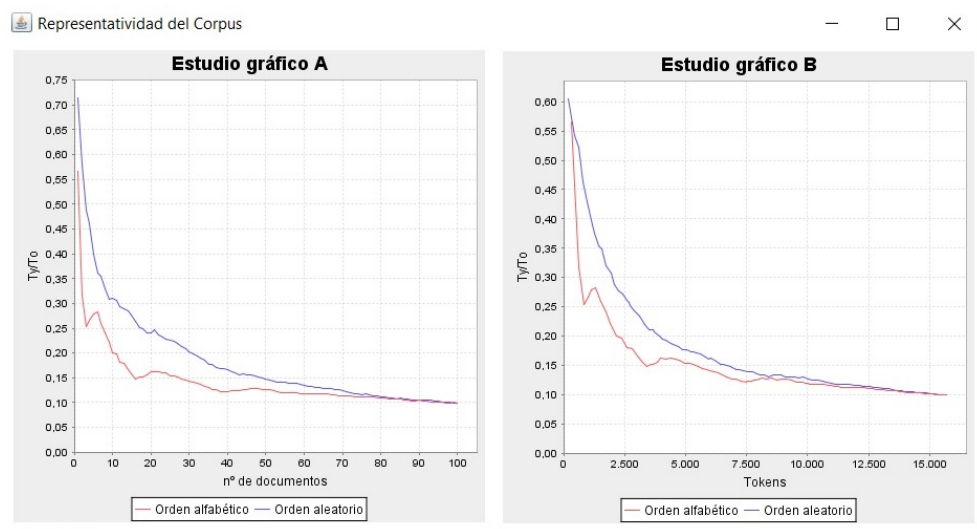


Figura 6. Representatividad cualitativa del subcorpus en español de C-GEFEM calculada con ReCor.

Por lo que respecta al subcorpus en español, en la Figura 6 se puede comprobar que este es representativo a partir de los 80 documentos y aproximadamente 12500 palabras.

Si comparamos los gráficos producidos por ReCor, se aprecia que el programa ha utilizado diferentes umbrales en el Estudio gráfico B para mostrar la representatividad cuantitativa, debido a que la longitud de las fichas descriptivas de embutidos es más amplia en inglés que en español.

En consecuencia, nuestro corpus se caracteriza por ser representativo desde el punto de vista cualitativo, a través de los parámetros de diseño y el protocolo de compilación, así como cuantitativo, mediante los resultados del programa ReCor.

3. Metodología de análisis

Puesto que en la literatura previa no hemos encontrado ningún modelo que identifique los movimientos típicos de las fichas descriptivas de embutidos, teniendo en cuenta que “you determine the code for a chunk of data by careful reading and reflection on its core content and meaning” (Miles, Huberman & Saldana, 2014: 73), en una primera fase hemos revisado una muestra de diez fichas de producto en cada una de las lenguas (español e inglés) y hemos diseñado una propuesta de etiquetas basándonos en nuestro conocimiento previo y en la observación de los posibles movimientos y pasos que podrían constituir la estructura retórica de dichos textos en cada lengua. Este procedimiento ha sido utilizado previamente por Cristobalena Frutos (2016) para estudiar los manuales de instrucciones de electrodomésticos.

Las etiquetas propuestas (Figura 7), que se corresponden con los posibles movimientos y pasos, se emplearán para anotar C-GEFEM con ayuda de las distintas herramientas desarrolladas por el grupo de investigación interuniversitario ACTRES.

<Additives>	<FollowOn>	<ProductCode>
<Address>	<ImageDriedmeats>	<ProductInformation>
<AllergenInfo>	<ImageLogo>	<ProductName>
<Brand>	<ImageNutrition>	<RecyclingInfo>
<BrandDescription>	<Ingredients>	<ReturnAddress>
<Comments>	<Manufacturer>	<ShareOn>
<CompanyName>	<ManufacturerAddress>	<Slogan>
<ConceptInfo>	<Map>	<StorageInfo>
<ContryofOrigin>	<NutritionalValues>	<SuitableFor>
<CuringPeriod>	<Other>	<Telephone>
<Description>	<PackagingInfo>	<Webpage>
<Download>	<PackedCountry>	<Weight>
<eMail>	<Preparation&Use>	<WriteReview>

Figura 7. Propuesta de etiquetas.

3.1. El Constructor de etiquetadores

Con la ayuda del Constructor de etiquetadores desarrollado por el grupo interuniversitario ACTRES introducimos las posibles etiquetas que harán referencia a los movimientos y a los pasos para cargarlas y que estén disponibles en el Etiquetador de movimientos retóricos®.

3.2. El etiquetador de movimientos retóricos

A continuación, con la ayuda del Etiquetador de movimientos retóricos®⁵ procedemos a etiquetar los textos que constituyen nuestro corpus comparable con los movimientos y pasos.

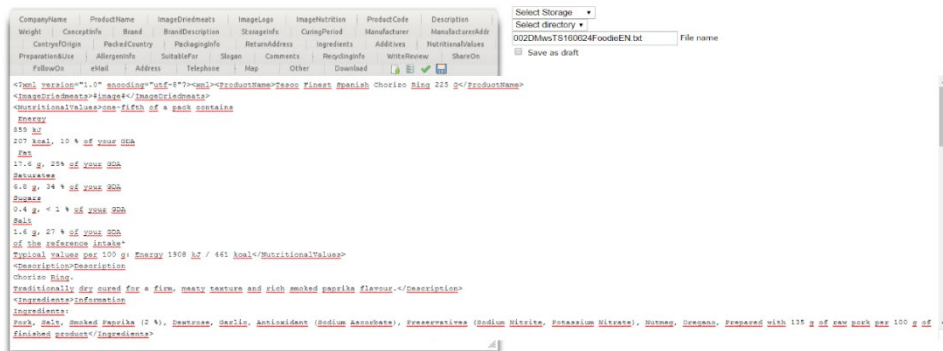


Figura 8. Captura de pantalla del Etiquetador de movimientos retóricos®.

De hecho, este programa está diseñado para anotar los movimientos retóricos en un corpus, así como para gestionar y almacenar los archivos una vez etiquetados. La ventana principal proporciona un menú con los géneros disponibles, se selecciona un directorio, un texto y, a continuación, aparecen las etiquetas que hacen referencia a los movimientos y a los pasos.

Por tanto, procedemos a anotar nuestro corpus con las etiquetas que hemos seleccionado a partir de nuestro conocimiento sobre el campo. Anotamos los dos subcorpus que constituyen C-GEFEM, en primer lugar, el corpus en español para, a continuación, anotar el corpus en inglés.

3.3. El visor de corpus comparables bilingües

Una vez que los textos están anotados con los movimientos y los pasos, estos se pueden comparar a través del visor de corpus comparables bilingües⁶, que incluye un menú que permite analizar y contrastar la información retórica, así como un analizador de concordancias. Además, las búsquedas se pueden limitar a un determinado movimiento o paso.



Figura 9. Interfaz del visor de corpus comparables.

Con esta herramienta procedemos a realizar un análisis de las etiquetas empleadas para verificar si la propuesta de etiquetas que se desprende de la observación de diez textos en cada lengua (ver Figura 7) se corresponde con los movimientos y los pasos del género de las fichas descriptivas de embutidos. Para llevar a cabo dicha verificación, en primer lugar, comprobamos el porcentaje de movimientos y pasos en cada uno de los subcorpus anotados, la ocurrencia de dichos movimientos y pasos, el porcentaje de textos en los que se incluyen dichos movimientos y pasos, así como el número de palabras total de cada movimiento o paso. El fin último es desarrollar un prototipo de estructura retórica formada por movimientos y pasos en cada una de las lenguas de trabajo: el español y el inglés.

4. Análisis y resultados

De las 39 etiquetas previstas para anotar los movimientos de las fichas descriptivas de embutido, 8 no se han utilizado durante la mencionada anotación de los textos en español y 3 no se han empleado en el etiquetado del subcorpus en inglés, puesto que en dichas fichas no había estructuras textuales que se correspondiesen con las mencionadas etiquetas. Dichas etiquetas se ofrecen la Tabla 1.

Tabla 1. Etiquetas no utilizadas.

Subcorpus en español	Subcorpus en inglés
<PackedCountry>	<Telephone>
<ReturnAddress>	<Map>
<Slogan>	<Download>
<Comments>	
<RecyclingInfo>	
<WriteReview>	
<FollowOn>	
<Map>	

En total, en español se utilizan 31 etiquetas, es decir, 31 movimientos y pasos, en tanto que en el subcorpus en lengua inglesa se han empleado 35 etiquetas, equivalentes a 35 movimientos y pasos. A continuación, procedemos a detallar la estructura retórica en cada una de las lenguas.

4.1. Estructura retórica de las fichas descriptivas de embutidos en español

En la Tabla 2 se muestra la incidencia de las etiquetas tras proceder a anotar el subcorpus en español de C-GEFEM.

Tabla 2. Etiquetas retóricas utilizadas en el subcorpus C-GEFEM en español.

Movimientos y pasos	Frecuencia	Nº veces	% de texto perteneciente a la etiqueta	Nº palabras
<ProductName>	100 %	115	5,06 %	685
<ImageDriedmeats>	96 %	96	0,71 %	96
<Description>	96 %	106	24,79 %	3354
<Ingredients>	90 %	93	15,49 %	2095
<StorageInfo>	85 %	90	6,36 %	860
<Weight>	84 %	87	3,23 %	437
<NutritionalValues>	66 %	66	16,24 %	2197
<Manufacturer>	52 %	52	3,23 %	437
<ManufacturerAddress>	51 %	51	4,02 %	544
<PackagingInfo>	46 %	46	3,81 %	516
<Brand>	41 %	42	0,87 %	118
<BrandDescription>	41 %	42	0,87 %	118
<Preparation&Use>	40 %	40	5,14 %	695
<CuringPeriod>	35 %	35	3,19 %	431
<CompanyName>	29 %	30	2,08 %	281
<ConceptInfo>	29 %	32	1,42 %	192
<ProductCode>	21 %	21	0,27 %	37
<AllergenInfo>	19 %	20	1,86 %	251
<Additives>	16 %	16	1,03 %	140
<CountryofOrigin>	7 %	7	0,26 %	35
<Other>	5 %	5	0,29 %	39
<Suitable for>	4 %	4	0,09 %	12
<eMail>	3 %	3	0,04 %	6
<Address>	3 %	3	0,10 %	14
<Telephone>	3 %	3	0,04 %	6
<ImageLogo>	2 %	2	0,09 %	12
<Download>	2 %	2	0,24 %	33
<ImageNutrition>	1 %	1	0,01 %	1
<ShareOn>	1 %	1	0,01 %	1

De los datos de la Tabla 2 se desprende que en todas las fichas descriptivas de embutidos aparece la denominación del producto (100%), en el 96% una imagen y la descripción del producto, así como los ingredientes (90%). Asimismo, en el 85% de las fichas se recoge información relativa a la conservación y utilización. En el 84% se indica el peso del producto. Asimismo, más de la mitad de las fichas (64%) incluyen información nutricional.

Tras observar con el visor de corpus comparables el orden de las distintas etiquetas, las fichas descriptivas de embutidos en español estarían compuestas por los siguientes movimientos y pasos, cuya frecuencia de uso la hemos representado con estrellas, siendo cinco estrellas (*****) el símbolo de obligatoriedad (81% - 100%), cuatro estrellas (****) una alta aparición (61% - 80%), tres estrellas (***) una frecuencia media (41% - 60%), dos estrellas (**) poca frecuencia (21% - 40%) y una estrella (*) una escasa aparición (1% - 20%):

1. DENOMINACIÓN DEL PRODUCTO (*****)
2. IMAGEN EMBUTIDO (*****)
3. MARCA (***)
 - 3.1. NOMBRE (***)
 - 3.2. LOGO (*)
 - 3.3. DESCRIPCIÓN (***)
4. INFORMACIÓN CONCEPTUAL (**)
5. CÓDIGO DE PRODUCTO (**)
6. DESCRIPCIÓN DEL PRODUCTO (*****)
7. INFORMACIÓN DEL PRODUCTO
 - 7.1. PESO (*****)
 - 7.2. INGREDIENTES (*****)
 - 7.3. ALÉRGENOS (**)
 - 7.4. INFORMACIÓN NUTRICIONAL (*****)
 - 7.5. CONSERVACIÓN (*****)
 - 7.6. UTILIZACIÓN (**)
8. ORIGEN (*)
9. ENVASADO (***)
10. OPERADOR (***)
 - 10.1. DIRECCIÓN (***)

Una vez extraídos por frecuencia y ordenados por la posición que ocupan en el texto los movimientos y los pasos en español, procedemos a repetir el proceso en lengua inglesa para detectar posibles asimetrías entre lenguas.

4.2. Estructura retórica de las fichas descriptivas de embutidos en inglés

Los resultados que se desprenden de la anotación retórica del corpus en lengua inglesa se ofrecen en la Tabla 3.

Tabla 3. Etiquetas retóricas utilizadas en el subcorpus C-GEFEM en inglés.

Movimientos y pasos	Frecuencia	Nº veces	% de texto perteneciente a la etiqueta	Nº de palabras
<ProductName>	98 %	119	2,78 %	620
<ImageDriedmeats>	87 %	87	0,39 %	87
<Description>	81 %	105	12,59 %	2805
<Ingredients>	78 %	81	10,77 %	2399
<Preparation&Use>	60 %	76	10,94 %	2437
<StorageInfo>	59 %	73	6,19 %	1380
<NutritionalValues>	59 %	81	16,42 %	3658
<ConceptInfo>	53 %	56	2,72 %	605
<AllergenInfo>	49 %	58	2,32 %	518
<Weight>	47 %	59	1,51 %	337
<PackagingInfo>	46 %	59	1,95 %	434
<CountryofOrigin>	46 %	53	2,12 %	473
<ManufacturerAddress>	36 %	36	2,02 %	449
<Brand>	27 %	29	0,30 %	66
<ProductCode>	27 %	29	0,19 %	42
<BrandDescription>	27 %	29	0,30 %	66
<WriteReview>	26 %	26	0,26 %	57
<ReturningAddress>	24 %	24	1,98 %	432
<Comments>	22 %	22	12,97 %	2890
<Other>	17 %	17	1,72 %	384
<Manufacturer>	13 %	13	0,51 %	113
<RecyclingInfo>	13 %	13	0,71 %	159
<Suitable for>	12 %	12	0,66 %	146
<Additives>	10 %	10	0,30 %	66
<PackedCountry>	10 %	10	0,24 %	53
<FollowOn>	9 %	9	0,24 %	54
<CompanyName>	8 %	8	0,09 %	19
<CuringPeriod>	7 %	7	0,48 %	108
<Address>	4 %	4	0,11 %	24
<ImageLogo>	3 %	3	0,05 %	11
<Slogan>	3 %	3	0,11 %	24
<eMail>	2 %	2	0,01 %	2
<ImageNutrition>	2 %	3	0,02 %	5
<ShareOn>	1 %	1	0,04 %	8

De la observación de los datos precedentes se desprende que en todas las fichas descriptivas de embutidos en lengua inglesa se incluye la denominación del producto (98%), una imagen (87%), la descripción del producto (81%), los ingredientes (78%), la utilización (60%), la conservación (59%) y el valor nutricional (59%).

Tras comprobar con el visor el orden de las etiquetas, las fichas descriptivas de embutidos en inglés estarían compuestas por los siguientes movimientos y pasos:

1. PRODUCT NAME (*****)
2. WEIGHT (***)
3. PRODUCT IMAGE (*****)
4. PRODUCT CODE (**)
5. CONCEPTUAL INFORMATION (***)
6. DESCRIPTION (*****)
7. INFORMATION
 - 7.1. BRAND (**)
 - 7.1.1. BRAND DESCRIPTION (**)
 - 7.2. STORAGE (***)
 - 7.3. ORIGIN (***)
 - 7.3.1. PACKED COUNTRY (*)
 - 7.4. PREPARATION AND USE (*****)
 - 7.5. PACKAGING INFO (***)
 - 7.6. RECYCLING (*)
 - 7.7. OTHER INFORMATION (**)
8. INGREDIENTS (*****)
 - 8.1. ADDITIVES (*)
 - 8.2. ALLERGENS (***)
 - 8.3. SUITABLE FOR (**)
9. NUTRITIONAL VALUES (***)
10. MANUFACTURER (*)
 - 10.1. MANUFACTURER ADDRESS (**)
11. RETURN TO ADDRESS (**)
12. REVIEW (**)
- 12.1. COMMENTS (**)

Una vez establecido el prototipo de estructura retórica de las fichas descriptivas de embutidos tanto en español como en lengua inglesa procedemos a examinar las diferencias retóricas de este género textual en dichas lenguas.

4.3. Comparativa de resultados

Las fichas descriptivas de embutidos en ambas lenguas contienen tres movimientos de obligada aparición: ‘denominación del producto’, ‘imagen del producto’ y ‘descripción del producto’. De hecho, tanto en español como en inglés las fichas comienzan presentando la denominación del producto. A continuación, en español aparece ‘imagen embutido’, en tanto que en lengua inglesa suele recogerse el movimiento ‘peso’ entre ‘denominación del producto’ e ‘imagen embutido’. En tercer lugar, en español suele presentarse la ‘marca’ con tres pasos: ‘nombre’, ‘logo’ y ‘descripción’, mientras que en lengua inglesa esta información se ofrece como paso dentro del movimiento ‘información del producto’, que ocupa el séptimo lugar.

Tras ‘marca’, en español suele describirse ‘información conceptual’, que también se recoge en lengua inglesa, pero en este caso le antecede ‘código de producto’. Este movimiento se presenta en español tras ‘información conceptual’.

Por otro lado, el movimiento relativo a ‘descripción del producto’ ocupa el sexto lugar y, a continuación, se incluye en ambas lenguas el movimiento ‘información del producto’. Dicho movimiento se compone de seis pasos en español: ‘peso’, ‘ingredientes’, ‘alérgenos’, ‘información nutricional’, ‘conservación’ y ‘utilización’. No obstante, en lengua inglesa está formado por siete pasos: ‘marca’, ‘conservación’, ‘origen’, ‘utilización’, ‘envasado’, ‘reciclaje’ y ‘otra información’. De hecho, en este movimiento detectamos ostensibles diferencias entre lenguas, puesto que solo coinciden los pasos relativos a ‘conservación’ y a ‘utilización’. Asimismo, las frecuencias de uso son dispares, por ejemplo, ‘conservación’ se emplea con una frecuencia de 85% en español y de 59% en inglés; ‘utilización’ con 40% y 60%, respectivamente.

Por otra parte, ‘ingredientes’ es un movimiento en las fichas descriptivas de embutidos en inglés, en tanto que en español es un paso recogido dentro del movimiento ‘información del producto’. Dicho movimiento en lengua inglesa se compone de tres pasos: ‘aditivos’, ‘alérgenos’ y ‘adecuado para’. Además, ‘información nutricional’ se corresponde con otro movimiento independiente en lengua inglesa, mientras que en español se incluye como paso de ‘información de producto’. Sin embargo, en español ‘origen’ y ‘envasado’ constituyen dos movimientos independientes, en tanto que en lengua inglesa son dos pasos que forman parte de ‘información de producto’.

Asimismo, en español se describe ‘operador’, que también se incluye en la estructura en inglés. Además, en esta última lengua se describen otros dos movimientos: ‘dirección de devolución’ y ‘reseña’, que no están reflejados en la estructura retórica en español.

Por último, detectamos que siete movimientos en lengua inglesa que no aparecen en las fichas descriptivas de embutidos en español, tales como ‘reseña’ (26%), ‘dirección de devolución’ (24%), ‘comentarios’ (22%), ‘información sobre reciclaje’ (13%), ‘país de envasado’ (10%), ‘Síguenos en las redes sociales’ (9%) y ‘eslogan’ (3%).

CONCLUSIONES

En primer lugar, constatamos que, si las empresas cárnicas españolas se limitan a traducir el contenido del español al inglés, las fichas descriptivas de embutidos no funcionarán en el contexto comunicativo anglosajón porque la estructura retórica de este género no es idéntica en ambas lenguas, tal y como ocurre en otros géneros en el par de lenguas inglés-español (Valero-Garcés, 1996; Labrador et al., 2014).

Además, los datos expuestos proporcionan a los usuarios patrones retóricos que pueden ayudarles a la hora de redactar fichas descriptivas de embutidos en lengua inglesa, puesto que como señala Bhatia (2004: 145), “generic competence is an important contributor to professional expertise”. Asimismo, dichos datos han servido para desarrollar un generador de fichas descriptivas de embutidos (GEFEM)⁷, que asistirá a los traductores y redactores españoles a elaborar fichas descriptivas de embutidos en lengua inglesa. No obstante, convendría abordar en futuros trabajos los patrones lexicogramaticales empleados en cada uno de los movimientos y pasos en las lenguas española e inglesa.

Para finalizar, con los resultados obtenidos, consideramos necesario sensibilizar y formar a los traductores y redactores en las particularidades que tienen los distintos géneros en las lenguas española e inglesa para lograr que la comunicación especializada en este ámbito sea lo más completa y precisa posible. Dado que la estructura retórica varía considerablemente entre lenguas, los traductores y redactores que trabajan en la industria cárnica tendrán en cuenta estas variaciones de la estructura retórica y se verán obligados a adaptar las fichas descriptivas de embutidos a la lengua de llegada durante el trasvase interlingüístico de la información.

REFERENCIAS BIBLIOGRÁFICAS

- Austermühl, F. (2001). *Electronic tools for translators*. Manchester: St. Jerome.
- Bhatia, V. K. (1993). *Analysing genre: Language use in professional settings*. Londres: Longman.
- Bhatia, V. K. (2004). *Words in written discourse*. Londres: Continuum.
- Biber, D., Connor, U. & Upton, T. A. (Eds.) (2007). *Discourse on the move. Using corpus analysis to describe discourse structure*. Ámsterdam: John Benjamins. DOI:10.1075/scl.28
- Bowker, L. (2002). *Computer-aided translation technology: A practical introduction*. Ottawa: University of Ottawa Press.
- Chesterman, A. (1998). *Contrastive functional analysis*. Ámsterdam/Philadelphia: John Benjamins.
- Corpas Pastor, G. (2008). *Investigar con corpus en traducción: Los retos de un nuevo paradigma*. Frankfurt am Main: Peter Lang.
- Corpas Pastor, G. & Seghiri, M. (2007a). Specialized corpora for translators: A quantitative method to determine representativeness. *Translation Journal*, 11(3) [en línea]. Disponible en: <http://translationjournal.net/journal/41corpus.htm>

- Corpas Pastor, G. & Seghiri, M. (2007b). Determinación del umbral de representatividad de un corpus mediante el algoritmo N- Cor. *SEPLN: Revista de la Sociedad Española para el Procesamiento del Lenguaje Natural*, 39, 165-172.
- Corpas Pastor, G. & Seghiri, M. (2009). Virtual corpora as documentation resources: Translating travel insurance documents (English-Spanish). En A. Beeby, P. Rodríguez Inés & P. Sánchez-Gijón (Eds.), *Corpus Use and Translating* (pp. 75-107). Amsterdam/Philadelphia: John Benjamins.
- Corpas Pastor, G. & Seghiri, M. (2010). Size matters: A quantitative approach to corpus representativeness. En R. Rabadán (Ed.), *Lengua, traducción, recepción. En honor de Julio César Santoyo / Language, translation, reception. To honor Julio César Santoyo* (pp. 112-146). León: Universidad de León, Área de Publicaciones.
- Cristobalena Frutos, A. (2016). *Análisis contrastivo inglés-español de los manuales de instrucciones de electrodomésticos / English-Spanish contrastive analysis of instruction manuals for household appliances*. Tesis doctoral, Universidad de León, León, España.
- EAGLES (1996). *Preliminary recommendations on corpus typology*. Documento técnico EAGLES EAG-TCWG-CTYP/P [en línea]. Disponible en: <http://www.ilc.cnr.it/EAGLES/corpus/typ/corpus.html>
- Egins, S. & Martin, J. R. (2000). Géneros y registros del discurso. En T. A. van Dijk (Coord.), *El discurso como estructura y proceso* (pp. 335-372). Barcelona: Gedisa.
- Egins, S. & Martin, J. R. (2003). El contexto como género: Una perspectiva lingüística funcional. *Revista Signos. Estudios de Lingüística*, 36(54), 185-205. DOI: 10.4067/S0718-09342003005400005
- García Izquierdo, I. (2002). El género: Plataforma de confluencia de nociones fundamentales en didáctica de la traducción. *Discurso, Série Estudos de Tradução*, 2, 13-21.
- Heinemann, W. (2000). Textsorten. Zur diskussion um basisklassen des kommunizierens. Rückschau und ausblick. En K. Adamzik (Ed.), *Textsorten: Reflexionen und Analysen* (pp. 9-29). Tübinga: Stauffenburg.
- Kwan, B. S. C. (2006). The schematic structure of literature review in doctoral theses of applied linguistics. *English for Specific Purposes*, 25(1), 30-55. DOI: 10.1016/j.esp.2005.06.001Get
- Labrador, B., Ramón, N., Alaiz-Moretón, H. & Sanjurjo-González, H. (2014). Rhetorical structure and persuasive language in the subgenre of online advertisements. *English for Specific Purposes*, 34, 38-47.

- López Arroyo, B. & Roberts, R. P. (2014). English and Spanish descriptors in wine tasting terminology. *Terminology*, 20(1), 25-49. DOI: 10.1075/term.20.1.02lop
- López-Arroyo, B., Fernández-Antolín, M. & de Felipe-Boto, R. (2007). Contrasting the rhetoric of abstracts in medical discourse. Implications and applications for English-Spanish translations. *Languages in Contrast*, 7(1), 1-28. DOI:10.1075/lic.7.1.02lop
- MAPAMA (2017). *Marco Estratégico para la Industria de Alimentación y Bebidas*. Madrid: MAPAMA [en línea]. Disponible en: <http://www.mapama.gob.es/es/alimentacion/temas/industria-agroalimentaria/marco-estrategico/>.
- Martin, J. R. (1997). Analysing genre: Functional parameters. En C. Frances & J. R. Martin (Eds.), *Genre and institutions: Social processes in the workplace and school* (pp. 3-39). Londres/ Nueva York: Continuum.
- Martínez Hincapié, J. D. (2015). El género Tesis Doctoral de Historia y Física: Descripción y variación retórico-funcional. En G. Parodi & G. Burdiles (Eds.), *Leer y escribir en contextos académicos y profesionales: Géneros, Corpus y Métodos* (pp. 113-154). Santiago de Chile: Ariel.
- Martín-Martín, P. (2005). *The rhetoric of the abstract in English and Spanish scientific discourse*. Berna: Peter Lang.
- McEnery, T. & Hardie, A. (2012). *Corpus linguistics: Method, theory and practice*. Ámsterdam/Philadelphia: John Benjamins.
- Méndez-Cendón, B. (2009). Combinatorial patterns in medical case reports: An English–Spanish contrastive analysis. *The Journal of Specialised Translation*, 11 [en línea]. Disponible en: http://www.jostrans.org/issue11/art_mendez.php
- Miles, M. B., Huberman, A. M. & Saldana, J. (2014). *Qualitative data analysis. A Methods Sourcebook*. Thousand Oaks: SAGE.
- Navarro, F. (2015). Análisis situado del plan de negocios en español y portugués: Perspectivas de emprendedores, docentes y estudiantes. *Calidoscópico*, 13(2), 189-200. DOI: 10.4013/cld.2015.132.05
- 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*. Berna: Peter Lang.

- Ortego-Antón, M. T. & Fernández Nistal, P. (2019). Estudio contrastivo de la terminología de embutidos en inglés y en español con ParaConc y tIConc a partir del corpus paralelo P-GEFEM y del comparable C-GEFEM. En M. Seghiri (Ed.), *El uso de los corpus lingüísticos como herramienta pedagógica para la enseñanza-aprendizaje de lenguas, traducción e interpretación* (pp. 23-48). Berna: Peter Lang.
- Parodi, G. (Ed.) (2015). *Géneros académicos y géneros profesionales. Accesos discursivos para saber y hacer*. Valparaíso: Ediciones Universitarias de Valparaíso.
- Parodi, G. & Burdiles, G. (Eds.) (2015). *Leer y escribir en contextos académicos y profesionales: Géneros, Corpus y Métodos*. Santiago de Chile: Ariel.
- Rabadán, R. (2016). Proposals in meeting minutes. An English-Spanish corpus-based study. *Languages in Contrast*, 16(2), 213-238. DOI: 10.1075/lic.16.2.03rab
- Sánchez Trigo, E. (2005). Investigación traductológica en la traducción científica y técnica. *TRANS: Revista de Traductología*, 9, 131-150.
- Seghiri, M. (2006). *Compilación de un corpus trilingüe de seguros turísticos (español-inglés-italiano): Aspectos de evaluación, catalogación, diseño y representatividad*. Tesis doctoral, Universidad de Málaga, España [en línea]. Disponible en: <http://hdl.handle.net/10630/2715>.
- Seghiri, M. (2015). Determinación de la representatividad cuantitativa de un corpus ad hoc bilingüe (inglés-español) de manuales de instrucciones generales de lectores electrónicos/Establishing the quantitative representativeness of an E-Reader User's Guide *ad hoc* corpus (English-Spanish). En M. T. Sánchez Nieto (Ed.), *Corpus-Based Translation and Interpreting Studies. From Description to Application* (pp. 125-146). Berlín: Frank & Timme.
- 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
- Sinclair, J. (2005). Corpus and text-basic principles. En M. Wynne (Ed.), *Developing Linguistic Corpora: A Guide to Good Practice* (pp. 1-16). Oxford: Oxbow Books.
- Swales, J. M. (1990/2001). *English in academic and research settings*. Cambridge: Cambridge University Press.
- Swales, J. M. (2004). *Research genres. Explorations and applications*. Cambridge: Cambridge University Press.

- Valero-Garcés, C. (1996). Contrastive ESP rhetoric: Metatex in Spanish-English economics texts. *English for Specific Purposes*, 15(4), 279-294.
- Venegas, R., Zamora, S. & Galdames, A. (2016). Hacia un modelo retórico-discursivo del macrogénero Trabajo Final de Grado en Licenciatura. *Revista Signos. Estudios de Lingüística*, 49(51), 247-279. DOI: 10.4067/S0718-09342016000400012
- Williams, I. A. (2010). Cultural differences in Academic Discourse. Evidence from first-person verb use in the methods sections of medical research articles. *Special Issue of International Journal of Corpus Linguistics*, 15(2), 214-239.
- Zannettin, F. (2012). *Translation-driven corpora. Corpus resources for descriptive and applied translation studies*. Manchester/Kinderhook (NY): St. Jerome Publishing.

NOTAS

- ¹ El presente trabajo ha sido realizado en el marco del proyecto titulado “Producción textual bilingüe semiautomática inglés-español con lenguajes controlados: parametrización del conocimiento experto para su desarrollo en aplicaciones web 2.0. y 3.0.” (REF. FFI2016-75672-R).
- ² ACTRES es un grupo de investigación interuniversitario coordinado por la Dra. Rosa Rabadán (Universidad de León) [en línea]. Disponible en: <http://actres.unileon.es/?lang=es>
- ³ Este tipo de corpus también ha recibido las denominaciones de “precision corpus”, “do-it-yourself corpus” y “ad hoc corpus” (Corpas Pastor, 2008: 91).
- ⁴ El programa ReCor fue diseñado por Seghiri (2006) y Copas Pastor y Seghiri (2007a, 2007b, 2009, 2010) y se encuentra patentado a través de la Oficina Española de Patentes y Marcas [en línea]. Disponible en: <http://patentados.com/invento/nuevo-metodo-determinar-representatividad-corpus.html> (Fecha de consulta: 12/04/2019).
- ⁵ [en línea]. Disponible en: <http://contraste2.unileon.es/web/es/tagger.html> (Fecha de consulta: 12/04/2019).
- ⁶ [en línea]. Disponible en: <http://contraste2.unileon.es/web/es/browser.html> (Fecha de consulta: 12/04/2019).
- ⁷ Una demo de GEFEM [en línea]. Disponible en: <http://contraste.unileon.es/appChacinas/#/> (Fecha de consulta: 12/04/2019).