



Universidad
de Alcalá

**LAS DIFERENCIAS ENTRE TEXTOS ORIGINALES Y
TEXTOS TRADUCIDOS SOBRE CÁNCER
COLORRECTAL: ANÁLISIS DE CORPUS**

**DIFFERENCES BETWEEN ORIGINAL AND
TRANSLATED TEXTS ON COLORECTAL CANCER:
ANALYSIS OF A CORPUS**

2019-2020

**Máster Universitario en Comunicación Intercultural,
Interpretación y Traducción en los Servicios Públicos**

Presentado por:

D./D^a PAULA MACÍA CARAYOL

Dirigido por:

Dr./Dra. D./D^a RAQUEL LÁZARO GUTIÉRREZ

Alcalá de Henares, a 20 de agosto de 2020

Índice

0. Resumen / Abstract.....	4
1. Introducción.....	5
2. Objetivos.....	6
3. Motivación.....	6
4. Hipótesis.....	7
5. Metodología.....	8
6. Marco teórico.....	10
6.1. Corpus lingüístico.....	10
6.1.1. Definición.....	10
6.1.2. Criterios generales para el diseño de un corpus.....	11
6.1.3. Representatividad.....	12
6.1.4. Otros problemas para el diseño de un corpus.....	14
6.2. Traducción médica.....	16
6.2.1. Género médico.....	18
6.2.1.1. Características de los <i>Fact Sheet for Patients</i>	20
6.2.2. Problemas en la traducción de textos médicos	22
6.2.2.1. Terminología.....	22
6.2.2.2. Otros problemas.....	23
6.3. Traducción de páginas web.....	25
6.3.1. Características de las páginas web.....	27
6.3.2. Proceso de traducción de una página web	28
6.3.3. Problemas en localización.....	30
7. Investigación.....	32
7.1. Introducción.....	32
7.2. Descripción del corpus.....	33
7.3. Modelo de análisis.....	34
7.4. AntConc.....	34
7.5. Análisis y resultados.....	35
7.5.1. Preparación de la aplicación AntConc.....	35

7.5.2. Gerundio.....	36
7.5.3. Voz Pasiva.....	40
7.5.4. Terminología	45
7.5.5. Siglas.....	49
7.5.6. Nombres de medicamentos.....	52
7.5.7. Tratamiento del lector.....	56
8. Conclusiones.....	60
9. Traducción.....	68
10. Bibliografía.....	88
10.1. Fuentes consultadas y utilizadas.....	88
10.2. Fuentes de los textos que conforman el corpus.....	92
10.2.1. Subcorpus de textos originales.....	92
10.2.2. Subcorpus de textos traducidos.....	94

0. Resumen / Abstract

Este estudio tiene como objetivo comparar textos redactados originalmente en español con textos traducidos del inglés a esta misma lengua. Esto se lleva a cabo a partir de la creación de un corpus comparable de alrededor de 70.000 palabras formado específicamente para este trabajo. El corpus lo conforman textos informativos extraídos de páginas web de entidades sanitarias que tratan el cáncer colorrectal. Aunque existe una gran variedad de elementos a los que hay que prestar especial atención en los textos médicos y sus traducciones, esta investigación se centra solamente en algunos de ellos como son el gerundio, la voz pasiva, el grado de especialización de la terminología empleada, las siglas, los nombres de medicamentos y la forma de dirigirse al lector. De este modo, se comprueba que ambos tipos de textos presentan una calidad de redacción bastante elevada en general, pero que existen ciertos aspectos que presentan los textos traducidos que pueden considerarse más adecuados para el tipo de texto que se está tratando.

Palabras clave: corpus comparable, traducción médico-sanitaria, cáncer colorrectal, gerundio, voz pasiva.

This essay aims to compare texts that have been originally written in Spanish with texts that have been translated from English into Spanish. This is carried out through the creation of a comparable corpus of around 70,000 words that has been created specifically for this study. This corpus includes informative texts whose main subject is colorectal cancer. They were compiled from different healthcare websites. Although there is a wide variety of elements to which pay special attention in medical texts and their translations, this essay only focuses on some of them such as gerunds, passive voice, the level of specialisation of terminology, acronyms, medicine's names and the way of addressing the reader. Thus, it is proven that both types of texts show drafting quality. However, decisions made in translated texts are more suitable for this type of texts.

Key words: comparable corpus, medical translation, colorectal cancer, gerunds, passive voice.

1. Introducción

La aparición de Internet ha supuesto un antes y un después en la historia de la humanidad. Es cierto que antes de que cualquier persona pudiera acceder a la red, existían otras formas de comunicación: el teléfono, el correo ordinario, la radio o la televisión. No obstante, la inmediatez que ofrece Internet no puede compararse a ningún otro invento anterior. Demostrarlo es tan sencillo como comparar la pérdida de usuarios que reciben cada vez más el resto de dispositivos o servicios con la cantidad de niños e incluso personas ya entradas en edad que quieren hacer uso de un móvil con 4G. Si bien es cierto que se sigue escuchando la radio, se ven las noticias o ciertos programas del corazón en la televisión o se hace uso del correo para ciertas cuestiones, no puede compararse a la cantidad de series que se ven a través de los diferentes servicios de distribución de contenido audiovisual como los que ofrecen Netflix, HBO o Amazon Prime, los mensajes que se mandan a través de aplicaciones como WhatsApp o Instagram, las escuchas que reciben las canciones en Spotify o el número de veces que se introduce cualquier información en la barra de búsqueda de Google para conocer cualquier noticia. Y no es solo que Internet se encuentre actualmente en lo más alto de la pirámide de la comunicación, es que también es la base de muchos trabajos, como aquellos que requieren un GPS, búsqueda instantánea de información o uso de programas conectados a la red. En resumen, Internet se ha ganado un puesto muy importante en las vidas de las personas, ¿pero en qué puede influir este en la medicina o, más concretamente, en el cáncer colorrectal?

Como se comentaba anteriormente, Internet sirve para comunicar y comunicarse, en muchas ocasiones, consiste en informar. Cuando se utiliza la palabra “informar” no solo se hace referencia a lo que ocurre alrededor del mundo. Existen tantas ciencias y disciplinas actualmente que es imposible que una sola persona tenga conocimiento de todo lo que implica cada una de ellas. Internet facilita el almacenamiento de la información sobre cualquier tema. Aunque el hecho de que cualquiera tenga acceso a esta red y pueda agregar cualquier información sin ningún tipo de filtro reduce su grado de fiabilidad. No obstante, el conocimiento actual sobre la red es tan amplio que la mayoría de usuarios ya hacen uso de ciertas estrategias para reconocer cuándo se puede o no confiar en una página web.

Si Internet es una de las fuentes de información más completas, asequibles y consultadas, cualquier empresa, entidad u organización querrá formar parte de esta red, ya sea a través de anuncios en otras páginas web o a partir de la creación de la suya propia. Aunque pueda parecer que esto es más una táctica de marketing mediante la cual las empresas promocionan sus productos o servicios, algunas entidades no pretenden recibir un beneficio monetario a cambio. Y es ahí cuando entra la temática que se va a tratar en este estudio: el cáncer colorrectal.

El cáncer es una de las enfermedades con más afectados, ya sea por los diferentes tipos que la componen o porque no se ha encontrado todavía ninguna cura definitiva para ninguno de ellos. El hecho de que sea tan agresivo ya que le envuelve un elevado índice de mortalidad, conlleva dos cosas: por un lado, que cualquier persona perteneciente a un país desarrollado la conozca y la tema; y, por otro lado, que las entidades sanitarias quieran informar y formar al público general sobre esta enfermedad para poder prevenirla en el mayor número de casos.

Actualmente, hospitales, organizaciones sanitarias y aseguradoras de todo el mundo ofrecen información sobre ciertas enfermedades y condiciones con el objetivo de informar a los posibles pacientes y ayudarles a actuar o a prevenir. A pesar de que los profesionales que trabajan para estas empresas tienen unos elevados conocimientos sobre todas estas nociones, es necesario que adapten el lenguaje a un lego. Para ello, esta información se agrupa en un género textual caracterizado por su vocabulario asequible, su funcionalidad y su simplicidad. No obstante, a veces los trabajadores sanitarios encargados de esta redacción se topan con grandes dificultades a la hora de resumir una gran cantidad de conocimientos que han ido adquiriendo durante toda su carrera profesional a un mensaje básico y reducido, cuando probablemente nunca hayan recibido formación en redacción. Por otro lado, los profesionales de la traducción también se enfrentan a ciertos problemas terminológicos o a una toma de decisiones que afectará a la recepción del texto por parte de un público con otra lengua meta que, por su cultura u otras razones, puede implicar la adaptación de ciertos aspectos.

2. Objetivos

El objetivo principal de mi Trabajo de Fin de Máster es descubrir las diferencias, en caso de que las haya, que se encuentran entre textos redactados originalmente en español peninsular y textos traducidos a este mismo idioma. Estas traducciones provendrán siempre del inglés, ya que combinar varias lenguas de origen podría alterar los resultados de la investigación. Para que el estudio sea mucho más específico, he decidido centrarme en una tipología textual y una temática especializada, estudiando, por ende, textos médicos destinados a un público no especializado sobre cáncer colorrectal. Para poder lograr mi objetivo, es necesario recolectar un gran número de textos de ambos tipos y realizar un análisis exhaustivo de cada uno de ellos, por lo que he creído que utilizar un corpus de textos traducidos y de textos originales podría ser una metodología apropiada. Por tanto, tengo un segundo objetivo, que se llevará a cabo antes que el objetivo principal, que consiste en ser capaz de crear un corpus comparable siguiendo unos criterios básicos para que los resultados de este estudio no se vean contaminados.

3. Motivación

Desde que los lingüistas empezaron a prestar especial atención a la traducción y a considerarla una disciplina digna de estudio, se han asignado distintos campos para cada una de sus ciencias. Uno de los más amplios y extensos es la traducción especializada y, dentro de ella, se pueden encontrar una gran variedad de subcampos que, en ocasiones, no se encuentran del todo delimitados. Es bastante obvio que los textos cuyo tema principal es la medicina contienen lenguaje especializado. No obstante, esta afirmación no implica que dichos documentos estén solamente destinados a un público previamente formado en esta ciencia. De hecho, desde que la red de Internet se volvió alcanzable para cualquier persona, este tipo de textos se ha visto cada vez más adaptado a un público general. Esto se debe a que los especialistas en medicina han observado que formar a una persona de a pie sobre esta disciplina puede prevenir ciertas enfermedades o condiciones. Es por esto que, actualmente, es muy habitual encontrar textos que contengan datos

generales sobre la gran mayoría de enfermedades con un vocabulario y discurso bastante accesibles. El problema que reside en esta premisa es que las organizaciones y entidades extranjeras dedican bastante tiempo a la elaboración de textos sencillos que formen e informen a un público cada vez más amplio, pero no muestran la misma preocupación a la hora de realizar sus traducciones. Por esta razón, comparar textos que se han trabajado únicamente y desde un primer momento en español con otros que se han elaborado en otro idioma y se han traducido posteriormente demostrará si esta diferencia en la redacción de textos supone un verdadero problema como podría parecer en un principio.

Por el lado de la traducción sanitaria, esta investigación no debería suponer ningún tipo de problema puesto que, al ser la medicina uno de los campos de conocimiento más antiguos del ser humano, se trata de una de las ramas más estudiadas en traducción. Sin embargo, la lingüística de corpus y las herramientas a la disposición de los investigadores para su análisis son dos terrenos mucho más recientes y que, por ende, se han estudiado en menor medida. Si bien es cierto que existe una gran cantidad de programas informáticos que ayudan en el manejo de corpus, para la realización de listas de sus formas (*types*), la adición de etiquetas o la proporción de líneas de concordancia, entre muchas otras tareas, el mayor problema reside en los criterios que deben seguirse para la realización de un corpus, especialmente a la hora de establecer un umbral de representatividad. El hecho de que esto pueda parecer un obstáculo a la hora de realizar esta investigación se convierte en una motivación más para seguir con el proyecto.

Otro de las razones que me lleva a haber escogido este tipo de estudio es el hecho de que engloba uno de los objetivos principales de una buena traducción: provocar en el lector la sensación de que esté leyendo una producción original y no una traducción. Cualquier traductor que se precie tratará siempre de realizar una obra que contenga el lenguaje más natural posible en la lengua de destino, llegando a realizar los cambios necesarios del texto original para que la lectura de su traslación sea lo más fluida posible. Es por esto que creo que un estudio que trate de analizar cuáles son los “fallos” que impiden que esto pueda tener lugar en un ámbito especializado como es la traducción biosanitaria dentro de un espacio reducido, pero a su mismo tiempo bastante consultado como es una página web, me parece de lo más interesante y necesario para conseguir unas traducciones cada vez más adecuadas y realistas.

4. Hipótesis

Como se mencionaba en uno de los apartados anteriores, este trabajo tiene dos objetivos. El primero de ellos consiste en realizar un corpus comparable. Esto implica la recopilación de un primer grupo de textos informativos sobre el cáncer colorrectal redactados originalmente en español y un segundo grupo de traducciones del inglés al español sobre esta misma tipología textual y temática, teniendo en cuenta que ambos grupos deben haber seguido unas directrices específicas de modo que puedan considerarse adecuados para esta investigación. El segundo objetivo consiste en hallar las diferencias que puede haber entre los textos traducidos y los textos originales y plantear si alguno de ellos presenta características generales más adecuadas para este tipo de textos.

Mi hipótesis inicial consiste en que el grupo de textos redactados en español presentará un discurso mucho más natural y consistente, así como un vocabulario que presta mayor atención a los matices. Uno de los motivos en los que he fundado mi hipótesis es que, a la hora de ir recolectando muestras, se podía comprobar que los autores de estos textos eran profesionales de la medicina. No obstante, en el caso del grupo restante, era prácticamente imposible comprobar la autoría de las traducciones, lo que podría significar que no se hubieran llevado a cabo por traductores profesionales o por traductores nativos de la lengua de destino. Además, el coste de una traducción aumentará en mayor o menor medida dependiendo del número de palabras que contenga el texto, las lenguas de origen y de destino y el nivel de especialidad. Si bien es cierto que el inglés y el español son dos de las lenguas con mayor número de hablantes y el nivel de especialidad no es todo lo elevado que podría llegar a ser, hay que tener en cuenta que la gran mayoría de textos aparecen en páginas web, lo que supone un precio mayor a la entidad al tratarse, no solo de traducción, sino de localización. Esto podría considerarse una razón para que la entidad de la página web encargue su traducción a una persona que no tenga conocimientos sobre este incremento del precio para ahorrar en costes. Finalmente, se puede exponer una razón más por la que los textos originales en español podrían presentar una mayor consistencia: el grado de conocimiento del autor. Aunque el traductor realizará una labor previa de análisis del texto y de documentación, si no es del todo cauto, podría cometer pequeños errores a la hora de utilizar ciertos vocablos o expresiones, emplear algún calco o no comprender ciertas ideas y representarlas de forma equívoca. De todos modos, estas son simples suposiciones y se ha realizado esta investigación para comprobar si esta hipótesis es cierta de forma total, parcial o es totalmente errónea.

5. Metodología

Para la realización de este estudio, he decidido utilizar una metodología mixta basada en el uso de corpus, más concretamente de un corpus comparable compuesto por textos originales en español y textos traducidos a este mismo idioma, para comprobar si existe alguna diferencia entre un tipo de texto específico redactado en una lengua y este mismo tipo de texto traducido a esa misma lengua. Para ello, ha sido primordial la realización propia de este corpus debido a que existe tal número de clases de corpus, tipologías textuales y temáticas que iba a ser prácticamente imposible encontrar un gran número de textos agrupados previamente por otro autor que siguiera lo mismos criterios que voy a necesitar para mi estudio.

Aunque hablaré de las distintas tipologías y clasificaciones de un corpus más adelante, quería destacar algunas de las características que presenta el que se va a utilizar en esta investigación para explicar cómo ha sido el proceso de recopilación de textos. Tras leer varios estudios con un diseño basado en el uso de corpus, comprobé que era usual utilizar corpus virtuales, es decir, agrupaciones de textos encontrados en la red. Como me pareció que Internet seguramente contendría una mayor selección de textos que cualquier otro recurso físico y su recopilación supondría un procedimiento mucho más sencillo que tener que escanear o copiar cualquier documento en papel, decidí realizar este corpus a partir de la compilación de textos encontrados en Internet. Tras tomar esta decisión, tenía bastante claro que, entre las distintas especialidades que habíamos estudiado en el máster,

la sanitaria había sido la que más había llamado mi atención. Por esto mismo, decidí investigar sobre las distintas tipologías de documentos con los que lidian los doctores, enfermeros y otros profesionales sanitarios, y empecé a explorar páginas web de entidades que estudian la medicina, dando con una modalidad bastante interesante destinada al público general: los *Fact Sheet for Patients*. Esta tipología se puede encontrar en la página web de cualquier entidad u organización con objetivos médicos, aseguradoras, hospitales o clínicas, entre otros y, por tanto, creo que es importante dedicarles su espacio ya que puede que sean uno de los textos médicos con mayor facilidad de acceso y, por tanto, más consultados por parte del público general.

Como decía anteriormente, antes de comenzar con este trabajo, tenía bastante claro que quería que mi estudio se centrara en textos médico-sanitarios ya que creo que es una temática a la que se le da bastante uso en la red y que ofrece una gran variedad de opciones. Es por esto que, tras investigar un poco sobre los distintos conceptos que la envuelven, como condiciones, enfermedades, entidades, tratamientos o medicamentos, entre muchos otros, escogí el cáncer colorrectal como tema principal de mis textos. La razón por la que quise tratar esta enfermedad es que el cáncer es uno de los términos médicos a los que la sociedad está más habituada puesto que se trata de una de las causas principales de morbilidad del mundo (Sociedad Española de Oncología Médica, 2018). Entre los más de 200 tipos de enfermedades que engloba (Asociación Española Contra el Cáncer, 2018), me decanté por el cáncer colorrectal, incluyendo su prevención, causas, síntomas, diagnóstico, fases y tratamiento, entre otras cuestiones, ya que se trata de uno de los tipos de tumores más diagnosticados en España en el año 2017 tanto en hombres como en mujeres (Sociedad Española de Oncología Médica, 2018). No obstante, es una enfermedad que puede curarse en más del 90% de los casos si se detecta a tiempo, lo que ha conducido a las distintas comunidades españolas a llevar a cabo programas de detección precoz para reducir esta incidencia y a elaborar textos con información básica sobre esta enfermedad para concienciar a la población. Estas razones me aseguraban la posibilidad de encontrar suficientes textos redactados originalmente en español, de calidad y con la temática y función que estaba buscando. Además, el cáncer colorrectal es también una de las principales causas de morbilidad a nivel mundial, lo que me garantizaba que hubiera textos que cumplieran con las mismas características tanto que hubieran sido redactados en lengua inglesa como traducidos al español.

Una vez decidido el tipo de estudio y material que iba a utilizar, comencé con la recopilación de textos para mi corpus, que se dividiría en dos subcorpus, uno formado por textos redactados originalmente en español y otro por textos traducidos a esta misma lengua. Decidí empezar por la compilación de textos originales porque, a pesar de que el español es una de las lenguas más habladas del mundo, no me servía cualquier variedad. Mi objetivo era encontrar textos en español peninsular para reducir el tipo de muestras y obtener así resultados mucho más específicos. Esto implicaba basarme solamente en fuentes procedentes de entidades españolas y que presentasen cierto grado de fiabilidad, lo que restringía el número de muestras disponibles. Por tanto, empezar a recopilar textos redactados en inglés que hubieran sido traducidos a una variedad neutra del español para luego comprobar que el número de muestras redactadas directamente al español peninsular era mucho menor iba a ser una pérdida de tiempo. Una vez recopilados todos los escritos posibles sobre cáncer colorrectal producidos por entidades sanitarias españolas, me centré en la búsqueda de textos traducidos hasta conseguir un número de

palabras lo más cercano posible al del subcorpus anterior. De este modo, se creó un corpus dividido en dos subcorpus compuestos por alrededor de 70.000 palabras cada uno.

Para poder comenzar con el análisis de ambos subcorpus, necesitaba un programa informático o aplicación que me ayudara a examinar todos los textos de forma simultánea. Buscar ciertos elementos en un grupo de 83 textos y posteriormente en uno de 67 de forma individual no solo sería una tarea interminable, sino que probablemente impediría identificar todos y cada uno de los ejemplos, un hecho que condicionaría los resultados del estudio. Para ello, opté por AntConc, un programa gratuito y de fácil acceso con el que ya había trabajado anteriormente en varias ocasiones y que había demostrado ser de gran utilidad. Antes de empezar con el análisis, se escogieron diferentes aspectos que podían resultar problemáticos en los textos que componen el corpus. Estos fueron adquiridos de las nociones teóricas consultadas previamente al inicio de este estudio basadas en investigaciones que tratan tanto la redacción de textos médicos, especialmente del género escogido, los *Fact Sheet for Patients*, y de páginas web como de sus respectivas traducciones. Estos elementos también se escogieron teniendo en cuenta el tipo de búsquedas que permite realizar la aplicación AntConc. Esto quiere decir que, aunque la presencia de ciertos elementos es bastante usual en la tipología textual elegida, estos pueden ser bastante difíciles de identificar en su totalidad. Un ejemplo que explica esta cuestión es la presencia de las faltas ortográficas. Aunque se trata de un aspecto importante, sobre todo en páginas web, resulta bastante complicado de buscar si se tiene en cuenta que estos errores abarcan desde tildes hasta la alteración de letras, pasando por la falta de concordancia entre artículo y palabra o sujeto y verbo, entre muchos otros. De este modo, concluí que el gerundio, la voz pasiva, la terminología, las siglas, los nombres de medicamentos y la forma de dirigirse al lector serían elementos apropiados por dos motivos. Por un lado, son aspectos que ya han sido considerados problemáticos en otras investigaciones y, por otro lado, el programa informático escogido no presenta ningún problema para su búsqueda en cada subcorpus al completo.

Una vez escogidos los elementos que se pretende examinar, se procede con la búsqueda y se analizan las cifras obtenidas para poder, finalmente, realizar una comparación entre los resultados del análisis de ambos subcorpus. De este modo, se podrá confirmar si la hipótesis propuesta al inicio del estudio, donde se afirmaba que existirían diferencias entre ambos subcorpus y se añadía que los textos originales serían más cuidadosos, naturales y fluidos que los textos traducidos, es cierta. Asimismo, también se observan las decisiones tomadas tanto por los redactores como por los traductores y se señala cuál podría ser, de acuerdo con todo lo aprendido en el marco teórico, la opción más apropiada para cada uno de los elementos escogidos para el análisis.

6. Marco teórico

6.1. Corpus lingüístico

6.1.1. Definición

De acuerdo con el Diccionario de términos clave del Centro Virtual Cervantes, desde un punto de vista especializado, la palabra corpus puede comprender dos significados. El

primero de ellos “designa la recopilación de material lingüístico hecha con un propósito de investigación concreto, ya sean muestras de oraciones, de enunciados o de textos” (Centro Virtual Cervantes, en línea) mientras que el segundo hace referencia a “una recopilación extensa de textos (escritos, orales o de ambos tipos) recogidos con el fin de servir como muestra representativa de una lengua, como conjunto de datos lingüísticos reales que reflejen el uso de la lengua (o del tipo de lengua específico) del cual quieren ser representativos” (Centro Virtual Cervantes, en línea). Esta última definición se acerca más al tipo de corpus al que se hará referencia a lo largo de todo este trabajo.

Pérez define un corpus como “cualquier colección que contenga más de un texto” (Pérez, 2002, en línea). No obstante, es consciente de que existen una serie de características que esta recopilación de textos tiene que poseer para que pueda ser considerado un corpus como tal. Es por esto que se apoya en las definiciones de distintos autores como Francis (1982) o Sinclair (1991) para mostrar tres implicaciones con las que actualmente cumple cualquier corpus: el formato magnético de sus textos, la representatividad adecuada y una serie de criterios que se tratarán más adelante. Una definición que siga estas pautas y que parece bastante acorde con el término que vamos a tratar sería la propuesta por Sinclair: *“a corpus is a collection of pieces of language text in electronic form, selected according to external criteria to represent, as far as possible, a language or language variety as a source of data for linguistic research”* (Sinclair 2004, citado en Navarro Colorado 2015).

6.1.2. Criterios generales para el diseño de un corpus

Muchos autores han clasificado en sus estudios los distintos tipos de corpus de los que disponemos en la actualidad. Esta distinción se basa en una gran variedad de criterios que van marcando el camino del investigador a la hora de diseñar dicho corpus. Antes de empezar a recopilar los textos, es primordial tener claro qué tipo de corpus se va a llevar a cabo. Es por esto que voy a mostrar a continuación varios de los criterios que he conseguido recopilar a partir de los estudios de diversos autores.

Aunque ninguno de los estudios consultados coincide totalmente con el número exacto de criterios existentes ni su denominación, todos ellos tienen en común el hecho de que es esencial determinar el tamaño del corpus, lo que hace referencia al número de textos o, en ocasiones, de palabras que va a contener. Algunos de ellos, como Torruella y Llisterri (1999) o Tolchinsky (2014), incluyen en esta misma categoría la distribución de los textos. Otro nivel bastante presente en gran parte de las clasificaciones es la especificidad de los textos, es decir, si la temática textual va a ser general o especializada. Determinar si las muestras serán escritas, orales o transcritas es otra de las pautas que debe tenerse en cuenta antes de la recopilación. Pérez incluyó en su clasificación de criterios para corpus especializados la división entre textos originales, traducciones y textos producidos por hablantes nativos o no nativos en su categoría “condición lingüística de los textos” (Pérez, 2002, en línea), donde también aparecía la decisión del medio de producción del texto. El número de lenguas que se va a utilizar y cuáles serán estas lenguas es importante también para saber qué textos se van a seleccionar, lo que significa que un corpus puede dividirse en monolingüe o multilingüe. Esta última categoría incluye las subcategorías de paralelo, comparable o de textos en distintos idiomas de acuerdo con los profesores Sierra, Bel y Lázaro (2018, en línea). Estos mismos

autores también señalaban la importancia de determinar la espontaneidad, la autoría, el tiempo, la accesibilidad, el propósito, la calidad, la codificación o anotación y la documentación, categorías que también aparecen en otros estudios de forma independiente. Flores Acuña (2014) quiso incluir también la temática y el tipo y género textual, un criterio apoyado también por Navarro Colorado (2015). Este último autor insertó en su clasificación la división entre el uso de textos completos o fragmentos, del mismo modo que lo hizo Tolchinsky (2014). Finalmente, tenemos el hecho de que un corpus contenga muestras solamente virtuales, muestras recogidas en papel o una combinación de ambas y la distinción entre textos ficticios o reales. En el caso de tratar solamente con un corpus especializado, Pérez (2002) señala la importancia de la tipología textual, como ya se había mencionado anteriormente, el nivel de tecnicidad y el receptor del texto.

Como conclusión de este apartado, se puede comprobar que, para la recopilación de un corpus, es necesario tomar varias decisiones previamente: ¿utilizar recursos solamente en línea, muestras en papel o una combinación de ambos? ¿Centrarse en un ámbito especializado o crear un corpus que represente el habla en un contexto general? ¿Recoger textos completos o solamente los fragmentos que interesen? Entre estas, habrá muchas más preguntas que un investigador deberá hacerse antes de recopilar un corpus. Todo ello se realizará con el fin de coleccionar un número de textos que presenten las mismas características de modo que el corpus sea representativo del habla en un contexto específico.

6.1.3. Representatividad

Algunos autores consideran la representatividad como otro de los criterios que deben tenerse en cuenta a la hora de compilar textos para la realización de un buen corpus (Sierra, Bel & Lázaro, 2018). De hecho, podría decirse que se trata de uno de los aspectos más importantes a la hora de diseñarlo puesto que una investigación llevada a cabo a partir de un corpus que no tenga la suficiente representatividad podría contaminarla por completo. No obstante, otros investigadores creen que la representatividad es el motivo de la creación del resto de criterios que existen para la elaboración de un corpus (Pérez, 2002). Lo cierto es que este término siempre ha presentado una gran controversia en la lingüística de corpus. Ya lo mencionaba Pérez en *Explotación de los corpóra textuales informatizados para la creación de bases de datos terminológicas basadas en el conocimiento*: “representatividad, estandarización y tipología de los corpóra han sido tres de los temas más debatidos entre la comunidad científica, con opiniones diversas recogidas en varios artículos y propuestas, algunas de ellas hechas en el seno de importantes proyectos europeos” (Pérez, 2002, en línea). Por este mismo motivo, se ha reservado un espacio en el que debatir el sentido de la representatividad en el diseño de corpus.

En el mundo de la lingüística de corpus, siempre ha existido la polémica entre un corpus de calidad frente a uno de cantidad. Esto separaba a aquellos que defendían la realización de un corpus cuantioso frente a aquellos que estaban a favor de un corpus representativo que asegurase muestras suficientes para representar un ámbito, lo que no implicaba un

número mínimo de textos. Este es un tema bastante estudiado y sobre el que han escrito autoras como Corpas y Seghiri:

la pregunta surgiría en este punto pues, a pesar de que los adelantos tecnológicos nos permiten gestionar grandes volúmenes de información, cabe plantearnos éstos si son verdaderamente necesarios o si corpus más reducidos pueden ser igualmente representativos —al menos en un campo concreto de especialidad (Corpas & Seghiri, 2006, p.5).

Como se explicaba anteriormente, siempre ha existido ese enfrentamiento entre la cantidad, categoría que representa a los corpus grandiosos, frente a la calidad, que encarna a los de menor tamaño, pero con muestras suficientes como para considerarse representativo. Esta confrontación presenta un gran problema y es que, al fin y al cabo, ambas opciones tienen su punto de razón. En el caso de un corpus de temática general, la compilación de un muy elevado número de documentos podría ser, dentro de lo que cabe, una tarea sencilla, ya que existen muchos textos que representan el habla general en recursos tanto digitales como físicos. Y, cuantos más documentos se obtengan, más muestras habrá que simbolicen de forma escrita o en audio la representación de una lengua. No obstante, cuando se opta por realizar un corpus especializado, donde se requiere la mayor especificidad posible, puede llegar a ser bastante complicado acercarse siquiera al número de palabras o textos que puede contener un corpus general. Asimismo, no solo los corpus especializados presentan problemas a la hora de conseguir suficientes muestras. Un corpus que requiera traducciones, como pueden ser los comparables o los paralelos, puede complicar la tarea del investigador a la hora de querer recopilar suficientes muestras. Esto es debido a diversos factores que “dificultan el equilibrio de la muestra, como la escasez de traducciones en ciertos países, el estatus de la traducción en un período determinado, la reputación del traductor, el sexo del autor de los originales y de los traductores, etc.” (Rojo, 2013, p.116), lo que no significa que, por el hecho de no alcanzar un número de muestras igual al de un corpus general, no vaya a ser representativo, ya que probablemente no exista tal número de datos en ámbitos de especialidad.

Por otro lado, aunque dependerá de los criterios de uno mismo dependiendo de las necesidades que requiera el corpus que se vaya a diseñar, es primordial conseguir un número suficiente de muestras que pruebe que los datos que se están extrayendo del análisis de todos los textos son válidos y aceptables. De no tener una cifra lo suficientemente elevada como para tener una gran variedad de textos que muestren todas las posibilidades que existen dentro de una misma tipología textual y temática, los resultados obtenidos podrían no estar representando realmente esa tipología o esa temática en especial. Es por esto que, en ocasiones en que dicha tipología textual o temática sí presente suficientes muestras como para compilar un corpus con una cifra considerable de escritos, es necesario encontrar una forma de saber cuándo se puede dejar de recolectar textos. Esto se debe a que, a veces, los corpus se recopilan por necesidad del momento, como podría ocurrir con un corpus ad hoc, en el que se necesitan muestras para ayudar en una traducción, pero el traductor no puede utilizar todo el tiempo que se

le haya ofrecido para el encargo recopilando textos que en conjunto contengan más de un millón de palabras.

6.1.4. Otros problemas para el diseño de un corpus

Como se mencionaba en el apartado anterior, la representatividad es uno de los temas que más complican el uso de corpus en una investigación, ya que, para que los resultados sean fiables, es necesario que el corpus escogido presente una colección suficiente de textos de calidad que sigan unos mismos patrones. Encontrar corpus en la red no es una tarea complicada. De hecho, Sierra, Bel y Lázaro ofrecen una lista de corpus de fácil acceso como el Corpus Diacrónico del Español, el Corpus de Referencia del Español Actual, el British National Corpus, el Corpus del Español de Mark Davies, el Archivo de textos hispánicos, la Base de datos ETDEWEB, el Corpus del Derecho Ambiental Mexicano, entre otros (2018, en línea). El instituto de Lingüística Aplicada de la Universidad de Cádiz (en línea) facilita el Hansard Corpus, el Corpus paralelo ACTRES, el Fono.ELE y el CHILDES como muestras de corpus multilingües, especiales y paralelos, pero también propone otras opciones monolingües para lenguas determinadas como el inglés o el francés. Martin Weisser (2016) ofrece una gran lista de corpus especializados en inglés, entre los que se encuentran el Wolverhampton Business English, el Saarbruecken Corpus of Spoken English, el Oxford Psycholinguistic Database, el Business Letters Corpus o el Air Traffic Control Corpus, entre muchos otros. El problema a todo esto reside en que, a pesar de que se ha comprobado que existe un gran número de corpus de calidad y de fácil acceso, es bastante difícil que alguno de estos cumpla con todas y cada una de las características que requiera el investigador para su propio trabajo. La solución que se presenta ante esta dificultad es la recopilación de un corpus propio, lo que, aunque a simple vista podría parecer una buena solución, hay que considerar las complicaciones que se extienden a partir de esta alternativa.

Al principio de este apartado, se explicaba que para realizar una investigación a partir de una metodología basada en el uso de corpus es necesario que dicho corpus presente una colección suficiente de textos de calidad que sigan unos mismos patrones. El término calidad implica que el corpus que se vaya a utilizar provenga de una fuente fiable, como es el caso de los ejemplos que se han mostrado anteriormente. No obstante, en ocasiones, ninguno de los corpus ya compilados por otros profesionales cumple con los requisitos de la temática de estudio al que uno quiere dedicarse y debe diseñar uno propio. En ese caso, para que cumpla con el criterio de calidad, es primordial que los documentos que lo conforman hayan sido escritos por autores fiables o seleccionados a partir de fuentes que proporcionen cierta seguridad. Esto, en ocasiones, puede suponer un problema debido a que algunas entidades encargan el escrito de textos a distintos profesionales, tanto internos como externos a esta, lo que hace que no aparezcan firmados, sino que simplemente pertenezcan a la entidad como tal. Esto dificulta el conocimiento de su autoría, por lo que es primordial centrarse en la profesionalidad de dicha entidad. No obstante, este no es el único obstáculo que se presenta para un investigador que quiera utilizar una metodología basada en el uso de un corpus que él mismo haya creado. En el caso de uso de traducciones, por ejemplo, es esencial demostrar que los documentos han

sido redactados previamente en otra lengua, buscando sus autores en el idioma original, la lengua principal de la entidad o si aparece firmada por algún traductor.

Conseguir un tamaño adecuado para un corpus que va a ser diseñado por uno mismo es una tarea bastante engorrosa, ya que se habrán seleccionado unos rasgos muy específicos para la investigación y habrá que conseguir que todas las muestras cumplan con esas características. Es cierto que se dispone de fuentes tanto físicas como digitales en abundancia, y que, con la propagación del uso de Internet, el aumento de textos en la red es cada vez mayor. Sin embargo, en según qué temáticas demasiado específicas, sigue siendo bastante complicado encontrar una cantidad suficiente de textos que además presenten cierta calidad o fiabilidad. Esto significa que, si uno decide basar su corpus en textos que traten exclusivamente la insensibilidad congénita al dolor, una enfermedad muy poco común, a pesar de que resulte difícil encontrar documentos que la traten, no puede utilizarse como fuente Wikipedia, donde cualquier usuario tiene la posibilidad de editar la información que proporciona, solo para conseguir una muestra más. En términos de tamaño de un corpus, el único problema no es no encontrar una cantidad suficiente sino el hecho de que no está estipulada la cantidad exacta que se requiere para determinar el tamaño de un corpus representativo.

Otra de las tareas que complican, en cierto modo, la elaboración de un corpus es el hecho de tener que pasar los textos a un formato electrónico, en caso de que hayan sido seleccionados en formato físico, para poder utilizar ciertas herramientas que ayuden a analizarlos. Aunque este pequeño problema tiene una fácil solución, supone un largo periodo de trabajo de escaneado y guardado. En caso de que los textos hayan sido encontrados en formato digital, se omite este paso, pero será necesario igualmente eliminar todo tipo de formato, lo que también implica un buen uso de tiempo.

Los corpus comparables, que son el tipo de corpus que se usarán en esta investigación, también suponen complicaciones que hay que tener en cuenta antes de realizar cualquier investigación. La primera de ellas es que los dos grupos de textos que pueden localizarse dentro del corpus no presentarán jamás los mismos textos como podría ocurrir en un corpus paralelo. Esto implica que se pueden hallar traducciones de quinientas palabras sobre un tema y que los textos originales que traten dicho tema estén formados por cinco mil. El problema que esto supone es que, si se ha decidido utilizar los textos completos, es decir, sin extraer fragmentos de ninguno de ellos para estudiarlos como un conjunto, aunque se solucione igualando el número de palabras en cada grupo para obtener un corpus equilibrado, aquel grupo que tenga una menor cantidad de textos presentará seguramente un número menor de diferencias en su análisis. Esto se podría deber a que el autor o traductor, dependiendo del caso en que esto ocurra, habrá intentado mantener una cohesión y coherencia a lo largo de todo el texto, utilizando los mismos términos o el mismo discurso. Por otro lado, la segunda complicación que puede presentar este tipo de corpus es que, como se mencionaba anteriormente, puede que no quede del todo claro cuándo un texto es original y cuándo una traducción. Sí que es cierto que, en muchas ocasiones, se señalan los nombres de los autores o de los traductores al principio o al final del texto. No obstante, en el caso de los textos utilizados para páginas web, es difícil que esto aparezca. Es por esto mismo que habrá que hallar los orígenes de esa página, esto es, el país de dónde proviene, si existe la opción de otros idiomas para comprobar si, en esa opción, sí aparecen los autores, si solo se ha traducido un texto, pero se pueden encontrar

más fragmentos en otros idiomas o si hay algún elemento en otro idioma con una explicación que ayude a conocer su procedencia.

6.2.Traducción médica

La traducción comenzó como una actividad cuyo único objetivo era la comunicación entre dos grupos de personas. Todo apunta a que surgiera con las primeras transacciones comerciales (Montalt & González, 2014). Desde entonces, esta actividad ha persistido y se ha ido introduciendo en todas y cada una de las disciplinas y ciencias que existen en la actualidad. Una de las características básicas del ser humano es su necesidad de vivir en comunidad. Lo que en un primer momento pudo implicar la convivencia entre un grupo reducido de personas ha evolucionado hasta la situación que existe hoy en día, donde la globalización se considera uno de los pilares fundamentales de nuestra sociedad. Esto ha supuesto que todos los ámbitos de especialidad precisen los conocimientos de los traductores para poder difundir sus novedades y descubrimientos al resto del mundo.

La medicina, al igual que ocurre con la traducción, es una ciencia muy antigua. Como es de esperar, los seres vivos siempre han tenido esa tendencia a luchar por la permanencia y supervivencia de su especie. De ahí que una de sus funciones básicas sea la reproducción. No obstante, no basta con reproducirse para impedir su extinción. Es necesario combatir las distintas amenazas que le acechan para que el ser humano sobreviva y prospere. De ahí que la medicina sea una de las disciplinas que más ha evolucionado con los años, permitiendo que la esperanza de vida de un ser humano ronde los 70-80 años, cuando hace unas décadas no superaba los 40-50. Esto convierte a la traducción médica en uno de los tipos de traducción científica más antiguos. No obstante, esto no significa que, por ello, sea a la que más atención se le presta. De hecho, es más bien todo lo contrario: el volumen de investigaciones sobre traducción médica es bastante más reducido de lo que cabría esperar y la oferta de programas de formación es escasa (Muñoz Miquel, 2009).

Todos los conocimientos médicos que se han ido adquiriendo con los años se han ido registrando por escrito. Actualmente, existen una barbaridad de escritos relacionados con la medicina. Esta gran variedad de documentos se debe a que cada uno posee un objetivo distinto lo que significa que, por ejemplo, un manual de anatomía patológica destinado a los estudiantes de Medicina no contendrá la misma información que el historial médico de un paciente que ha acudido al hospital por un esguince. Con esto quiero decir que el campo de la medicina es tan extenso que parece necesario delimitar los aspectos que cubrirá este estudio.

Para empezar, se explicarán brevemente algunas nociones sobre medicina que debe conocer todo traductor que vaya a dedicarse a este campo. En el libro *Medical translation step by step* (2014), Montalt y González exponen ocho aspectos, mencionados también por otros autores, que diferencian la traducción médica de otros tipos de traducción y que, por tanto, deberían tenerse en cuenta antes de comenzar cualquier encargo relacionado con este ámbito. El primer punto que hay que tener en cuenta cuando se recibe un encargo de traducción médica es que existen distintas especialidades que comprenden esta disciplina (Montalt & González, 2014; Muñoz Miquel, 2009). El funcionamiento de un

hospital se rige por un número determinado de departamentos que se encargan de la gran variedad de situaciones que pueden llevar a una persona a acudir a este tipo de centro. Es importante tener suficientes conocimientos sobre esto ya que puede que exista un mismo término dentro de más de una de estas especialidades que tenga connotaciones diferentes según el campo en que se esté tratando. Como se ha mencionado anteriormente, los textos que envuelven el corpus de este estudio tratan el cáncer colorrectal, por lo que se podrían clasificar tales documentos dentro de la medicina interna, “una especialidad médica con una larga tradición, que aporta, una atención global al enfermo adulto de carácter preferentemente hospitalario” (Sociedad Española de Medicina Interna, en línea), especialidad que, por tanto, incluye el diagnóstico y tratamiento de cualquier tipo de cáncer, entre otras cuestiones.

Otra de las destrezas que debe presentar un traductor médico es la capacidad de comprender cualquier noción médica, incluyendo la terminología (Montalt & González, 2014; Muñoz Miquel, 2009), ya que “*whereas the literary translator’s main focus is normally on aspects such as register, rhythm, puns, character’s attitude, or cultural references, the medical translator’s priority is to deal adequately with factual complexity and accuracy*” (Montalt & González, 2014, p.20). Otros autores, como Congost Maestre, también han utilizado esta comparación entre traducción médica y literaria para resaltar ciertas cuestiones propias de la traducción médica que se están infravalorando. En este caso, la autora señala que, “suele aceptarse que la traducción técnica es más fácil que la literaria solo por el hecho de que los términos son más exactos, sin connotaciones” (Congost Maestre, 1994, p.22). Sin embargo, rebate esta afirmación a partir de dos razonamientos: el primero hace alusión a que el traductor, antes de comenzar con la traducción, debe comprender a la perfección un texto plagado de tecnicismos de los que quizá no ha escuchado antes una parte, que será mayor o menor según si ya ha tratado esa temática o no. El segundo razonamiento se basa en la elevada probabilidad de encontrar terminología especializada para la que puede que no exista una traducción. Todo esto implica que, para dedicarse a la traducción médica, es extremadamente importante comprender todas las ideas que se estén plasmando en un texto y ser capaz de trasladarlas a la lengua meta de forma clara y adecuada, ya que hay que valorar que se está trabajando con lo más preciado que tiene una persona: su salud. Como es normal, es prácticamente imposible conocer todos los conceptos que abarcan cada una de las ramas de la medicina y más todavía si se tiene en cuenta que solo unos pocos traductores médicos poseen títulos universitarios proporcionados por las facultades de ciencias de la salud. Es por esto que es de vital importancia conocer recursos fiables que faciliten información precisa y correcta, tanto a la hora de traducir como de documentarse y formarse previamente.

Además de los dos aspectos mencionados anteriormente, también es primordial que los traductores conozcan las situaciones comunicativas en las que puede darse cualquier tipo de mensaje médico y los géneros que la envuelven (Montalt & González, 2014; Muñoz Miquel, 2009), que sean conscientes de la mala calidad de redacción que pueden presentar estos textos, y que mantengan la confidencialidad, privacidad y respeto hacia los pacientes, entre otras acciones en las que intervenga cualquier cuestión ética o moral (Montalt & González, 2014). Marsh (1999) apunta también la importancia de tener en cuenta a quién va dirigida la traducción, ya que tendrá consecuencias en las estrategias que se vayan a adoptar.

Una vez mencionadas todas las nociones generales que hay que considerar antes de embarcarse en la traducción médica, se prestará especial atención a una de ellas, el género textual, que será lo que, posteriormente, ayude a discernir las peculiaridades del tipo de texto que se ha escogido. De este modo, se obtendrá una visión general de lo que se estudiará más adelante en el análisis, en el que se comprobará si las traducciones cumplen con los aspectos que debe poseer un texto médico para considerarse adecuado.

6.2.1. Género médico

Un texto es un mensaje oral o escrito que hace posible el acto de la comunicación. Según la situación o el contexto en que se den estos mensajes, las necesidades a las que respondan, los temas que traten o las intenciones que tenga su autor, corresponderán a un género u otro. Esto es algo que ya afirmaban algunos investigadores, apuntando que el género textual cubre más aspectos de los que podría parecer a simple vista:

genre is not just a set of formal characteristics —structure, length, tenor, degree of specialization of the information contained, and so— that must be fulfilled, but is also a communicative activity carried out by researchers whose purpose is to convince readers of their conclusions, gain prestige, make the discipline advance, and so on (Montalt & González, 2014, p.57).

Esto significa que para determinar el género de un texto no solo influyen su forma y contenido, sino que su finalidad también juega un papel importante. Esto es algo bastante lógico si se tiene en cuenta que tanto una novela ambientada en un hospital como un manual de medicina pueden contener los mismos conceptos y un nivel de especialización bastante similar en ciertos puntos, pero sus intenciones y objetivos como texto nunca serán los mismos. Mientras que el primero tendría una función poética y el entretenimiento como finalidad, la función del segundo sería referencial ya que su fin es transmitir información de forma objetiva.

Montalt y González organizan los géneros médicos a partir de tres clasificaciones en las que se tienen en cuenta tres aspectos: “*overall rhetorical purpose of the writer*”, “*overall social function*” y “*communication gaps*” (Montalt & González, 2014, pp.58-59). El primero de ellos hace referencia a la función textual de un texto y se dividiría en instructiva, expositiva y argumentativa. El segundo gira entorno a su función, pero en un sentido mucho más específico de la palabra, centrándose en cómo afecta a los receptores la lectura de este texto. Entre estas se puede encontrar la prevención de enfermedades, la transmisión de nuevos descubrimientos o la venta de productos a profesionales, entre otros. El último aspecto divide los textos según el nivel de especialidad que presenten tanto el emisor como el receptor y la relación que los envuelven.

Los textos escogidos para el corpus que se analizará en esta investigación se caracterizan por tener una función expositiva, generalmente, ya que tratan de presentar unos hechos con el fin de informar objetivamente a un lector. Es decir, proporcionan una definición sobre el cáncer colorrectal, muestran sus síntomas, ofrecen distintos tratamientos, etc. No

obstante, algunos de ellos también presentan una función instructiva ya que exponen ciertas conductas que ayudan a disminuir las posibilidades de desarrollar este tipo de cáncer o facilitar instrucciones sobre cómo seguir ciertos tratamientos, entre otros. Estas dos funciones abarcan desde los folletos informativos para pacientes y manuales hasta el historial de un paciente, un atlas o una reseña. Si se considera el segundo aspecto mencionado anteriormente, es decir, la función social, los documentos recopilados para este corpus tendrían como fin ayudar en la prevención de enfermedades, educar al público general o concienciar sobre ciertas acciones que pueden poner en riesgo la salud de cualquier persona, tres ejemplos que engloban la primera subcategoría. Teniendo en cuenta que los textos escogidos para esta investigación prestan especial atención a las causas que provocan esta enfermedad y a la importancia que tiene acudir a revisiones a partir de los 50 años para poder diagnosticar esta enfermedad a tiempo, ya que en la mayoría de los casos en los que esto ocurre, se consigue una cura, esa es seguramente la clasificación que mejor representa a nuestros escritos. Junto a esta categoría, se proporcionan varios ejemplos de géneros que poseen estas funciones como son los comunicados de prensa o la información para pacientes. Finalmente, existe una última clasificación que divide los géneros médicos según el nivel de conocimiento especializado de los receptores y emisores. Teniendo en cuenta que los textos escogidos para este estudio han sido redactados por doctores u otros profesionales del ámbito para pacientes o personas que no tengan estos conocimientos, podrían situarse dentro de la primera clase, donde se encuentran, por ejemplo, las hojas informativas para pacientes o los folletos.

Tras haber revisado tres clasificaciones diferentes de los géneros médicos y haber podido conocer varios de los términos que definen ciertos tipos de textos médicos, se podría concluir que el que mejor se adecúa a los textos que se van a tratar durante toda esta investigación es *Fact Sheet for Patients*. El término como tal describe a la perfección que son exactamente estas muestras: son textos destinados a pacientes que proporcionan informaciones objetivas sobre ciertos conceptos médicos. La definición proporcionada por Montalt y González (2014) parece bastante adecuada para determinar las características principales de este tipo de textos, exceptuando un fragmento en el que se explica que también suelen conocerse como *patient leaflet* o *patient information brochure*. Aunque, en un principio, no me concordó esta parte de la definición, ya que el resto coincidía a la perfección con los documentos que estaba tratando, tecleé dicho término en Internet para ver sus primeras entradas. Lo primero que observé fueron un montón de imágenes de folletos emitidos por entidades sanitarias, por lo que la definición era totalmente correcta. El problema que presenta esta afirmación en caso de querer aplicarla a las muestras recopiladas para este trabajo que hacía que no acabara de convencerme es que difícilmente alguien podría denominar “folleto” a un texto que encuentra plasmado en una página web, y mucho menos si no tiene ni siquiera la opción de descarga en formato pdf. Además, si bien es cierto que los folletos suelen compartir el tipo de discurso que presentan los escritos de este corpus, difieren en ese estilo llamativo clásico de los folletos, que suele presentar dibujos u otros elementos que puedan llamar la atención del lector. No obstante, tras ciertas búsquedas, terminé comprobando que algunas entidades, como es el caso del Health information for Western Australians, lleva directamente al tipo de textos que se han escogido para este estudio, es decir, escritos en formato de página web (html) que tratan nociones médicas con un discurso sencillo para

que cualquier persona sea capaz de entenderlo, independientemente de su nivel de conocimientos especializados, si se pincha sobre un título llamado “*Fact Sheet for Patients*”. Como esa prueba no me parecía suficiente, decidí seguir investigando sobre los géneros textuales en el ámbito médico hasta encontrar nuevas clasificaciones. El estudio *Los géneros de Información para pacientes en el contexto español: una primera aproximación* (Sánchez Castany, 2013) trata los tres principales géneros médicos destinados a pacientes, entre los que claramente se encuentran los *Factsheets*. En ese mismo apartado, Sánchez Castany apunta que este género textual presenta textos de diferentes macroestructuras dependiendo de ciertos factores como el tema principal, el formato o el receptor, por lo que no estaría de más hacer una clasificación de subgéneros entre los que se encontrarían los “folletos, dípticos, trípticos, pósters, historietas, cuentos infantiles, webs o vídeos” (Sánchez Castany, 2013, p.837). Esto, junto con el hecho de que las características que se nombran en la definición de este término son exactamente las mismas que presentan los textos pertenecientes a mi corpus, terminaría de confirmar que el término escogido para estos textos extraídos de páginas web, *Fact Sheet for Patients* o simplemente *Factsheets*, es totalmente correcto.

6.2.1.1. Características de los *Fact Sheet for Patients*

Los *Fact Sheet for Patients* surgen a raíz de la necesidad tanto de controlar y acabar con ciertas enfermedades o condiciones médicas como de prevenirlas, una función que tenían principalmente los profesionales sanitarios, pero que, con el desarrollo de Internet, el acceso a información de cualquier tipo se ha vuelto posible para todos. Esto ha provocado un efecto positivo en la salud y bienestar de las personas gracias al incremento de conocimiento y concienciación sobre ciertos aspectos sobre la salud. Es por esto que se quiere recalcar de nuevo que su función sería generalmente la de informar a un público general sobre ciertas cuestiones relacionadas con la salud y el bienestar de las personas.

Cuando se habla de *Fact Sheet for Patients*, se hace referencia a un género textual que involucra tanto a los textos científicos, debido a su temática especializada en medicina, como a los textos de divulgación, debido a que su objetivo principal es difundir información sobre ciertos aspectos médicos. Si de algo se diferencia este género de un texto “puramente” médico es que estos últimos presentan un emisor y un receptor especialistas en la materia a tratar, mientras que los documentos que se sitúan dentro de la categoría de textos médicos de carácter divulgativo podrían haber sido redactados por una persona que conoce la materia, incluso autodidacta, pero no necesariamente especializada en dicho tema (Félix Fernández & Alarcón Navío, 1998) y, en general, suelen ir destinados a un lego. Por otro lado, Montalt y González (2014) recalcan en su definición de este tipo de textos en específico que sus autores suelen ser organizaciones sanitarias locales, nacionales o internacionales, asociaciones de pacientes, institutos de investigación, hospitales o sociedades médicas, lo que se contrapone totalmente a lo expresado por los autores anteriores. En el caso de las muestras seleccionadas para esta investigación, la mayoría provienen de clínicas, hospitales, aseguradoras y organizaciones y entidades médicas. No obstante, también es cierto que, para los textos redactados originalmente en español, cuya oferta era mucho menor, se ha escogido algún que otro texto redactado por alguna fuente de divulgación, como un periódico. Es

necesario recalcar que, a pesar de esto, se ha tenido en cuenta que los textos provenientes de estas últimas entidades siguieran los mismos patrones que el resto de los escritos.

En relación con el apartado anterior, creo que es importante mencionar que, independientemente de si los autores son especialistas en la materia o no, está claro que se han basado en fuentes bastante similares y de una fiabilidad considerable. Esto significa que, para la redacción de los *Fact Sheet for Patients*, aunque el destinatario sea un lego, es necesario extraer la información de fuentes especializadas que dominen el tema a tratar y que proporcionen la mayor cantidad de información y explicaciones posibles, como son los manuales, tratados o artículos de medicina. A partir de estos textos con un alto nivel de especialización, se realizaría una adaptación para un público general, lo que implicaría el uso de los términos más conocidos y la evasión de las formas latinas, por ejemplo.

En este apartado, se hará un pequeño inciso en el aspecto terminológico surgido en la sección anterior. Una de las características que representa a cualquier texto científico y, por ende, a los *Fact Sheet for Patients*, es la presencia de terminología. Si bien el uso de términos facilita la comprensión de ciertos conceptos muy específicos y favorece la economía del lenguaje, algunos autores expresan la necesidad de “encontrar el equilibrio entre la precisión terminológica y la fácil accesibilidad a la información médica por parte del gran público” (Félix Fernández & Alarcón Navío, 1998, p.173) en este tipo de textos, hallando como solución el uso de un nivel lingüístico básico para que pueda ser comprendido por todos los lectores. También es cierto que la difusión por parte de los medios de comunicación de muchos de los términos que envuelven el campo sanitario ha permitido que la redacción de textos médicos de carácter divulgativo sea mucho más sencilla. Esto quiere decir que existen términos muy específicos para los que no se puede encontrar un equivalente en un lenguaje más coloquial, como el nombre de ciertas enfermedades. No obstante, gracias a los textos de prevención de enfermedades surgidos de una preocupación cada vez mayor por la salud (Félix Fernández & Alarcón Navío, 1998), la sociedad se ha ido familiarizando con ellos y no es necesario buscar un sinónimo que se ajuste a un nivel de menor especialidad.

Si se considera que este tipo de textos tienen un carácter divulgativo, es decir, que su objetivo es llegar a un gran público para emitir ciertas informaciones, lo más lógico sería preparar un discurso accesible a cualquier persona. Por tanto, esta adaptación a la que se hacía referencia anteriormente no solo afecta a la terminología, sino que se extiende también hacia cualquier aspecto discursivo. Para empezar, las oraciones deben ser bastante claras y concisas debido a que se está tratando una temática que ya de por sí no es muy habitual en el día a día del lector y puede que le cueste concentrarse en la lectura. Elaborar oraciones cortas, con un orden claro de sujeto + verbo + predicado, evitará muchas confusiones. Asimismo, y al contrario de lo que ocurriría en un texto literario, es bastante común repetir los términos clave constantemente en lugar de buscar sinónimos (Marsh, 1999) a pesar de que pueda sonar repetitivo. Esto se debe a que en ocasiones los términos poseen ciertos matices que hacen que sus sinónimos no tengan un significado exactamente igual. El uso de definiciones, explicaciones e incluso metáforas y ejemplificaciones es bastante común en este tipo de textos. En un contexto mucho más especializado suelen omitirse porque se da por hecho que una persona que lea tal tipo de textos ya conoce los conceptos más básicos. Finalmente, es también importante tener en

cuenta el orden de la información que se ofrece. Esto implica que se puede tratar el tema desde la información más general hacia la más específica, o que se seguirá un orden más lineal. Esta última técnica supone que, en el caso de una enfermedad, se comience por las causas que la provocan, se continúe por los síntomas y el diagnóstico, y se finalice con el tratamiento y las posibles consecuencias de haber padecido tal enfermedad.

Además del lenguaje, existen más aspectos que caracterizan a los *Fact Sheet for Patients* y que, casualmente, los diferencian de textos con un alto nivel de especialidad. El primero de ellos hace referencia a la temática. Si bien es cierto que ambos tratan conceptos especializados que pertenecen al ámbito de la medicina, los textos “puramente” científicos se centran en temas exclusivamente científicos. Por el contrario, los textos de carácter divulgativo pueden permitirse la libertad de poseer un rigor científico más reducido (Félix Fernández & Alarcón Navío, 1998).

Finalmente, existe una última característica que me parece importante mencionar antes de comenzar al siguiente apartado. Los textos médicos abarcan géneros tan distintos entre sí que pueden encontrarse en cualquier fuente y formato. La diferencia entre un texto de un alto nivel de especialización con un *Fact Sheet for Patients* reside en el hecho de que los primeros pueden encontrarse en conferencias, revistas de divulgación especializadas u obras monográficas (Félix Fernández & Alarcón Navío, 1998), entre otros, mientras que los textos médicos de carácter divulgativo aparecen en revistas, periódicos y otras fuentes de divulgación, como Internet.

6.2.2. Problemas en la traducción de textos médicos

6.2.2.1. Terminología

Cada ciencia o disciplina posee un lenguaje que describe los distintos conceptos que las envuelven. Las palabras que conforman este lenguaje pueden aparecer en otros campos, pero se caracterizan por tener un significado específico dentro de un área de especialidad. Cuando este tipo de lenguaje se utiliza con el fin de que exista una comunicación dentro de un campo científico, se le llama tecnolecto, jerga profesional o lenguaje científico (Congost Maestre, 1994). En la actualidad existe tal número de campos y áreas de especialidad, que conseguir encontrar un solo término para cada concepto existente es una tarea prácticamente imposible por dos motivos: o no existe un término en la lengua meta para tal concepto o existe más de uno. De ahí que surgiera la necesidad de una normalización terminológica que combatiera “ambigüedades y la pérdida de información en un mundo como el nuestro” (Congost Maestre, 1994, p.26). Como bien señala Pérez, esta normalización no solo surge debido a “las posibles confusiones en la utilización de los términos” sino que “el deseo de establecer criterios lingüísticos claros para la designación y la descripción de los conceptos” (Pérez, 2002, en línea) es otra de sus mayores causas. Es por esto mismo que la principal tarea de aquellos que se dedican a la normalización terminológica es la de combatir la sinonimia y homonimia, buscando el mejor término para designar un concepto.

Aunque la normalización pueda parecer una muy buena iniciativa para ayudar a los traductores a encontrar el término más adecuado para denominar conceptos

pertenecientes a campos especializados, la rapidez con la que aparecen nuevos conceptos a los que hay que asignar una denominación dificultan la aplicación de una solución ante este caos terminológico. Esto quiere decir que, debido a que la comunicación de progresos y novedades en cualquier ámbito debe ser instantánea, prácticamente no se concede un margen de tiempo a los técnicos y lingüistas para escoger un buen término en la lengua meta antes de que se realicen las primeras traducciones que contengan dicho concepto.

Anteriormente se mencionaba que la rapidez que se exige para traducir información sobre cualquier adelanto o invento provoca que traductores, técnicos y medios de comunicación decidan cómo trasladar un concepto de una lengua a otra sin saber muy bien cómo, ya que no existe ningún tipo de normalización sobre ese significado que acaba de surgir. En ocasiones, debido a que todos estos progresos provienen del extranjero, se opta por utilizar anglicismos o barbarismos.

Otro problema que surge de la falta de una terminología normalizada en el ámbito científico-técnico y probablemente uno de los más graves es que la constante innovación en estos ámbitos dificulta todavía más la presencia de términos estancos. Esto causa esa escasez de diccionarios o vocabularios actuales en estas ramas debido a la continua creación y actualización que desacredita a todos los anteriores. En palabras de Congost Maestre, “[un] vocabulario técnico, aunque sea reciente, en seguida pierde actualidad porque pronto aparecen otros de nueva gama que lo postergan, hasta el punto de que podríamos decir que el vocabulario técnico queda obsoleto de un curso a otro” (Congost Maestre, 1994, p.28).

Ante esta situación, Newmark (1988) presenta una solución que combate la necesidad de buscar un nombre para un concepto para el que todavía no existe una traducción (citado en Congost Maestre, 1994, p.30). Consiste en el uso de un término descriptivo, lo que supone que, en lugar de utilizar un término o expresión terminológica que no haya sido normalizada, se utilice una descripción de lo que designa dicho término. No obstante, Congost Maestre expresa que

el uso del término técnico en el texto de la lengua terminal es ciertamente preferible en aquellos casos donde el fragmento es técnico y hay pruebas evidentes de que quizás se ha usado el término o expresión descriptiva sólo porque el término técnico es raro o falta en la lengua terminal (Congost Maestre, 1994, p.30).

6.2.2.2.Otros problemas

Si se intenta realizar una clasificación de las distintas dificultades que supone llevar a cabo una traducción médica, se puede observar que, aunque no existe una misma lista que se reproduzca en todos los estudios que tratan esta temática, sí que coinciden varios de los conceptos que las componen. Muchos de ellos mantienen cierta relación con la terminología o con el léxico en general, por lo que he decidido empezar por presentar las distintas subcategorías que envuelven al lenguaje médico. Entre ellos, se encuentran los

false friends o falsos amigos, los préstamos, los calcos, los neologismos, las expresiones idiomáticas (Félix & Ortega, 1998; Congost Maestre, 1994; Muñoz Miquel, 2016) o

los epónimos [que] identifican un descubrimiento o un invento con un antropónimo, asociado con él, que tal vez no se reconozca en la lengua de otro país: por ejemplo, el páncreas menor se conoce también como páncreas de Willis o de Winslow, pero solamente en inglés (Newmark, 1988 en Congost Maestre, 1994, p.32).

Los textos médicos de carácter divulgativo, como es el caso de los *Fact Sheet for Patients*, se caracterizan también por un elemento que es necesario saber identificar para poder trasladar correctamente a otra lengua: las metáforas lexicalizadas. Estas son muy comunes en este tipo de textos puesto que ayudan al emisor a completar las explicaciones de ciertas ideas que puedan resultar difíciles de procesar para un lego. Estas metáforas ayudan al lector a visualizar ciertos conceptos y a comprender mejor el mensaje.

Asimismo, el receptor de los *Fact Sheet for Patients* no suele ser una persona especializada en la temática. Es por esto que es bastante común hallar distintos registros lingüísticos según el tipo de texto médico que se vaya a tratar, lo que engloba desde términos altamente especializados a palabras o construcciones coloquiales (Marsh, 1999). Esto es algo que confirmaba Congost Maestre y ofrecía varios ejemplos. Es posible encontrarse con un lenguaje “de andar por casa” como “paperas” o “anginas”, con términos que posean cierto grado de formalidad, como amigdalitis o difteria, y latinismos, que podrían clasificarse dentro de un nivel académico. En los textos escogidos, lo más lógico sería encontrar un léxico entre coloquial y formal que pueda ser comprendido por cualquier persona. Además de esto, es necesario saber reconocer cuándo un término o una construcción de carácter más formal va a ser comprendido por el lego o no. De no ser así, el traductor debería ser capaz de reformularlo o adaptarlo a partir de un “doblete terminológico” que posea tal concepto (Muñoz Miquel, 2016).

Otra dificultad que puede presentar cualquier texto médico es la presencia de nombres de marcas en lugar de productos. Esta tendencia es bastante común en el habla de cualquier lengua por lo que ha conseguido introducirse en ciertos lenguajes especializados. En reiteradas ocasiones, se podrá comprobar un uso frecuente del nombre comercial de un fármaco antes que su nombre original. En España, es bastante habitual usar Nolotil (nombre comercial) para apelar al Metamizol (medicamento) o Seguril para recurrir a la Furosemida. En cambio, en inglés, es mucho más común encontrar los compuestos químicos para tratar los fármacos. Como traductor, todo esto implica la necesidad de reconocer tales términos y ser capaz de trasladarlos a uno que el receptor de la lengua meta pueda comprender.

El léxico no es el único factor que hay que tener en cuenta a la hora de realizar una traducción médico-sanitaria. La gramática también juega un papel muy importante y existen ciertos rasgos de esta que son habituales en todos los textos médicos y que hay que saber reflejar. La objetividad, la exactitud y la economía expresiva son tres de estas características que podrían aplicarse tanto a la gramática como al léxico. Como es obvio,

un emisor que pretenda informar sobre una condición o enfermedad querrá ser directo, claro y conciso; querrá utilizar los términos que expresen con mayor exactitud las ideas que pretende transmitir, siempre que sean lo suficientemente conocidos por el lector como para que los comprenda al instante. Por tanto, y a diferencia de otros géneros textuales, no esperará de sus resultados un mensaje bello y adornado. Esto se ve reflejado en el uso de frases concisas y formas nominalizadas. Uno de los problemas que puede suponer esta necesidad de reducir al máximo el mensaje sería el uso de abreviaturas, siglas y acrónimos (Muñoz Miquel, 2016), lo que no sería un problema si estuvieran normalizadas. Por otro lado, la objetividad se expresa a partir de ciertos elementos como oraciones en pasiva que restan importancia al individuo que genera la acción, dándosela así a quién le afecta, o modos impersonales. A todo esto, se añade la presencia del gerundio médico o de muchos adjetivos para un solo sustantivo (Muñoz Miquel, 2016) que, aunque no representen la concisión o la objetividad de los textos médicos, se recurre mucho a ellos en lengua inglesa, lo que se ve reflejado en las traducciones al español.

Otro problema que presenta la traducción médica es la gran cantidad de fuentes que se pueden encontrar tanto en formato electrónico como físico. Aunque a primera vista esto podría parecer una ventaja más que un problema, hay que saber reconocer qué recursos presentan cierto grado de fiabilidad. De no ser así, el traductor podría acabar utilizando un término erróneo y, teniendo en cuenta que se trata de la salud de las personas, no es lo más idóneo. Por tanto, debe contrastarse la información encontrada en una fuente que no parezca del todo fiable con otras que proporcionen seguridad para asegurar la adecuación del término.

Las diferencias culturales pueden suponer también cierta dificultad para el traductor a la hora de trasladar un texto de una lengua a otra. Cuando se menciona “elementos culturales” en un contexto médico se hace referencia a los sistemas sanitarios de cada país, a sus normas sociales, creencias sobre la salud y la enfermedad o incluso las denominaciones de cargos médicos (Muñoz Miquel, 2016). Dentro de este problema, Marsh (1999) destacaba las diferencias en medidas y convenciones entre comunidades, en la denominación de ciertas sustancias químicas, enfermedades o síndromes e incluso en las variaciones en la ortografía de algunos nombres propios. Dependiendo del tipo de texto y lo que pretenda transmitirse con él, será necesario realizar una adaptación a la cultura de la lengua meta o añadir explicaciones o términos que ayuden al lector a comprender el papel de ese cargo o cómo funciona el sistema sanitario de un país que no es el suyo para evitar cualquier tipo de confusión o malentendido.

6.3. Traducción de páginas web

Actualmente, Internet es una de las fuentes de información más consultadas en todo el mundo. Es rápida, de fácil acceso tanto para introducir como para extraer datos y bastante asequible. De hecho, es prácticamente imposible vivir sin acceso a Internet en un país desarrollado, ya que cada vez más trabajos requieren esta herramienta para su funcionamiento, los estudiantes la utilizan a diario para la recopilación de información o para resolver dudas e incluso la mayoría de personas llevan consigo un dispositivo con 4G a todas partes, lo que les permite estar siempre en contacto con cualquier individuo, comprar, vender, entretenerse mediante redes sociales, entre muchas otras actividades.

Internet es uno de los factores más influyentes en la globalización. Gracias a este, cualquier persona puede conocer lo que pasa a más de diez mil kilómetros de distancia sin necesidad de desplazarse. Puede acceder a textos, vídeos o imágenes sin importar su autor, su procedencia o su fecha de publicación. Hoy en día, si se difunde algún escrito en Internet, lo más probable es que sea con la intención de que llegue al mayor número de personas posible, ya que este recurso lo permite. Es por esto que muchas entidades aprovechan esta condición para tratar de alcanzar un público más amplio con aquello que pretenden transmitir en sus páginas web. Como ya disponen del canal, lo único que les hace falta es poder trasladar sus textos a todos aquellos idiomas que les permitan llegar a los destinatarios que les interesen.

La traducción de una página web puede suponer una gran ventaja para la entidad u organización a la que representa ya que le abre las puertas a una mayor cantidad de visitantes. Es cierto que, actualmente, la mayoría de las personas presentan cierto nivel de inglés, ya sea en mayor o menor medida. No obstante, los lectores sienten preferencia por aquellos textos escritos en su lengua materna por dos motivos. Por un lado, comprender un texto en otra lengua supone un mayor esfuerzo para el usuario y, por tanto, más tiempo y, por otro lado, los textos que aparecen en estos sitios web pueden contener cierta terminología cuyo equivalente en un segundo idioma se desconozca. Esto significa que, en el caso de tener dos opciones, una escrita en inglés y otra en su lengua materna, el público escogerá aquella que le presente más facilidades a la hora de retener la información. Es por eso que las grandes compañías o entidades internacionales ofrecen su contenido en distintos idiomas.

El problema que presentan este tipo de traducciones es que suelen ser bastante costosas y conllevan un largo proceso debido a que las técnicas empleadas no consisten solamente en el simple traslado del mensaje de una lengua a otra. La traducción de páginas web va mucho más allá. Los elementos culturales y todas las especificaciones técnicas que están envueltas en el correcto funcionamiento de una página web son dos de los factores a los que deberá prestar especial atención el traductor. De no ser así, la página podría provocar el efecto contrario al que se pretende conseguir. Sin embargo, muchas de los directivos de las entidades no son conscientes de este problema y encargan estas tareas a personas o incluso máquinas que no están preparadas para realizar un trabajo de calidad. Es por esto que creo que es de suma importancia recalcar que “economizar durante el proceso de traducción por no comprar nuevos nombres de dominio sería mala decisión de marketing: los gastos que se hacen traduciendo la página profesionalmente al final se amortizan muy rápido” (Berends, 2007, p.25).

Como en este estudio se va a lidiar con textos encontrados en páginas web de entidades sanitarias, se tratarán primeramente las páginas web como tal para reconocer las distintas características que las componen. Posteriormente, se hablará del proceso que debe seguirse para traducir correctamente una página web y de las tres estrategias principales que requiere este tipo de textos. Finalmente, se profundizará en los problemas y dificultades de la traducción de páginas web, también conocida como localización, para poder realizar el análisis del corpus recopilado para esta investigación.

6.3.1. Características de las páginas web

Para analizar el contenido de un sitio web es necesario prestar atención a dos aspectos: el estilo y la sintaxis (Berends, 2007). A pesar de que este estudio se centrará en el contenido textual, se realizará una breve mención al aspecto de un sitio web y los distintos elementos que lo componen.

El aspecto de una página web no viene regido por unas normas. Es decir, el creador de dicha página tiene total libertad para situar el mensaje que pretenda transmitir de la manera que crea conveniente. Puede optar por dividir su mensaje en párrafos o unir todo el contenido en un solo bloque; puede llenar la pantalla de colores vivos o utilizar solamente tonos neutros; puede insertar imágenes y enlaces que lleven a otras ventanas para no saturar la página principal; puede crear una página de introducción del contenido o exponer directamente la información que le interesa emitir al receptor. La cuestión es que, aunque el estilo puede ser muy variable, el éxito o fracaso de la institución depende en gran parte de este.

Berends (2007) propone tres aspectos a tener en cuenta a la hora de crear una página web. El primero de ellos es la claridad y la facilidad. En esta categoría hace referencia a la importancia de una estructura clara, con títulos y subtítulos que señalen la información que completa el párrafo, para que así, el lector pueda acudir directamente a aquello que está buscando. Añade, además, que, según algunas investigaciones, el receptor prefiere textos divididos en apartados de cinco renglones. Esto implica que la información debe ser directa y concisa y estar estructurada de forma clara y cómoda para su lectura. Otro de los elementos a los que da importancia es a la alineación de textos, preferiblemente a la izquierda, o a la elusión de columnas ya que dificultan una lectura ligera. El último aspecto hace referencia a la simplicidad de la página. El creador de la página debe reducir la información y el uso de otros elementos al mínimo ya que, de lo contrario, saturará el sitio web. Esto se aplica tanto a colores, tipos de letra, imágenes, cuadros u otros elementos que puedan añadirse para complementar el texto principal.

Por otro lado, esta misma autora habla también de la sintaxis y del léxico (Berends, 2007). A pesar de los distintos objetivos que puedan poseer las páginas web, como transmitir información, convencer o publicitar un producto, o la gran variedad de temáticas que pueden tratarse, como medicina, cocina o informática, las características lingüísticas serán, en gran parte, las mismas. Teniendo en cuenta que Internet es una red a la que tiene acceso cualquier persona, lo más lógico será que las páginas web estén diseñadas para todos los públicos. De este modo, el alcance será mucho mayor que si se diseña solamente para un grupo en concreto. Es por esto que lo más aconsejable es utilizar un léxico sencillo y accesible. Esto puede ser un poco más complicado en ciertos ámbitos especializados como las ingenierías o el derecho. Aun así, existen ciertos campos como la medicina que han ido desarrollando vocablos menos técnicos para muchos de sus conceptos más básicos. En cuanto a la sintaxis, se tratan varias nociones que deberían tenerse en cuenta. Para empezar, los títulos que se dan a cada una de las secciones que posee la página web deben “cubrir los rasgos principales del contenido del texto” (Berends, 2007, p.9). Eso no implica que deban tener una gran longitud sino todo lo contrario: deberían ser lo más concisos posibles. A su misma vez, deben tener ese toque llamativo que incite al público a querer leer esa sección. Otra técnica que suele utilizarse para preceder a un párrafo es el uso de preguntas personales ya que dirigirse directamente al lector incrementa el interés de este por el contenido. De hecho, suele ser bastante común en las páginas web con un

objetivo informativo ya que se presentan las preguntas en forma de la duda que podría tener el visitante del sitio. Otra característica que pueden presentar las webs es la presencia de frases que conduzcan a otras pestañas con nueva información para no saturar la página principal. Estos enlaces pueden aparecer de diferentes formas. Aunque el “haz click aquí” es bastante común, es preferible evitarlo y sustituirlo por “más información” o cualquier expresión que indique el contenido que se encontrará en dicho enlace (Berends, 2007). Para terminar, las páginas web suelen contener imágenes que adornan y complementan el sitio y que atraen en cierto modo al lector. Es por esto que es imprescindible saber redactar un pie de foto ya que lo más probable es que sea uno de los primeros textos que se lean de la página. Para que el pie de foto tenga el efecto deseado no se debe describir la imagen ya que es algo que puede apreciarse a simple vista, sino que se deberá ofrecer información relacionada con lo que aparece en dicha imagen.

6.3.2. Proceso de traducción de una página web

Cuando ya se conocen algunos de los aspectos que deben contener los sitios web para que estos tengan cierto éxito entre el público, hay que determinar los pasos que deben seguirse para realizar una traducción adecuada. Tanto Berends (2007) como Corte (2002) coinciden en que uno de los primeros pasos consiste en identificar la función o el objetivo del texto. En el caso de este estudio, las páginas web de donde se han extraído los textos tienen una función principalmente informativa. Corte (2002) incide un poco más en este aspecto e insiste en diferenciar entre determinar el efecto que pretende conseguir la página a simple vista y los objetivos de la entidad como tal. Ambas autoras coinciden también en la importancia de identificar a los usuarios y sus intereses. Todo esto afectará claramente en la forma de dirigirse al lector ya que habrá que dejar bastante claros estos objetivos. Corte (2002) añade otras etapas al proceso de traducción de una página web como determinar a los competidores, conocer los requisitos legales específicos del país a cuya lengua se vaya a traducir y los requisitos técnicos, tener en cuenta las características culturales, religiosas y sociales de los usuarios a los que va dirigido el sitio web y considerar la legislación en materia de protección de la información personal y copyright.

Una vez identificados todos los aspectos externos que pueden afectar al texto y a la forma en la que se dirigirá la traducción, habrá que sopesar las distintas estrategias disponibles para este tipo de documentos. El traslado de una página web de una lengua a otra no consiste solamente en la traducción del texto, sino que hay que realizar una adaptación de ciertos elementos del contenido e incluso del software. Es por esto que las estrategias a las que se hará referencia a continuación no se centran tanto en los aspectos lingüísticos del texto sino más bien en los aspectos culturales.

La primera estrategia a la que se hará mención es la internacionalización. Es la técnica que más utilizan las grandes marcas ya consolidadas. La intención de esta estrategia es la de transmitir el mismo mensaje y causar el mismo efecto a todas las comunidades, independientemente de su lengua o cultura. Esto implica mantener todo tipo de contenido textual, aspecto, estructura, imágenes y símbolos. Solamente se realizan cambios en casos donde puedan darse malentendidos como en las fechas, números, moneda o información de contacto, entre otros (Berends, 2007). Según algunos autores como Corte (2002), esta estrategia consiste en retirar la información local y crear una especie de “plantilla” que

pueda aplicarse a todas las lenguas y culturas. De hecho, algunos autores consideran que a los sitios web se les aplica primero la internacionalización para poder realizar posteriormente una localización (LISA, 1998 en De la Cova, 2017).

Por otro lado, está la localización. Esta podría considerarse la estrategia opuesta a la internacionalización ya que sus objetivos son totalmente contrarios. Consiste en realizar “una adaptación total de un sitio web a un idioma y cultura diferente” (Berends, 2007, p.27). Algo que caracteriza a esta técnica es que al utilizarla no debería notarse que se ha hecho una traducción (Corte, 2002; Berends, 2007). Esto quiere decir que, si originalmente la página ha sido escrita en inglés porque la entidad es británica, el lector no debe identificar ningún elemento perteneciente al Reino Unido ni su lengua. En muchas páginas web suelen añadirse datos del país al que pertenece el organismo o la empresa. Por ejemplo, la versión francesa de la página de Nestlé tiene un apartado llamado *Nestlé en France* donde habla de los orígenes de la empresa y de sus objetivos actuales de forma muy resumida, pero a su misma vez, de forma muy cercana a los visitantes de habla francesa ya que todo el texto está escrito en primera persona del plural; no obstante, en la versión española existe un apartado donde también se habla de la historia de la empresa pero de manera mucho más extensa y desde un punto de vista más frío con respecto al lector. De este modo, se puede ver el objetivo de una buena localización que es adaptar el contenido de la web a una región, lengua y cultura meta en lugar de solo traducir la información de manera que no se pueda notar que se ha redactado originalmente en otro idioma.

La globalización es una técnica que comprende aspectos de las dos estrategias mencionadas anteriormente (Corte, 2002; Berends, 2007). Esto es algo que ya confirmaba la organización LISA en su definición del término: “*globalization addresses the business issues associated with taking a product global. In the globalization of high-tech products this involves integrating localization throughout a company, after proper internationalization and product design*” (LISA, 1998, p.3 en De la Cova, 2017, p.90). De la Cova definía este concepto como “el proceso que engloba todas aquellas estrategias y tareas necesarias para que un producto se comercialice y tenga éxito en un mercado determinado” (De la Cova, 2017, p.90). A esto añade que podría considerarse una estrategia empresarial puesto que todas las decisiones tomadas inciden en los objetivos de la empresa y en el tipo de mercado al que va dirigido el producto. Otro aspecto que cabe recalcar es que este tipo de adaptación puede utilizarse en varios países sin la necesidad de hacer grandes cambios (Corte, 2002). Esto significa que, aunque no supone realizar una adaptación total como podía ocurrir en la localización, sí que es necesario emplear ciertas estrategias, como la exclusión de alguna información (Berends, 2007).

Por tanto, se podría concluir que tanto la internacionalización como la globalización son estrategias que se acercan más al texto original mientras que la localización se adapta a la lengua meta. Asimismo, tanto la internacionalización como la localización se consideran procesos largos y costosos, pero que salen rentables a largo plazo ya que suponen la apertura de nuevos mercados y un aumento de ventas (Corte, 2002).

6.3.3. Problemas en localización

Tras haber definido ciertas nociones relacionadas con las páginas web, se procederá a tratar su relación con la traducción. Aunque existen autores que opinan que la

diferenciación entre traducción y localización está desapareciendo, De la Cova (2017) explica de forma muy clara las principales disparidades entre ambos conceptos. La primera de ellas se centra en que la localización se utiliza principalmente en productos digitales que requieren una adaptación que no es aplicable a ningún otro tipo de traducción. Lo cierto es que, si se piensa en otras clases de traducción, como la jurídica o la científica, por ejemplo, se puede comprobar que ninguna de ellas requiere ninguna especificidad técnica más allá del texto en sí como ocurre con la localización. También añade que el grado de adaptación cultural suele ser bastante más elevado que en otro tipo de textos. Los cambios culturales suelen estar muy presentes en ciertos tipos de traducción como la audiovisual o la literaria, pero no siempre son obligatorios. Si una serie de televisión está basada en la convivencia de seis amigos en la ciudad estadounidense de Nueva York, no será necesario adaptar el tipo de moneda o el prefijo de los números de teléfono. En cambio, si se trata de una página web que preste ciertos servicios en diferentes países, habrá que adaptar los precios y las formas de contacto según el país al que se haga referencia. Finalmente, añade que la localización es considerada uno de los diferentes servicios que puede proporcionar una empresa de traducción según la norma ISO/FDS 17100 (De la Cova, 2017, p.101).

Ahora que se ha comprobado que la localización puede considerarse un tipo de traducción independiente, se tratarán los distintos problemas que puede dar a la hora de llevarla a cabo. Internet es una fuente de información que tiene la gran ventaja de permitir constantes actualizaciones a los elementos que lo componen. Introducir ciertos cambios en una página web no supone el mismo tiempo ni coste que implica eliminar una palabra de una obra que ya ha sido impresa en papel. De hecho, cualquier empresa o institución que se precie, programará pruebas de calidad a sus páginas web cada cierto tiempo para asegurarse de que no haya errores ni información desactualizada. No obstante, es común que se pasen por alto ciertos fallos a los que la vista ya se ha acostumbrado tras leer un mismo texto en reiteradas ocasiones.

Tras leer distintas investigaciones, la clasificación más completa y que mejor se adapta a este estudio es la realizada por Olvera-Lobo y Castillo-Rodríguez (2018). Estas autoras proponen una extensa serie de recomendaciones para la traducción de páginas web que ayudará en la identificación de problemas de los textos traducidos al español del corpus utilizado en este trabajo. Sobre los distintos bloques que abarcan dicha clasificación, se prestará especial atención a los aspectos léxicos, a los morfosintácticos y a los gramaticales. Posteriormente, se hará un breve repaso a los aspectos culturales y técnicos.

Como el objetivo de este estudio se centra en el contenido textual de los documentos recopilados para el corpus, se tratarán, primeramente, los problemas léxicos que puedan darse en la traducción de páginas web. El léxico es una parte fundamental de cualquier texto ya que en este recae la mayor parte del contenido del mensaje. Es por esto que es imprescindible evitar cuanto menos cualquier tipo de error. Lo primero que debe eludirse en la traducción de una página web y que podría aplicarse a otro tipo de traducciones es el uso de calcos, préstamos o *false friends*. Aunque muchos de estos pueden haberse introducido en la lengua meta de tal modo que sean muy difíciles de reconocer, es fundamental evitarlos a toda costa, sobre todo los *false friends*, ya que pueden llegar a modificar el significado del mensaje. Dos aspectos que se mencionan en el apartado léxico están directamente relacionados con la terminología: la falta de coherencia y la

falta de precisión. Aunque una buena redacción en español siempre ha sido marcada por la presencia de sinónimos que rehúyan la repetición, en el caso de traducciones especializadas, lo mejor es hacer hincapié en los mismos términos para evitar confusiones sobre si dos términos hacen referencia al mismo concepto o no. De hecho, es muy probable que, aunque dos vocablos puedan parecer sinónimos totales, no lo sean y contengan pequeños matices que cambien el mensaje que pretende transmitirse. Esto afecta directamente al hecho de escoger el término adecuado, un problema al que se le ha dado mucha importancia en otros estudios como el de Corte (2002). El uso de un término distinto al que se utiliza en el texto original podría cambiar el mensaje y traer graves consecuencias. Finalmente, es importante prestar atención al uso incorrecto de siglas y acrónimos y a la presencia de siglas en inglés junto a una traducción del significado de las palabras que la componen en español ya que se trata de una estrategia que es preferible eludir.

El segundo apartado que puede ser de ayuda para este estudio hace referencia a la gramática y la sintaxis. El primer aspecto al que se hace referencia es el incumplimiento de reglas gramaticales y sintácticas, como pueden ser los errores de concordancia y género o el calco de estructuras sintácticas (Olvera-Lobo y Castillo-Rodríguez, 2018). El estilo también afecta directamente en la calidad de redacción. Es por esto que hay que hacer uso de un registro y de una variedad adecuados y evitar los errores tanto ortográficos (Corte, 2002) como tipográficos. Dentro de estos cuatro conceptos, se prestará especial atención al correcto uso de variedad lingüística y a los errores ortográficos respectivamente. Por un lado, a pesar de que es bastante obvio que un traductor de español que trabaje en España será capaz de utilizar los términos o vocablos adecuados a su variedad lingüística, es posible que la persona a la que se le encargue la traducción de una página web, al igual que puede ocurrir con cualquier otro encargo, no sea precisamente nativa de español y no conozca muy bien las palabras empleadas en cada región. Esto podría afectar directamente al acceso de la página, puesto que, si un español quiere buscar una óptica donde comprarse gafas, utilizará este mismo término. Si la página web de un pequeño comercio español tiene en su página el término de un objeto utilizado para corregir los problemas de visión como “lentes”, los visitantes españoles serán mucho más reducidos que los de otras regiones que hablen español. Por otro lado, aunque se puede extraer el texto de la página y traducirlo con un procesador de textos que señale las erratas, siempre se puede olvidar algún error. Por ejemplo, si se quiere utilizar un “sé” del verbo “saber” y, en el momento de redacción, el traductor no añade la tilde, puede que el procesador no lo señale como error debido a que existe el pronombre “se” sin acento y tanto el traductor como el corrector, al leer demasiadas veces el texto en busca de fallos, normalice ese “se” y no lo reconozca como errata.

Finalmente, están los aspectos culturales y los técnicos. Los aspectos culturales abarcan cualquier elemento propio de una región que varíe al dirigir el texto a otro país o cultura. Esto englobaría la moneda, el formato de la fecha, la hora, los enlaces o los números de teléfono. En este apartado es importante tener cuidado con las explicitaciones innecesarias y, sobre todo, incorrectas y las imágenes con significado cultural. Corte (2002) también hace un pequeño apunte al uso de elementos visuales y señala la especial atención que debe prestarse al uso de imágenes incongruentes. Esto puede deberse a la omisión de alguna información innecesaria en la lengua meta, el descuadre del texto debido a que una lengua ocupe más espacio que otra, en entre otros motivos. En el caso

de los aspectos técnicos, debería atenderse a la selección del idioma, la estructura o la búsqueda dentro del sitio web. En el caso de este estudio, es de vital importancia que se hayan traducido los documentos descargables, el contenido audiovisual y los enlaces. De acuerdo con Corte (2002), dejar algunos fragmentos en el idioma original es uno de los mayores problemas que suelen encontrarse en la traducción de páginas web. Es bastante lógico que este sea uno de los errores más frecuentes si se tiene en cuenta que la traducción de un documento consiste en ir trasladando el texto de izquierda a derecha, de arriba abajo (en caso de tratar con las lenguas de trabajo de esta investigación, que serían el español y el inglés). Sin embargo, la traducción de una página web implica otros enlaces, pestañas, cuadros de diálogo, recuadros de distintos tamaños y letras de colores, lo que influye en el hecho de que al traductor pase por alto algún fragmento.

7. Investigación

7.1. Introducción

Antes de comenzar con la investigación como tal, es necesario tratar el proceso que va a seguirse a lo largo de esta sección. Para llevar a cabo cualquier investigación científica, lo primordial es escoger una metodología. De acuerdo con muchos manuales dedicados a esta temática, se han fijado tres diseños o métodos de investigación: los cualitativos, los cuantitativos y los mixtos (Rojo, 2013). Cada una de estas categorías abarca una gran cantidad de modelos que no están sujetos a estas de forma fija, sino que pueden variar y alterarse según el investigador crea conveniente para su estudio. Esto significa que la variedad de métodos es tan amplia y presenta tal capacidad de transformación que cualquier objetivo que uno se proponga debería poder ajustarse a cualquiera de estos tres métodos de investigación.

Para dejar claro cuáles son las diferencias existentes entre los tres tipos de diseño, en el libro *Diseños y métodos de investigación en traducción* aparece un párrafo donde se explica de forma bastante gráfica esta diferenciación:

sería posible colocar los tres tipos de diseño a lo largo de un continuo virtual que iría de los diseños puramente cualitativos (esto es, los basados únicamente en la introspección y capacidad analítica del investigador) a los puramente cuantitativos (es decir, los basados fundamentalmente en el análisis objetivo de datos cuantificables). Entre ambos extremos se encontraría toda una gama de diseños mixtos (que, como su propio nombre indica, serían una combinación de los dos anteriores) (Rojo, 2013, p.20).

En el caso de este estudio, se plantea una metodología basada en el uso de un corpus. Este diseño podría parecer en un primer momento un método bastante cuantitativo ya que “el uso del corpus proporciona una alternativa más sólida y empírica a la simple intuición o introspección del investigador” (Rojo, 2013, p.47). De hecho, y como aclara Rojo más adelante, “su objetivo más inmediato es cuantitativo, identificando regularidades y

patrones de uso mediante medidas de frecuencia y asociación” (Rojo, 2013, p.47). Sin embargo, es cierto que también podría considerarse un diseño cualitativo ya que, para terminar de concretar la investigación, es necesario ofrecer respuestas a esos resultados numéricos.

El propósito de haber escogido el uso de corpus es tener la posibilidad de examinar una gran cantidad de textos con la ayuda de herramientas informáticas para identificar las diferencias entre textos con la misma tipología y temática que hayan sido redactados en español de forma original y traducidos a esta misma lengua. Esto quiere decir que el objetivo del estudio es centrarse en si alguno de los dos grupos presenta rasgos definitorios, inclinaciones o tendencias a la hora de ser redactados, como podrían ser la explicitación, la normalización o la simplificación (Rojo, 2013).

7.2.Descripción del corpus

Para concretar todavía un poco más este método de investigación, del que ya se sabe que es mixto, es decir, una combinación de técnicas propias de un diseño cualitativo y de uno cuantitativo, y que, además, se hace a partir del uso de un corpus, es necesario explicitar los criterios que se han seguido para la creación de dicho corpus. Como se ha mencionado anteriormente, este corpus es un corpus textual, puesto que los documentos que lo componen son textos en formato escrito, y comparable, ya que contiene dos tipos de textos: por un lado, documentos redactados originalmente en español y, por otro, textos traducidos del inglés a esta misma lengua. Se trata de 83 textos que han sido redactados originalmente en español y 67 que han sido traducidos a este idioma. Esta diferencia se debe a que los textos redactados por entidades de habla hispana ofrecían una cantidad de información menor a los traducidos. A pesar de esta diferencia, el número de palabras de ambas partes del corpus es de 70.700 palabras aproximadamente. Para ser más exactos, son 70.754 en los traducidos y 70.731 en los originales, lo que supone una diferencia mínima de 23 palabras. Se ha decidido que sea igual el número de palabras en lugar de la cantidad de textos puesto que si una de las partes del corpus ofrecía más contenido y texto, la cifra de aquellos posibles problemas que se quieran identificar podría ser mayor en aquel grupo que contuviera un mayor número de palabras. Este corpus también se caracteriza por haber sido compuesto a partir de textos obtenidos de Internet, lo que lo convertiría en un corpus virtual. Estos mismos textos se recolectaron de forma completa, es decir, no se eliminó ningún extracto, y la temática es el cáncer colorrectal, por lo que podría considerarse un corpus especializado. No obstante, esta temática se trata desde un punto de vista bastante sencillo, por lo que el nivel de tecnicidad no es muy elevado, ya que los receptores de este tipo de textos son un público general no especializado. Otros aspectos que cabe destacar sobre los textos que engloban este corpus son la fecha en que fueron realizados, y es que esta es relativamente reciente ya que, al haber sido extraídos de páginas web cuya intención es proporcionar información actual sobre el cáncer colorrectal, se van actualizado cada poco tiempo, así como la calidad que presentan. Todos los textos escogidos han sido extraídos de fuentes fiables como hospitales, clínicas, aseguradoras u otras entidades relacionadas con el ámbito sanitario, por lo que la calidad de estas está demostrada.

7.3. Modelo de análisis

Tanto la traducción médica como la traducción de páginas web poseen una serie de dificultades o problemas a los que hay que prestar especial atención a la hora de llevar a cabo un encargo de estas características. No obstante, es muy común olvidar algunos de estos rasgos propios de este tipo de traducciones y que pasen desapercibidos por un lector que no se ha formado en esta disciplina. El objetivo de este estudio es demostrar que existen diferencias entre los textos médicos redactados originalmente en español y aquellos traducidos a esta misma lengua. Mediante este estudio, se pretende que, a partir del programa informático AntConc, del que se hablará detenidamente más adelante, se puedan encontrar con mayor facilidad algunos de estos elementos que aparecen en las traducciones que restan naturalidad a la lengua española como son los gerundios o la pasiva y, de este modo, comprobar si es cierto que su presencia en este tipo de traducciones es considerable o si, por el contrario, brilla por su ausencia. La idea de utilizar este modelo de análisis viene de De Felipe (2007). Esta investigadora realizó un estudio a partir del uso de un corpus comparable formado por textos jurídicos traducidos al español y redactados originalmente en español. Este trabajo pretendía demostrar la normalización de la lengua a partir del uso del gerundio y la voz pasiva. Como estos dos elementos son bastante comunes también en los textos médicos, ya que su presencia no se debe tanto al tipo de texto que se vaya a utilizar, sino a la lengua origen, se ha optado por utilizar el mismo modelo de análisis.

No obstante, basar solamente en dos elementos un estudio que trata de demostrar que un tipo de textos se produce de forma más natural y acaba considerándose más asequible para el lector que otro no parece un fundamento de peso. Es por esto que se escogieron además otros elementos que también han sido considerados problemáticos por otros autores en estudios previos como son la terminología, especialmente la convivencia de términos coloquiales con términos altamente especializados, la traducción de siglas, el uso de nombres de medicamentos o la forma de dirigirse al lector.

7.4. AntConc

AntConc es un programa informático diseñado principalmente para el análisis de grandes cantidades de texto o, más concretamente, de corpus. Esta aplicación resulta bastante útil ya que existen patrones difíciles de identificar a simple vista o tras una lectura rápida o que, de ser capaz de conseguirlos, es muy fácil pasar por alto algunos ejemplos. Este software dispone de una gran variedad de funciones que abarcan desde la creación de glosarios hasta la observación de palabras o grupos de palabras en sus propios contextos. Para ello, dispone de distintas herramientas que facilitan este análisis de los textos. La primera de ellas de la que se va a hablar es “Concordance”. Esta pestaña permite al usuario observar listas de aparición de la palabra que haya introducido en la barra de búsqueda acompañada del texto que la rodea. Asimismo, ofrece la fuente de donde procede dicha palabra, lo que resulta bastante útil si el corpus está compuesto por textos producidos por diferentes autores o entidades. “Concordance Plot” proporciona una visualización de la localización de la palabra introducida a lo largo de un texto. Las palabras se muestran como líneas negras dentro de un rectángulo y representan la altura a la que se encuentran dentro del archivo, que se presenta en forma de dicho rectángulo. Junto a esta se encuentra

“File View” que lo que proporciona es una muestra mucho mayor del contexto en que se encuentra la palabra buscada ya que se muestra el archivo completo. Mientras “Clusters” expone las palabras que acompañan con mayor frecuencia al término buscado, “Collocates” señala aquellas que tienen una mayor probabilidad de estarlo. “Word List” proporciona una lista de todas las palabras que componen el corpus por orden de frecuencia y “KeyWord List” realiza comparaciones entre dos corpus.

Otra función que se debe conocer de esta aplicación para poder entender ciertos caracteres que se utilizarán a lo largo de esta investigación es que existen distintos operadores que permiten buscar más de una palabra a la vez. Realizar una búsqueda que no esté restringida a un solo término, sino que permita la obtención de palabras con cierta relación morfosintáctica es una posibilidad que ofrece este programa. Para ello, existen una serie de caracteres que ayudan en la búsqueda de más de un término y que aparecen muy bien explicados en el artículo de Froehlich (2018). El asterisco (*) sustituye la posibilidad de que haya uno o más caracteres o incluso ninguno. El símbolo de suma (+) sustituye un carácter o ninguno. El símbolo de interrogación (?) sustituye un solo carácter. La diferencia entre la interrogación y el asterisco reside en que, si solamente se quisiera buscar un adjetivo en sus dos géneros como, por ejemplo, delgado y delgada, lo más útil sería utilizar una interrogación que solo permita que cambie un solo carácter y se introduciría “delgad?” en la barra de búsqueda. Por otro lado, si se introdujese en la palabra “dependiendo”, solamente se obtendrían los resultados de dicho verbo. No obstante, si lo que se pretende buscar son todos los gerundios que se encuentren en el texto, consultar un diccionario e introducir todos los verbos conjugados en gerundio no parece una idea eficaz. Para ello, el asterisco ayuda a encontrar todas aquellas palabras, independientemente del número de letras que le precedan, que acaben por “ndo” que es la terminación de los verbos en gerundio. Además de los operadores de búsqueda relacionados con caracteres, también existen aquellos con funciones para palabras completas. La arroba (@) se utiliza para hacer referencia a una palabra o a ninguna. La almohadilla (#) sirve para sustituir cualquier vocablo. Finalmente, la barra vertical (|) se utiliza para buscar dos palabras a la vez y, para ello, hay que introducir ambas en la barra de búsqueda separadas por dicho símbolo.

7.5. Análisis y resultados

7.5.1. Preparación de la aplicación AntConc

La aplicación AntConc solo funciona con textos en formato txt. Por ello, a la hora de recopilar los textos para la creación de este corpus, se fueron guardando en dicho formato. Al tratarse de un corpus comparable, es decir, formado por dos subcorpus, se empezó por introducir los textos escritos originalmente en español en el programa. Después, se abrió una ventana nueva de esta misma aplicación para añadir los textos traducidos. De este modo, se podría hacer el análisis de subcorpus por separado para hacer, posteriormente, una comparación.

Antes de empezar a analizar los textos que componen este corpus, se ha decidido utilizar una Stop List para eliminar todo aquel contenido vacío de significado. Esto quiere decir que se ha buscado una lista de palabras que no aporten ningún contenido al texto como

son las preposiciones o los determinantes para profundizar en aquellas que proporcionan un mayor significado, como son los sustantivos y los adjetivos. El objetivo de utilizar esta lista es que, al crear una lista de los términos más frecuentes, no aparezcan artículos ni preposiciones, dos tipos de palabras que se emplean con mucha frecuencia, pero que no aportan ningún valor a la investigación.

La Stop List escogida ha sido obtenida de un blog llamado Outbook (2010, en línea). Antes de introducirla en el programa, se ha comprobado que todas las entradas fueran palabras vacías, se han eliminado algunas de ellas que podían ser útiles en el contexto de este trabajo, como es el caso de ciertos verbos, y se han añadido otras, como es el caso de algunos determinantes que aparecían solamente en un solo género y número. Para introducirla en AntConc, ha sido necesario clicar sobre la opción “Tool Preferences”, entrar en la categoría “Word List” y bajo “Word List Range”, escoger la opción de “Use a stoplist below”. A continuación, solo hubo que apretar el botón “Open”, cargar la Stop List y presionar “Apply” antes de cerrar la ventana.

7.5.2. Gerundio

El gerundio es una construcción con valor adverbial que indica tanto la duración de un proceso como una acción anterior a la que expresa el verbo principal (Mendiluce Cabrera, 2002). A pesar de tener una definición bastante clara, esta forma verbal plantea muchas dificultades tanto en su uso general en español como en su traducción a esta misma lengua. Si se empieza por tratar el gerundio en español como tal, se puede confirmar que aparte de los usos mencionados anteriormente, existen muchos otros a los que los hispanohablantes están acostumbrados a oír o utilizar, pero que no son correctos. Mendiluce Cabrera (2002) menciona algunos de los más comunes como el “gerundio del Boletín Oficial del Estado”, el “gerundio especificativo”, el “gerundio relativo al complemento directo” y el “gerundio de posterioridad”. A esto añade que hay dos casos en los que algunos de estos errores estarían permitidos pese a no ser todavía una opinión unánime: “cuando expresa una acción inmediatamente posterior a la del verbo principal” o “cuando, pese a referirse al complemento directo, el gerundio es «hirviendo» o «ardiendo»” (Mendiluce Cabrera, 2002, p.74).

Si el gerundio en español como tal ya presenta problemas de uso, la traducción del gerundio inglés es todavía más complicada. Aunque el equivalente en inglés de un gerundio en español debería ser un verbo terminado en -ing, no funciona en el sentido contrario, es decir, no todas las palabras terminadas en -ing se traducen como un gerundio en español. El primer error consiste en traducir un verbo terminado en -ing como un gerundio en lugar de como un infinitivo. En inglés, el infinitivo que se conoce en español puede escribirse de dos formas: como un verbo en su forma base precedido por un *to* o como un verbo terminado en -ing. Esto variará dependiendo del tipo de palabra que le preceda. Por ejemplo, si el verbo tiene delante una preposición, se utilizará un verbo terminado en -ing. En cambio, si la palabra antesora es un adjetivo, le seguirá un infinitivo. Traducir un gerundio inglés por un gerundio español o un infinitivo también dependerá del verbo que tenga delante en la lengua inglesa. Verbos como *delay*, *enjoy* o *recommend* van seguidos de un gerundio mientras que *come*, *fail* o *request* preceden a un infinitivo. Asimismo, hay otros verbos a los que pueden ir seguidos de ambas opciones

ya sea porque el significado de la oración es el mismo, como ocurre con *love* o *begin*, o porque este varía, como ocurre con los verbos *forget* o *remember*. La cuestión es que los gerundios en inglés no siempre se traducen como gerundios en español. Por tanto, hay que tener en cuenta en qué situación de todas las mencionadas anteriormente se encuentran para decidir si se traduce como un gerundio o como un infinitivo.

Ahora que ya se han visto a grosso modo tanto los usos aceptados del gerundio como los incorrectos, se centrará este párrafo en un gerundio más específico de este estudio: el gerundio en el lenguaje médico. En medicina, el gerundio puede aparecer como sustituto de las oraciones de relativo, uso inadecuado ya que su función debería ser puramente adverbial. Al igual que los adverbios, presenta distintas categorías: modal, temporal o condicional, entre otras. Otro uso que suele darse al gerundio en medicina es aquel que conforma la perífrasis verbal estar + gerundio, o aquella que está formada por los verbos “continuar” o “seguir” seguidos de un gerundio. El gerundio ilativo o copulativo, también conocido como gerundio médico, es otro de los ejemplos que pueden verse con más frecuencia en este tipo de textos. Este es aquel que puede sustituirse fácilmente por una oración copulativa. Finalmente, el “destacando el hecho de que” y el “siendo” son también dos construcciones propias de la prosa médica. La conclusión a la que llega Mediluce Cabrera (2002) tras exponer todos estos ejemplos es que utilizar esta construcción es totalmente correcto ya que, tanto si tiene una función adverbial como si forma parte de una perífrasis verbal, su uso está aceptado. La cuestión es que, en la escritura científica, lo que empezó como un sustituto de conectores para evitar la repetición de estructuras se ha convertido en un patrón incorrecto que los escritores han ido adoptando al haberlas leído en otros artículos y haberse acostumbrado a ellas.

Como se ha mencionado anteriormente, el objetivo de este estudio es comprobar si el gerundio, entre otras cuestiones, se muestra con mayor frecuencia en los textos traducidos que en los textos originales o viceversa. Es por esto que se empezó la investigación introduciendo el subcorpus de textos originales en la aplicación AntConc. Después, en la pestaña llamada “Concordance”, se introdujo el grupo “*ndo” de modo que aparecieran todos los gerundios ordenados y en su contexto. El resultado de esta primera búsqueda fue de 410 palabras, de las cuales no todas eran gerundios debido a que los vocablos “cuando” y “segundo” también tienen esta terminación y no pueden considerarse verbos en gerundio. De este primer grupo de textos, se han conseguido extraer 93 verbos diferentes conjugados en gerundio, sin tener en cuenta la cantidad de veces que aparece cada uno de ellos. Por tanto, en este primer subcorpus se muestra una cantidad de 208 gerundios. La frecuencia con la que se utilizan cada uno de ellos varía bastante. Mientras que hay ciertos verbos que aparecen solamente una vez como “mejorando”, “palpando”, “detectando” o “provocando”, entre muchos otros, solo unos pocos se presentan más de 10 veces. Estos son “dependiendo” y “siendo”, que se manifiestan en 19 ocasiones cada uno. Anteriormente, ya se había mencionado lo común que es encontrar el verbo “ser” conjugado en gerundio en los textos médicos, por lo que no es de extrañar que sea uno de los más repetidos. Un ejemplo que muestra este último uso sería el siguiente:

Con relación a su localización, la mayoría de los tumores se localizan en el recto (37%), y el sigma (31%), *siendo* menos frecuentes en el colon ascendente (9%),

el ciego (8%), el colon descendente (5%), el colon transverso (4%), el ángulo hepático (4%) y el ángulo esplénico (2%).

También se han encontrado ejemplos del llamado gerundio médico:

Los programas de detección precoz del CCR permiten diagnosticar la enfermedad en sus primeras etapas, incluso antes de que aparezcan síntomas, *mejorando* el pronóstico de la enfermedad y la calidad de vida de los afectados y sus familias.

El gerundio aparecido en la oración anterior se considera un gerundio médico porque podría sustituirse perfectamente por una oración copulativa, es decir, por un “y mejorar”. Otra forma de comprobar que se trata del gerundio ya mencionado es a partir de otras partículas parecidas como “por lo que”, “así” o “pues”. Esto permite verificar que se trata de dos oraciones que tienen el mismo rango. Aunque este tipo de ejemplos es de los más comunes en la prosa médica, debería evitarse en la medida de lo posible y sustituirse por una construcción distinta que incluya uno de los enlaces expuestos anteriormente.

Otra muestra de uno de los diferentes usos que ofrece el gerundio es el siguiente:

En tumores más avanzados, fundamentalmente por afectación de los ganglios linfáticos de alrededor del mismo (estadio III), el pronóstico empeora, aunque la probabilidad de supervivencia continúa *siendo* superior al 50%.

Esta oración contiene un gerundio en forma de perífrasis, construcción que se ve repetida en distintas ocasiones a lo largo de este subcorpus. En este caso, se trata de una perífrasis formada a partir del verbo “continuar” junto con el verbo en gerundio. Este uso es uno de los pocos usos que están permitidos dentro de la gran variedad que ofrece esta forma verbal. De hecho, podría decirse que el empleo de gerundio en perífrasis es uno de los más frecuentes en este subcorpus. El siguiente caso muestra otro ejemplo de perífrasis, esta vez construida a partir del verbo “estar”:

Existe la posibilidad de que no se detecte el cáncer si éste no estaba *sangrando* cuando se realizó la prueba de Sangre Oculta en Heces.

Finalmente, se ha añadido un ejemplo de gerundio con valor adverbial, en este caso condicional:

Afortunadamente, gran parte de los factores de riesgo pueden ser evitados *adoptando* un estilo de vida saludable.

De los diferentes usos que presenta el gerundio en este subcorpus, se han querido mostrar distintas categorías, tanto de usos correctos como incorrectos. Es por esto que se han ofrecido ejemplos de gerundio médico, del verbo “ser” en gerundio, del gerundio en perífrasis y del gerundio con valor adverbial. No obstante, es importante recalcar que, de los distintos usos de gerundios que tienen lugar en este subcorpus, más de la mitad se consideran adecuados. De los usos correctos, abundan los gerundios con valor adverbial modal y las perífrasis mientras que de los usos incorrectos predomina el gerundio médico.

A continuación, se realizará el mismo procedimiento con el subcorpus de textos traducidos. Si se vuelve a introducir el grupo “*ndo” en la pestaña “Concordance”, se puede apreciar un número menor de resultados: 342. Aun así, no es la cifra exacta de gerundios puesto que también se han añadido otro tipo de palabras terminadas en “-ndo” como “cuando” o “fondo”. Por tanto, el número total de gerundios es de 164, una cantidad menor a la de los textos originalmente redactados en español, lo que contradice la hipótesis de que las traducciones tendrían un número mayor de gerundios. En cuanto a la frecuencia de los verbos, que son 71, ocurre lo mismo que con el subcorpus anterior. Se pueden encontrar varios verbos que aparecen solo una vez y otros de los que se hace uso hasta 23 veces, como es el caso del “dependiendo”. Solamente hay otro verbo, “incluyendo”, del que se hace uso más de 10 veces, un total de 21.

Como se decía anteriormente, la perífrasis es uno de los mayores usos que se le puede dar a un gerundio. Por ello, no es de extrañar que se hayan encontrado varios ejemplos en este subcorpus como el siguiente:

Los científicos están *buscando* las causas y las formas de prevenir el cáncer colorrectal, mejores maneras para detectarlo temprano (cuando es más fácil de tratar), así como las formas de mejorar los tratamientos.

También aparecen bastantes gerundios con valor adverbial, aunque en este caso con una función modal:

La criocirugía destruye el tumor mediante congelación con una sonda de metal. La sonda es guiada por la piel hasta alcanzar el tumor *usando* una ecografía.

Otro uso bastante común es el de un “incluyendo” que introduzca una serie de ejemplos:

Se han probado diferentes métodos en estudios clínicos, *incluyendo*: probar nuevos medicamentos de quimioterapia o medicamentos que ya se estaban usando contra otros cánceres.

Finalmente, se ofrece una muestra de gerundio ilativo o médico:

Podrías ser que también se recomiende más cirugía si el pólipo no se pudo extirpar completamente o si tuvo que ser extraído en muchos fragmentos, *dificultando* ver si había células cancerosas en los bordes.

Para comprobar que se ha identificado el gerundio correctamente, se vuelven a utilizar las técnicas presentadas en el análisis de gerundios del subcorpus anterior, es decir, a través de conectores como “y” o “con lo que”, con el “así” que enfatiza más el valor del gerundio o asegurando que ambas oraciones están al mismo nivel.

Por tanto, este análisis verifica que el subcorpus de textos traducidos también hace uso del gerundio de distintas formas. Si se hace un recuento de estos, una buena parte serían usos correctos, predominando el gerundio con valor adverbial modal y el gerundio en forma de perífrasis. Por otro lado, en cuanto a usos que es preferible evitar, se halla el gerundio médico y el uso del “incluyendo” para preceder ejemplos.

Los resultados obtenidos de este primer análisis muestran que, mientras se creía que el uso frecuente del gerundio podría deberse a traducciones incorrectas provenientes del inglés y que, por este mismo motivo, la cifra sería mayor en los textos traducidos, se ha podido comprobar que los textos escritos originalmente en español contienen una cantidad más elevada. A diferencia del inglés, el español es una lengua que trata de ser lo menos repetitiva posible ya sea con respecto a términos o a estructuras. Es por esto que la razón por la que los textos originales contienen más gerundios se deba seguramente a que lo que empezó como una forma de evitar ciertas construcciones que podían parecer menos formales para este tipo de textos como son las oraciones copulativas, han provocado que el gerundio se haya ganado un papel importante en el lenguaje técnico. En la actualidad, aquellos profesionales que se dedican a la prosa médica han ido imitando la forma de expresarse de sus antecesores en este campo, lo que ha causado la adopción de ciertas estructuras o construcciones cuyo uso no es del todo adecuado.

7.5.3. Voz Pasiva

La voz pasiva es aquella forma de lenguaje que indica que el sujeto de la oración está siendo afectado por el verbo principal, al contrario de lo que ocurre en la voz activa donde el sujeto realiza la acción que indica dicho verbo. El español tiene dos formas pasivas que se conocen como perifrástica y pronominal (Navarro, Hernández, & Rodríguez-Villanueva, 1994). La perifrástica es aquella formada por el verbo ser con su conjugación correspondiente y el verbo principal en participio mientras que la pronominal, también conocida como refleja, es aquella utilizada mediante un “se”.

Al igual que ocurría con el gerundio, esta forma de expresión no siempre se utiliza de forma correcta. Aunque es cierto que la pasiva perifrástica no tiene ninguna complicación, es lógico que, al tener un uso mucho menos frecuente, el número de errores al utilizarla no sea muy elevado. En el caso de las oraciones pasivas reflejas ocurre todo lo contrario. Al formarse a partir de un “se”, partícula que tiene otras funciones como la de complemento indirecto, la de pronombre personal con valor reflexivo o recíproco, la de pronombre con valor impersonal o la de morfema derivativo, es decir, que sirve para crear verbos a partir de otros que ya existen (Liroz, 2010), es muy común confundir el tipo de oración que se está tratando y no ser capaz de concordar los elementos que la componen de forma correcta. Uno de los usos indebidos de este tipo de oraciones es aquel que tiene lugar cuando intenta utilizarse un complemento de persona. Este tipo de oraciones solo puede construirse en tercera persona y tienen como función apelar a objetos o acciones de modo que cualquier ejemplo que incluya a una persona podría considerarse errónea. Navarro, Hernández, & Rodríguez-Villanueva explican esto a partir de un muy buen ejemplo:

Cuando el sujeto de la pasiva es una persona (*se vacunaron las embarazadas*), nace ambigüedad a causa del valor reflexivo o recíproco de *se*, de modo que no puede emplearse la pasiva refleja. Ante una frase como la que comentamos, el lector entendería que las embarazadas se vacunaron “a sí mismas” (Navarro, Hernández, & Rodríguez-Villanueva, 1994, p.102).

No obstante, no es el único ejemplo de uso incorrecto que puede darse en la pasiva refleja. Es muy común confundir el tipo de oración que se está tratando, lo que influye a la hora de concordar en número verbo y complementos. Las oraciones impersonales solo pueden conjugarse en tercera de singular ya que no poseen ninguna clase de sujeto, pero esto no implica que no puedan contener complementos referidos a personas. Es por esta misma razón que es bastante frecuente cometer el error de conjugar el verbo en plural al identificar un grupo de personas en la oración. Esto se entiende mucho mejor a partir del ejemplo siguiente: “se distribuyó a los pacientes en dos grupos” (Navarro, Hernández, & Rodríguez-Villanueva, 1994, p.102). Conjugar el verbo de otra forma implicaría un mal uso de esta oración impersonal. De hecho, utilizar este verbo en plural podría implicar un valor reflexivo de este “se”, lo que expresaría que los pacientes se distribuyen a sí mismos. Si se escoge otra oración impersonal y se omite cualquier tipo de complemento que haga referencia a una persona, se puede apreciar mejor por qué es un error pluralizar el verbo: “se trabaja poco en este hospital” (Navarro, Hernández, & Rodríguez-Villanueva, 1998, p.102).

Existen algunos elementos de la lengua que aportan al discurso ciertas características. Es un hecho que la voz pasiva ayuda a un texto a parecer más objetivo. Es por esto que abundan en los textos científico-técnicos, especialmente en los médicos. Aunque es cierto que en español la pasiva perifrástica no es tan común y utilizarla en exceso puede llegar a sonar poco natural, otras lenguas de las que se traduce al castellano sí que la anteponen ante la voz activa. Esto implica que el traductor tenga que estar muy alerta a la hora de transferir las oraciones pasivas puesto que, aunque su uso no es incorrecto, sí que muestra

cierta artificialidad. De hecho, son varios los autores que recalcan lo “incómodo” que puede llegar a ser leer un texto en español repleto de pasivas, lo que ha causado que se ofrezcan distintas propuestas que eviten una traducción poco fluida. Tres de las más sencillas consistirían en transformar la oración pasiva perifrástica en una oración activa, intercambiar una oración pasiva perifrástica por una pasiva pronominal o reformular la oración de modo que la voz pasiva desaparezca (Infomed, 2016).

Una vez tratados los dos tipos de oraciones pasivas que existen y su problema en las traducciones al español, se comprobará si los textos del corpus recopilado confirman la hipótesis de que las traducciones contienen una cantidad mayor de pasivas, sobre todo perifrásticas, que los textos originales. En este caso, la búsqueda de pasivas va a resultar más complicada que la búsqueda de gerundios puesto que las pasivas pueden encontrarse conjugadas en distintos tiempos verbales. Si se empieza por introducir en el subcorpus de textos traducidos “es *do”, se puede apreciar que aparecen 67 resultados. El objetivo de esto era encontrar todas las pasivas perifrásticas en presente de indicativo. No obstante, y como ocurría con los gerundios, no todos los resultados son válidos; solo 12 de estos son pasivas. Si se inserta “sido *do” en la barra de búsqueda, se pueden encontrar 4 resultados más. Introduciendo las partículas “fue *do” aparecen otros 3. Aplicando el futuro simple, “será *do”, se halla 1 nuevo resultado. Esto implica que aparecen solo 20 resultados de pasiva perifrástica en total. Es curioso que todos los resultados provengan de 4 fuentes distintas, pero que la mayoría de ellas se encuentren en los diferentes textos proporcionados por American Cancer Society. Esto explica que probablemente el resto de traductores estén concienciados sobre esta cuestión y que hayan olvidado algún ejemplo o hayan creído que en esa ocasión la oración pasiva perifrástica era la mejor opción. Por el contrario, el traductor o grupo de traductores que se encargó de los textos publicados por American Cancer Society ha hecho un uso mucho más repetitivo de este elemento por motivos que se desconocen.

Si se realiza una búsqueda de las pasivas reflejas que contiene este texto a partir del pronombre “se”, se obtiene una cifra de 1.401 resultados, de los que un elevado porcentaje son pasivas. El resto se consideran otro tipo de muestras de los diferentes usos que tiene el “se” y que se habían mencionado anteriormente, como el “se” impersonal o el pronombre personal que acompaña al usted. En cuanto a usos erróneos de pasiva refleja, no se ha encontrado ninguno.

A continuación, se mostrarán algunos ejemplos de los dos tipos de pasivas que se han tratado hasta el momento.

El catéter *es guiado* hacia la arteria hepática en el hígado a través de un pequeño corte en la piel del lado interno del muslo.

El riesgo es aún mayor si el familiar *fue diagnosticado* con cáncer cuando tenía menos de 45 años, o si más de un familiar de primer grado *es afectado*.

Puede ser mortal cuando *es diagnosticado* en las etapas finales, pero si *es descubierto* de forma temprana, a menudo puede *curarse*.

Primero, *se realiza* una cirugía para extirpar tanto cáncer como sea posible en el abdomen (vientre).

El primer ejemplo parece una de las mejores opciones de traducción de entre las tres muestras ofrecidas. Si se quisiera utilizar una oración activa como uno de los diferentes cambios que se proponían anteriormente para evitar la pasiva, habría que añadir un sujeto como el/la médico o el/la enfermero/a, dependiendo del profesional que se encargue de realizar dicha intervención. Una oración pasiva refleja tampoco sería una mala opción, es decir, cambiar el “el catéter es guiado” por “se guía el catéter”. Realizar un cambio de la oración podría resultar complicado en este contexto ya que podría derivar en la pérdida de algún matiz o de algún elemento importante de la oración, pero es otra de las opciones. Aun así, y como se decía anteriormente, no es que se trate de una construcción incorrecta, de hecho, en este caso no suena del todo artificial, pero varios autores recomiendan evitarla.

El segundo ejemplo contiene dos verbos en pasiva perifrástica. El primer caso es bastante similar a la primera oración que se ha analizado. Es un ejemplo de pasiva que, aunque puede modificarse, no suena artificial. No obstante, el segundo verbo en pasiva sí que sería conveniente cambiarlo puesto que el orden de la oración es un calco de la estructura inglesa. En este caso, una oración activa ayudaría al lector a comprender el sentido de la frase de forma mucho más rápida: “o si ha afectado a más de un familiar de primer grado”.

En el tercer ejemplo se pueden apreciar dos pasivas perifrásticas y una pronominal. En este caso, lo más probable es que se haya optado por alternar ambos tipos para no sonar muy repetitivo. Sin embargo, no debería ser una opción utilizar una estructura menos común o natural para una traducción solamente para evitar la repetición ya que no se trata de palabras que, si se usan constantemente, pueden hastiar al lector. Se está lidiando con un tipo de voz que, si no se quiere repetir, se puede omitir a través de un cambio en la estructura de la oración.

El cuarto párrafo presenta una oración pasiva refleja de las que podrían llegar a confundirse con una oración reflexiva. En este caso, solo hace falta ver el contexto y tener un poco de sentido común para verificar que la oración expresa que un médico es quien realiza la cirugía. No obstante, si se tratara de un verbo distinto, podría haber la posibilidad de que el lector confundiese una oración pasiva refleja con una oración reflexiva. Entre los distintos tipos de pasiva, este es el más habitual en este subcorpus.

Si se realiza la misma búsqueda en el subcorpus de textos originales, hay 10 resultados de pasiva perifrástica en presente de indicativo de los 16 resultados encontrados al introducir las partículas “es *do” en la pestaña de “Concordance”. Mientras que no aparece ningún resultado tras la búsqueda de la pasiva en pasado simple, se han encontrado 7 ejemplos tras introducir “sido *do”. En futuro aparecen otros 2 ejemplos más, lo que resultaría en un total de 19 pasivas perifrásticas. En este caso, las fuentes son

mucho más variables que en el subcorpus anterior ya que los 10 resultados de pasiva en presente provienen de 8 fuentes distintas, lo que indica que ya no se trata de un estilo de redacción de un traductor en concreto, sino que cada uno de los diferentes autores ha creído adecuado utilizar dichas pasivas perifrásticas en sus textos. Si se busca la pasiva pronominal a partir del pronombre “se”, aparecen 1.240 resultados, de los cuales más de la mitad son pasivas.

Como se ha hecho con el subcorpus de textos traducidos, se insertarán varios ejemplos de oraciones en voz pasiva a continuación.

La muestra, posteriormente, *se entrega* en el centro de salud y si el resultado es positivo, el usuario *es citado* por el médico de familia para continuar con el procedimiento diagnóstico.

CMED aún estas especialidades y cada caso *es discutido* para decidir el abordaje terapéutico más conveniente.

Si el cáncer *se encuentra* en un pólipo que *ha sido extirpado* durante una colonoscopia, es posible que sólo sea necesario acudir a revisiones médicas periódicas.

De estos tres ejemplos, el primero y el tercero son bastante similares. El autor utiliza primero una pasiva refleja y, probablemente para que no resulte repetitivo, emplea de forma posterior una pasiva perifrástica. Aunque ya se ha comentado que la pasiva perifrástica no es muy usual en español, es correcta y, en estos dos casos, se entiende el significado de la oración. Por otro lado, está la segunda oración que, aunque no tiene ninguna pasiva refleja en su interior, hace uso también de una pasiva perifrástica. En este caso vuelve a ocurrir lo mismo que con las otras dos oraciones: aunque existen otras opciones más comunes y frecuentes, no se trata de una oración ni incorrecta ni sin sentido, por lo que es totalmente válida.

Los resultados obtenidos de los análisis de los dos subcorpus indican que ambos tipos de textos siguen haciendo uso de oraciones pasivas, tanto perifrásticas como reflejas. En los dos casos, la presencia tanto de un tipo como de otro es bastante parecida ya que presentan unas cifras muy similares. Tales números indican que la pasiva perifrástica está muy por debajo de la refleja, algo bastante normal si se tiene en cuenta que la pasiva perifrástica es muy poco común en el habla española. No obstante, se puede llegar a la conclusión de que su uso se deba a la desesperación de los autores o traductores por utilizar estructuras variadas y no resultar repetitivos ya que el español es una lengua que, en términos de redacción, trata de cambiar constantemente de estructuras o de utilizar el mayor número de sinónimos posible para demostrar la riqueza que posee. Otro motivo que ayudaría a entender por qué se hace uso de la pasiva perifrástica en estos textos es el mismo por el que la prefieren tanto en otras lenguas: para dar importancia al que sería el complemento

directo en la oración activa. La cuestión es que, en la mayoría de casos, su uso está aceptado y no supone ningún problema de sentido. Aun así, un lector de español preferirá la voz activa o la pasiva refleja puesto que ambas presentan estructuras mucho más naturales y, por ende, podrían considerarse opciones más adecuadas. Esto explica por qué la pasiva pronominal muestra unas cifras de uso tan elevadas en comparación con la perifrástica.

7.5.4. Terminología

Al tratar en el marco teórico la traducción médica y las dificultades que esta implica, se pudo comprobar que la terminología era una de las mayores preocupaciones a las que debía prestar atención el traductor. Aunque se comentaba que, al tratarse de *Fact Sheet for Patients*, cuyo objetivo es el de informar a un público general sobre ciertas nociones médicas, el hecho de que sean redactados por profesionales de la salud puede implicar en cierto modo la convivencia de términos muy especializados con la de otros mucho más familiarizados por el lego. Es por este mismo motivo que en este apartado primero se va a realizar, a partir de la aplicación AntConc, una lista de los términos más utilizados en ambos subcorpus para ver si coinciden. Después, se comprobará si existe alguna diferencia entre los términos que se usan para tratar el cáncer colorrectal según si se trata de un texto original o de una traducción y si se encuentra una falta de coherencia al combinar términos más coloquiales con otros más propios del lenguaje especializado.

Anteriormente, ya se había comentado que se había introducido una Stop List que eliminara todas aquellas palabras vacías de contenido. Esto se había hecho con la intención de que, al querer una lista de los términos más frecuentes en ambos corpus, no aparecieran ciertas palabras con un uso bastante elevado, pero que no aportan nada a esta lista. De no añadir una Stop List, preposiciones como “de”, conjunciones como “y” u “o” y determinantes como “el” o “la” encabezarían la lista y complicarían la interpretación de los datos que se quieren obtener realmente.

El análisis terminológico se comenzó por el subcorpus de textos originales. Para poder obtener la lista de términos frecuentes fue necesario utilizar la herramienta “Word List”. Para que la aplicación comenzase a ordenar las palabras, había que clicar sobre el botón “Start”. A pesar de que se mostró una lista de 5.540 palabras, se centró la investigación solamente en los términos con un mayor uso en este subcorpus. “Cáncer”, “colon” y “colorrectal” encabezaban esta lista con un número de 1.076, 772 y 387 veces respectivamente. Tiene lógica que estos sean los más usados ya que el tema principal es el cáncer colorrectal. Lo más probable es que “colon” estuviese por delante de “colorrectal” debido a dos razones complementarias. La primera es que, si el texto trata sobre el cáncer colorrectal, es decir, una enfermedad que afecta tanto al colon como al recto, la palabra colon se utilice a menudo ya sea para describir dónde se encuentra el tumor, cómo afecta a esta parte del cuerpo, si tiene que describirse un tratamiento quirúrgico, cómo influye al colon o incluso para nombrar al cáncer en sí, entre otras posibilidades. El segundo motivo por el que la palabra “colon” sea tan frecuente se puede deber a que muchas entidades prefieren tratar este cáncer como dos enfermedades independientes. Al resultar el cáncer de colon más habitual que el cáncer de recto, es usual que los organismos médicos le presten algo más de atención. No obstante, se trata también

como un solo tipo de cáncer a través del término “cáncer colorrectal” porque, al ser dos partes del cuerpo que están unidas, es probable que la enfermedad les afecte simultáneamente.

A partir de la herramienta “Clusters/N-grams” es posible comprobar que se utiliza más el “cáncer de colon” (381) por encima de “cáncer colorrectal” (346) como conjunto. El cáncer de recto en comparación con el cáncer de colon tiene una importancia mucho menor ya que el término como tal aparece en solo 22 ocasiones. En la herramienta “Concordance” se ha podido observar que se utilizaba CCR como sinónimo de “cáncer colorrectal”. Por eso, se han buscado las siglas directamente para comprobar si era más propio de este tipo de textos utilizar el término como tal o abreviado en tres letras. Los resultados de CCR en todo el subcorpus han sido de 28 ejemplos, aunque hay que añadir que no provienen todos de la misma fuente, sino que son 6 entidades distintas las que lo usan. Esto resulta también interesante puesto que muchos autores suelen recurrir a las siglas, hasta el punto de crearlas, para eludir los términos de dos o más palabras que acaban ocupando un espacio bastante importante como ocurre con los textos que se emplean para páginas web o folletos. No obstante, parece que no ha sido la opción más aclamada por los autores de los textos.

Tras identificar cuáles eran los términos más frecuentes en los textos originales, se realizó el mismo proceso con el subcorpus de textos traducidos. Cuando la aplicación creó la misma lista de términos que aparecen con mayor frecuencia, los resultados fueron los mismos en cuanto a los tres términos más utilizados: cáncer, colon y colorrectal. Si se presta atención solamente a las cifras, puede observarse un cambio bastante elevado. La palabra “cáncer” aparecía 1.813 veces, casi 800 veces más que en el subcorpus anterior. “Colon” obtenía 728 resultados y “colorrectal” 693. El término abreviado de cáncer colorrectal que usaban los autores de textos originales, CCR, no aparece en ningún momento en el subcorpus de textos traducidos. Esto indica que los autores anglosajones no han creído necesario crear ninguna sigla para sustituir este término y tampoco lo han visto necesario los traductores. En mi opinión, no es algo estrictamente imprescindible ya que se trata de un término bastante corto y no resulta tan cansado de leer como podría ocurrir con otros términos compuestos por cuatro o cinco palabras. No obstante, al haberse comentado en el análisis del subcorpus anterior, he creído necesario buscarlo y hacer un breve comentario al respecto.

Si se sigue con el análisis de la frecuencia de uso de términos se puede comprobar que, mientras que la palabra “colon” consigue prácticamente las mismas cifras que en el otro subcorpus, la palabra “colorrectal” aumenta a más de 300 resultados de más. Una explicación que se le puede dar a esto es que, al ser el inglés una lengua que tiende a realizar oraciones mucho más cortas y sencillas, la repetición del sujeto deba hacerse más a menudo y, por tanto, deba repetirse la palabra “cáncer” de forma más constante. Además, el inglés es una lengua que no ve problemas en utilizar un mismo término o palabra una y otra vez, por mucho que esta aparezca de forma muy seguida. En cambio, el español tiende a buscar sinónimos, utilizar pronombres o, como bien se comentaba anteriormente, a alargar oraciones y a hacerlas mucho más complejas con tal de no tener que repetir un mismo término. Como traductores, intentar evitar una palabra de un texto médico a partir del uso de un sinónimo puede conllevar ciertos problemas si no se trata de un sinónimo total. Es por esto que, antes de hacerlo, deben tenerse muy claros los

significados de las palabras que se estén utilizando y por las cuales se quieran sustituir. Asimismo, siempre existe la opción de repetir el término las veces necesarias como puede haber ocurrido en los textos traducidos, un hecho que podría haberse evitado si el traductor lo hubiese creído adecuado. No obstante, parece que este ha llegado a la conclusión de que repetir los términos constantemente en lugar de buscar oraciones más diferentes entre sí, pero a la vez enrevesadas, era una mejor opción para unos textos que van dirigidos a un lego y cuya función es la de informar, no la de embellecer el lenguaje.

Si se vuelve a utilizar la herramienta “Clusters/N-grams”, se puede comprobar que en el subcorpus de traducciones se emplea con mayor frecuencia “cáncer colorrectal” (680) que “cáncer de colon” (307). Si se observa la cantidad de usos de “cáncer de recto” se puede observar que está muy por debajo del cáncer de colon con una cifra de solamente 59 resultados. Esto demuestra que los redactores de *Fact Sheet for Patients* tanto en lengua española como en lengua inglesa prestan una menor atención al cáncer de recto que al de colon.

La lista de frecuencia de usos de los términos que componen este corpus va a servir también para comprobar uno de los problemas que se trataba en el marco teórico que podrían presentar los *Fact Sheet for Patients*: la combinación de términos coloquiales con otros de mayor grado de especialidad. Para ello, hay que tener en cuenta que hay ciertos términos que se han introducido en las conversaciones cotidianas y que, por ese mismo motivo, podría llegar a parecer términos coloquiales. Por ejemplo, el cáncer es una de las enfermedades que más preocupa a la población ya que, aparte de afectar a una gran parte de esta, no se ha encontrado una cura definitiva. El hecho de que las personas estén tan familiarizadas con este término y que pueda tratarse el concepto en el día a día no le resta su valor especializado.

El subcorpus de textos originales posee una lista de 5.540 palabras. Aunque no toda la lista está formada por términos, sí que ayuda a discernir entre aquellos más coloquiales y aquellos más especializados, en caso de haberlos. Tras observar detenidamente esta lista, puede comprobarse que no se ha utilizado ningún término coloquial que favorezca la comprensión del texto. Un ejemplo de esto sería emplear la palabra “anginas” como sustituta de “amigdalitis”. En este caso, todos los términos, por muy habituales que puedan ser, son términos pertenecientes a la jerga médica. En todo caso, sí que se puede percibir un mayor uso de términos de un alto grado de especialidad que podrían dificultar la comprensión de un lector que no conozca del todo bien la temática. “Neoplasia”, “adenocarcinoma”, “exéresis”, “dehiscencia” o “RAS” son ciertos términos que podrían complicar la lectura fluida del lector. Lo interesante de esto es que, aunque cada uno de sus significados es muy particular, podría buscarse alguna alternativa más sencilla. Por ejemplo, si se busca la definición de “dehiscencia”, se trata de una “abertura espontánea de una zona que se había suturado durante una intervención quirúrgica” (WordReference, en línea). Utilizar una corta explicación a partir de esta definición sería probablemente una idea bastante más acertada si se tiene en cuenta que solo se pretende informar a los pacientes que quieren conocer la enfermedad y no formarles en medicina. Lo mismo ocurre con el término “neoplasia”, que se utiliza para hablar una “formación patológica de un tejido cuyos componentes sustituyen a los de los tejidos normales” (WordReference, en línea) o “una masa anormal de tejido provocada porque las células que lo constituyen se multiplican a un ritmo superior al normal” (Educalingo, en línea).

Esto significa que podría utilizarse como sinónimo de cáncer o tumor, dos términos a los que un lego está bastante más acostumbrado. Ejemplos como “RAS” son mucho más complicados de sustituir por un término conocido, pero sí que puede añadirse alguna explicación posterior sobre lo que es, técnica que se ha utilizado para tratar otros términos como “regorafenib”:

El regorafenib, un fármaco oral que inhibe diferentes vías de crecimiento tumoral, está aprobado en 3ª línea o posteriores en monoterapia, pero produce importantes efectos secundarios, por lo que se están estudiando nuevas dosis para mejorar el perfil de toxicidad.

Se podría deducir que el uso de ciertos términos como los que se han mostrado previamente, seguramente desconocidos por gran parte del público, se deba a que los redactores son profesionales del sector que se han habituado de tal forma a estos vocablos que dan por hecho que los lectores los conocerán también. No obstante, no se han encontrado términos coloquiales que se hayan querido utilizar con el objetivo de facilitar la comprensión del texto, como se había supuesto a partir de las nociones teóricas obtenidas.

A continuación, se observará también en la misma lista de términos, en este caso de los textos traducidos, si es posible encontrar términos coloquiales que convivan con otros especializados en los mismos escritos. Al igual que ocurría con el subcorpus de textos originales, no se ha hallado ningún término demasiado coloquial para el tipo de texto que se está tratando. No obstante, tampoco se han encontrado términos con un grado de especialización demasiado elevado. Aunque sí que es cierto que se han identificado ciertos términos que podrían suponer alguna dificultad de comprensión al lector, se ha solventado el posible desconocimiento. Por un lado, se han encontrado términos sin explicaciones posteriores, pero que tampoco son necesarias porque el sufijo que los acompaña ya muestra a qué está haciendo referencia. Por ejemplo, en este subcorpus aparece con bastante frecuencia la palabra “sigmoidoscopia”. Si bien no le sigue una definición clara de lo que es en la mayoría de textos, se sabe que el sufijo “-copia” se utiliza para hacer referencia a un procedimiento médico, ya que existen otros más comunes que también poseen esta terminación como, por ejemplo, “colonoscopia” o “endoscopia”. Además de esto, los autores de estos textos facilitan bastante la comprensión del escrito al lector puesto que sitúan este tipo de términos en contextos donde es muy fácil adivinar a qué se están refiriendo y puede verse en varios ejemplos:

Están disponibles otras varias técnicas de examen: analizar muestras de materia fecal en busca de sangre, realizar una *sigmoidoscopia* para ver el tercio inferior del colon, o usar una prueba radiológica tal como un enema de bario o una colonografía TC.

Sigmoidoscopia: se introduce una cámara diminuta con un tubo de plástico flexible en el recto, lo que ofrece una visión del recto y la porción inferior del colon.

Otro procedimiento diagnóstico utilizado para el cribado es la *sigmoidoscopia* (examen de la porción inferior del intestino grueso con un tubo de visualización, ver Endoscopia).

Si ninguna de las ideas anteriores puede aplicarse a dicho término, este siempre se encuentra dentro de un párrafo titulado “tipos de pruebas” o algún ejemplo similar que determina que el lector está frente a un procedimiento que ayudará al médico a diagnosticar al paciente.

Otros ejemplos que se quieren resaltar de términos asequibles para un lego gracias al contexto, a una explicación o a su terminación son el de “polipectomía” o el de “adenocarcinoma”. Aunque en varios casos se explica a qué se está haciendo referencia, también se trata de términos que, al tener unas terminaciones muy características como son “ectomía” o “carcinoma”, el lector puede deducir que se trata de un procedimiento médico, en este caso a una técnica endoscópica, o a un tipo de tumor.

De todos modos, existen ciertos términos médicos que no poseen estas terminaciones que ayudan al lector a identificarlos y de los que se hace un uso frecuente en estos textos. Por ello, los autores, que prefieren que sus lectores no se vean obligados a realizar búsquedas frecuentes en Internet debido a la incompreensión de ciertos términos, añaden ellos mismos sus propias definiciones. A continuación, se muestra un ejemplo:

Durante la prueba, usted coloca una muestra de heces en una tarjeta de prueba recubierta con una sustancia de origen vegetal llamada *guayacol*.

No obstante, tampoco se ha podido encontrar una gran cantidad de ejemplos que pruebe este rasgo ya que la mayoría de términos son totalmente asequibles para el lector y, de no ser así, contienen estos sufijos tan comunes en la terminología médica que ayudan a comprender la clase de concepto que se está tratando.

7.5.5. Siglas

La sigla es una palabra formada por el conjunto de iniciales de una expresión compleja que puede provenir de cualquier lengua (Villoria, en línea). Dos ejemplos extraídos del blog Proyecto Aula (en línea) podrían ser ONU (Organización de Naciones Unidas) o CDROM (Compact Disc Read Only Memory). Existen otros elementos similares a las siglas con los que no hay que confundirlas. Los acrónimos, por ejemplo, son palabras creadas a partir de la unión de dos o más palabras, cuyo nombre se forma a partir del principio de la primera palabra y el final de la última u otro tipo de combinaciones del

estilo. Un ejemplo de acrónimo podría ser ofimática (oficina informática) (Villoria, en línea). Las abreviaturas son otro ejemplo de palabra distinta a la sigla. Estas se caracterizan por ser abreviaciones de palabras que se realizan a partir de la eliminación de letras y que suelen llevar un punto al final. Ejemplos de esto podrían ser Dir. (director), SS. MM. (Sus Majestades) o íd. (ídem).

Las siglas, al igual que cualquier palabra o concepto lingüístico, tiene sus propias normas y es de vital importancia conocerlas para cumplirlas tanto como autor como como traductor. Para empezar, es primordial recordar que las siglas no llevan tilde, van en mayúscula y no requieren ningún punto como las abreviaturas. Esto puede apreciarse en los ejemplos mostrados anteriormente. No obstante, existen ciertas siglas que se han lexicalizado de tal manera que ya no se consideran siglas, sino palabras. Estas sí que irían en minúscula y con las tildes necesarias. Asimismo, las siglas suelen ir acompañadas de un artículo y este debe poseer el mismo género que la primera palabra que lo compone. Otro aspecto importante que cabe tener en cuenta es que el número no se señala a partir de una “s” final, sino que se marca a través del artículo que le preceda. El último factor que se debe conocer para poder analizar con totalidad las siglas de este corpus es que no pueden dividirse a no ser que se hayan lexicalizado.

Si se sigue analizando la lista de términos, se puede apreciar la presencia de una gran cantidad de siglas. Como en el subcorpus de textos originales no era un factor muy común como para dedicarle cierta atención en esta investigación, no se ha hecho ningún comentario al respecto, aparte del CCR como alternativa a cáncer colorrectal. No obstante, al comprobar la lista de términos de subcorpus de textos traducidos, ha ocurrido todo lo contrario. Las siglas iban apareciendo cada varios términos y se ha querido comprobar a qué hacían referencia para que se les estuviera dando tal importancia. Tras buscar en su contexto cada una de ellas, se ha podido constatar que hacen referencia tanto a conceptos médicos como a instituciones sanitarias. Un aspecto que llama la atención de estas siglas es que suelen ir acompañadas de su significado. No obstante, parece que no se llega a un acuerdo de si dicho contenido debería dejarse en español o en inglés, o de si debería señalarse el idioma en el que se ha creado la sigla, aunque este hecho puede comprobarse dependiendo de si encaja o no con la traducción propuesta del término.

El Physician Data Query (*PDQ*) es la base de datos integral del Instituto Nacional del Cáncer (*NCI*).

Los síndromes hereditarios más comunes asociados con el cáncer colorrectal son el síndrome de Lynch (cáncer colorrectal hereditario sin poliposis o *HNPCC*) y la poliposis adenomatosa familiar (*FAP*, siglas en inglés), pero otros síndromes menos comunes también pueden incrementar el riesgo de cáncer colorrectal.

La prueba de sangre oculta en heces (Fecal Occult Blood Test, *FOBT*) se usa para detectar sangre en las heces o la materia fecal.

Prueba anual de sangre oculta en las heces fecales (*FOBT*) o prueba inmunoquímica fecal (*FIT*) una vez cada 12 meses.

A partir de estos cuatro ejemplos, se puede constatar que, aunque es común utilizar siglas en inglés y añadir de dónde provienen después, parece que no está muy claro cuál es la mejor técnica. En el primer ejemplo, es lógico no traducir el nombre de la *Physician Data Query* y omitir el nombre en inglés, si no existe tal traducción de forma oficial ya que el lector podría estar interesado en esta y no hallar resultados al introducir el término en español. No obstante, no sería una mala opción añadir una traducción al nombre oficial para que el lector conozca a qué se dedica tal institución. Por el contrario, como sí que existe el término Instituto Nacional del Cáncer como traducción oficial de la *National Cancer Institution*, es mejor utilizar la opción en español por la misma razón: si el lector está interesado en informarse sobre esta entidad, al introducir el nombre en español, obtendrá más información en su lengua materna.

Si se analiza el segundo ejemplo, se pueden apreciar dos nuevas muestras de siglas con sus dos opciones de traducción, que difieren la una de la otra. Mientras que, en el primer ejemplo, es decir, el de HNPCC, el autor añade la sigla como sinónimo del término a partir de la conjunción “o”, más adelante prefiere utilizar las siglas entre paréntesis, aclarando que es el conjunto de letras que se utiliza en inglés para hacer referencia a tal término. Aunque ambas opciones son válidas ya que explican el contenido de las siglas, la mejor traducción debería seguir una coherencia a lo largo de todo el texto. Es decir, si se opta por una opción, seguir con ella durante todo el proceso de traducción. De no hacerlo, el lector puede llegar a confundirse al encontrar tantas formas distintas de tratar las siglas.

Aunque la traducción del tercer ejemplo es la más larga, podría ser una de las mejores propuestas de todas las que se han extraído. Esto se debe a que aparecería la traducción en español, que ya indica al lector de qué se está hablando, y, como quieren utilizarse siglas desde ese instante en adelante para no utilizar el término completo, se aclara de dónde proceden para no provocar confusiones, todo ello dentro de un mismo paréntesis.

El cuarto ejemplo sería similar al ejemplo uno, pero se expresa de forma más sencilla. Es decir, se deja el término en español y se añaden entre paréntesis las siglas en inglés para utilizarlas a partir de ese momento a lo largo del texto y evitar utilizar el nombre completo. El único problema que puede presentar esta traducción, como se ha comentado anteriormente, es que quizá sea algo confuso para el lector no poder comprobar de dónde provienen dichas siglas. Aunque no sea necesario indicar de qué lengua provienen como aparecía en el segundo ejemplo, estaría muy bien seguir la técnica del ejemplo tres ya que podría ser la más completa.

Si se procede a un análisis de siglas general, se puede concluir que no se han encontrado ninguno de los errores mencionados al principio de esta sección en todo el subcorpus de textos traducidos. El hecho de que la mayoría de siglas provengan del inglés no ha supuesto ningún problema para el traductor que, aunque en muchos casos solamente las colocaba entre paréntesis junto al significado de dicha sigla, ha sabido mantener las mayúsculas en todos los casos, ha omitido cualquier punto y, en el caso de utilizar

artículos, ha sabido concordarlos con la palabra principal de las siglas. A continuación, se muestran dos ejemplos que conservan las mayúsculas, no utilizan puntos al final como sí debe hacerse con las abreviaciones y mantienen la concordancia con el artículo si es necesario.

Después de la FOBT, podrá reanudar a sus actividades habituales.

Estos genes ayudan normalmente a reparar el ADN que se ha dañado.

7.5.6. Nombres de medicamentos

Otro aspecto que se quiere analizar de este corpus de textos sanitarios sobre cáncer colorrectal es la forma en que se nombran los fármacos o medicamentos. Para comenzar, es necesario hacer una distinción entre ambos términos ya que no significan lo mismo. Como bien se indica en el blog de traducción Trusted Translations,

un fármaco es una sustancia con actividad farmacológica, es decir, químicamente activa que tiene un efecto sobre el organismo; y el medicamento sería la forma comercial de éste, es decir, el medicamento no sólo sería el principio activo sino también la sustancia final que se ofrece comercialmente, lo que incluye excipientes, la presentación del producto (en polvo, grageas...), [etc.] (Trusted Translations, 2010, en línea).

Por tanto, la sustancia que realiza un cambio en el cuerpo de la persona que lo ingiere sería el fármaco mientras que la forma en que es vendida al público, ya que no podría tomarse si no se adapta al paciente o consumidor, sería el medicamento.

Una vez conocida esta diferencia, que ya sirve de ayuda al traductor incluso para no cometer ningún error tratando de utilizarlas como sinónimos, se verán las distintas formas que existen para hacer referencia a un medicamento o fármaco. La primera es a partir de la denominación común internacional (DCI) que recomienda la OMS. La DCI es el nombre genérico oficial atribuido a un fármaco (Academia Europea de Pacientes, en línea). Esta normalización ayuda a no confundir sustancias que puedan tener diferentes nombres de marca. Otra forma de apelar a un medicamento es a partir de su nombre comercial o nombre registrado, es decir, al nombre atribuido por la empresa que lo ha creado o a su marca. Finalmente, existen otras denominaciones comunes que no tienen ninguna validez internacional como pueden ser el nombre químico, una sigla, un código de investigación, un nombre erróneo o una denominación común nacional (Navarro, 2003). El hecho de que existan tres formas distintas presenta ciertos problemas de traducción. Aunque encontrar una traducción para la primera opción es una cuestión bastante fácil de resolver puesto que solo consistiría en buscar su equivalente al español

en la base de datos en línea de la OMS, la segunda y la tercera forma suponen mayores dificultades. A pesar de que encontrar una marca registrada es una tarea sencilla ya que suelen ir acompañadas del símbolo (®) o con mayúscula inicial, a veces hay que buscar una marca que sea equivalente en la lengua y cultura meta o incluso sustituirla por el nombre genérico, si no se halla ninguna o se cree que así se evitarán confusiones. La cuestión aquí es buscar un fármaco o medicamento equivalente que contenga los mismos componentes. Finalmente, averiguar una traducción para una denominación local sería la opción más complicada de resolver puesto que estas no aparecen de forma tan frecuente en los recursos de consulta habituales (Trusted Translations, 2010).

Además de esto, existen otros problemas que surgen a raíz de la traducción de medicamentos y fármacos. El primero de ellos es que cada lengua y cultura tiene una forma más común de tratarlos y es necesario mantener dicha forma para realizar una traducción adecuada. Esto quiere decir que, si en un país están acostumbrados a tratar los medicamentos a partir de sus nombres genéricos, al lector de tal lugar le costará reconocerlo si se utiliza el nombre comercial. Un ejemplo de esto es que el español tiende a utilizar los nombres comerciales mientras que el inglés prefiere los nombres genéricos. Por lo tanto, habría que considerar esta adaptación cultural si se quiere realizar una traducción lo más natural y comprensible posible para el receptor de la lengua y cultura meta. Otro problema que supone la aparición de nombres de medicamentos en un texto es el número de reglas ortográficas que los envuelven. Para utilizar correctamente una denominación común es necesario escribirla en minúscula y precedida por un artículo (Navarro, 2002). Es fundamental conocer esta información ya que la influencia del inglés ha contribuido a la aparición de nombres de fármacos sin artículo. No obstante, las marcas registradas o nombres comerciales siguen otras normas: deben aparecer con mayúscula inicial y sin ir acompañado de ningún artículo (Navarro, 2003a). Otro problema que surge a raíz de un incorrecto uso de mayúsculas y minúsculas es que, en ocasiones, son los textos originales en inglés los que utilizan de forma errónea estos términos, lo que puede llegar a confundir a un traductor, haciéndole creer que se encuentra frente a un nombre comercial cuando se trata del nombre genérico. Navarro explicaba esto con un buen ejemplo:

ante una frase como «children with Ampicillin rash are highly unlikely to develop an immediate or accelerated reaction in the future», muchos traductores pueden llegar a pensar, de forma equivocada, que *Ampicillin* es una marca comercial (invariable), cuando en realidad se trata de una INN, que tiene su correspondiente DCI en español: *ampicilina* (Navarro, 2003b, en línea).

Para hallar estos nombres de medicamentos o fármacos en el corpus, se acudió a la opción “Word List” de la aplicación AntConc, donde se buscaron los nombres de medicamentos que más se habían utilizado, primero en el subcorpus de textos originales, para observar de qué forma había querido plasmar el autor en lengua española los fármacos o medicamentos que el paciente debe tomar como tratamiento contra el cáncer colorrectal. Tras realizar las búsquedas pertinentes, se llegó a la conclusión de que este primer subcorpus prefiere el uso de nombres genéricos. Algunos ejemplos de esto podrían ser

los términos “cetuximab”, “panitumumab” y “regorafenib”. Estos tres vocablos son nombres genéricos de tres fármacos distintos. Se localizaron sus respectivos nombres comerciales, “Erbix”, “Vectibix” y “Stivarga”, y se introdujeron en la opción de “Concordance” para comprobar si se habían utilizado en algún momento. No apareció ningún resultado con ninguno de ellos por lo que se concluyó que los autores en lengua española de *Fact Sheet for Patients* sobre cáncer colorrectal preferían hacer uso de nombres genéricos de medicamentos antes que de nombres comerciales. Una de las razones a las que puede deberse esta decisión es que, al no ser fármacos conocidos ni habituales en la vida diaria de un paciente, como sí lo podría ser Dalsy, por ejemplo, no se les concede tanta importancia al uso de un nombre u otro.

Además de esto, se ha querido comprobar que no se estuviese cometiendo ningún error con respecto a las mayúsculas o minúsculas. Aunque en un primer momento, tras realizar la búsqueda de ciertos términos, ha aparecido uno de ellos en algunas ocasiones con mayúscula inicial y en otras en minúscula, se han querido observar estos ejemplos en su contexto para averiguar la razón de esta desigualdad, sobre todo teniendo en cuenta que una misma fuente utilizaba las dos versiones. Tras pinchar sobre los términos que aparecen con mayúscula en la lista, se ha comprobado que esta variedad se debe a que se utilizan como inicio de oración o tras un punto como puede verse en el siguiente ejemplo:

Tratamiento personalizado con anticuerpos anti-EGFR

Cetuximab y panitumumab son anticuerpos monoclonales cuya diana molecular es un receptor situado en la superficie de las células denominado factor de crecimiento epidérmico (EGFR).

En cuanto al tema de los artículos, se ha podido comprobar que, en algunas ocasiones, como en los ejemplos mencionados hasta el momento, los nombres genéricos no van precedidos de ningún artículo, un error de acuerdo con Navarro (2002). No obstante, también se han encontrado muestras, aunque una menor cantidad en comparación con la omisión del artículo, en las que los autores sí que lo utilizan, como en los ejemplos de la siguiente oración:

Los principales fármacos aprobados en España son las fluoropirimidinas (5-Fluorouracilo, Capecitabina, UFT y Utefos), *el oxaliplatino*, *el irinotecan*, *el raltitrexed*, *el cetuximab* (tumores RAS no mutado), *el panitumumab* (tumores RAS no mutado), *el bevacizumab*, *el aflibercept* y recientemente *el regorafenib*.

Lo que llama la atención de esta cuestión es que los autores no siguen una coherencia con respecto a este aspecto. De nuevo, se ha podido demostrar que, en un mismo texto, es posible encontrar ejemplos de nombres genéricos de medicamentos precedidos por artículos y sin artículo ninguno.

Tras comprobar que los textos redactados originalmente en español utilizaban nombres genéricos para tratar los medicamentos, se quiso averiguar si los textos traducidos a este mismo idioma seguirían la misma línea. En un primer momento, se especulaba que estos textos también utilizarían los nombres genéricos y dejarían de lado los nombres comerciales, pero tras buscarlos en la herramienta “Concordance” dentro de sus contextos, los resultados fueron inesperados. A continuación, se mostrarán varios ejemplos:

Se pueden usar medicamentos que atacan el EGFR para tratar algunos cánceres de colon o de recto en etapas avanzadas. Entre estos se incluye:

Cetuximab (Erbix)

Panitumumab (Vectibix)

Una de las combinaciones anteriores más un medicamento que ataca el VEGF (bevacizumab [Avastin], ziv-aflibercept [Zaltrap] o ramucirumab [Cyramza]) o uno que ataca el EGFR (cetuximab [Erbix] o panitumumab [Vectibix])

5-FU y leucovorín por sí solos o junto con un medicamento de terapia dirigida

Capecitabina, sola o con un medicamento de terapia dirigida

Irinotecán, solo o con un medicamento de terapia dirigida

Cetuximab solo

Panitumumab solo

(Stivarga) solo

Regorafenib Trifluridina y tipiracil (Lonsurf)

El primer ejemplo expuesto contiene dos de los medicamentos que se usaron como muestra del subcorpus de textos originales. La diferencia reside en que el traductor ha decidido mantener la decisión del redactor en lengua inglesa de añadir el nombre comercial junto al nombre genérico del medicamento. Esto permite al lector estar más formado y mejor informado sobre el tratamiento que va a seguir. Un hecho sorprendente sobre esto es que se pensaba que sería decisión del traductor la de añadir ambos nombres, pero se ha comprobado que esta son las formas que ofrece el texto original.

En el segundo ejemplo ocurre lo mismo. Se ha utilizado este ejemplo y no cualquier otro debido a la cantidad de medicamentos que se mencionan y al hecho de que algunos de ellos contengan corchetes. Los corchetes suelen ser una opción que utilizan los traductores para señalar que la información que se encuentra dentro ha sido añadida por ellos. Sin embargo, en este caso se ha podido comprobar que su uso se debe a que el autor sentía la necesidad de utilizar unos paréntesis dentro de otros paréntesis y los corchetes eran la mejor forma de diferenciar cuando empieza y acaba cada uno.

Asimismo, se ha querido prestar especial atención a las mayúsculas y minúsculas para constatar que no hubiera ningún error ortográfico. En este caso, se puede observar que los nombres genéricos aparecen en ocasiones con mayúscula inicial, pero esto se debe solamente a que se encuentran al principio de una oración. Puede confirmarse que es debido a esta razón y no por falta de conocimientos porque cuando el nombre genérico de un medicamento no encabeza ninguna línea aparece en minúsculas.

Finalmente, se ha querido comprobar si se utilizaban los artículos delante de los nombres genéricos, algo que no sería necesario en ejemplos como los anteriores en los que solamente se mencionan los nombres en una lista. No obstante, se han buscado estos nombres de medicamentos dentro de una oración completa y se ha demostrado que sí que van precedidos por artículos, como ocurre en el ejemplo siguiente:

Los problemas que *el panitumumab* causa en la piel pueden ser más graves y podrían ocasionar descamación de la piel.

Aun así, son muy escasas las veces en las que los nombres de medicamentos se utilizan en oraciones completas. Los medicamentos mencionados en este subcorpus tienden a aparecer enumerados en listas, un hecho que no permite a esta investigación afirmar que en el cien por cien de los casos estos términos se empleen de forma correcta.

Como conclusión a todo esto se puede decir que, mientras que los autores de habla española han optado por mencionar solamente los nombres genéricos de los medicamentos, los redactores de habla inglesa han preferido mostrar dos de las opciones disponibles de llamar a dicho medicamento. Aunque es cierto que puede que a ciertos pacientes no les interese mucho si el medicamento que van a tomar se llama de una forma u otra, el hecho de proporcionar toda la información que pueda ser útil da confianza y tranquilidad al lector. Además, se ha demostrado también que los traductores, a pesar de no tener por qué ser profesionales de la salud, tienen conocimientos sobre las normas ortográficas que envuelven a la denominación de medicamentos y fármacos, al contrario que los autores en lengua española, que no acaban de tener muy claro si deben utilizar artículos que precedan a los nombres genéricos o no.

7.5.7. Tratamiento del lector

A diferencia del inglés, el español posee dos formas de dirigirse a una segunda persona del singular: a través del “tú” o del “usted”. El uso de una u otra dependerá del grado de formalidad de la situación, de la persona a la que se quiere transmitir el mensaje o del grado de confianza con esa persona, entre otras opciones. Es decir, el “usted” posee un mayor grado de neutralidad y muestra cortesía, formalidad y distancia entre el emisor y el receptor. En cambio, el “tú” implica familiaridad, cercanía y confianza (Tomasena, en línea). Mayor Serrano (2008) señala en un estudio sobre folletos, una tipología textual bastante similar al género escogido para la elaboración de este corpus, varios ejemplos. Esta autora indica que, si los textos van dirigidos a ancianos, se recomienda evitar el lenguaje autoritario e imperativo y tratar al lector de “usted”. En cambio, en aquellos

redactados para niños o adolescentes, es preferible utilizar el pronombre personal “tú”. Tomasena, por su lado, en su artículo sobre la mejor forma de dirigirse a un receptor en páginas web, señala que lo importante es conectar con el usuario e incitarle a actuar y esto es algo que se logra independientemente del uso del “usted” o del “tú”. A esto aclara que existen tres factores que influyen en esta decisión: el público, lo que implica edad, género, procedencia y nivel económico, la imagen que se quiera transmitir o el producto o servicio que se ofrezca. A esto añade que, una vez tomada la decisión, tiene que mantenerse en todos los aspectos y no solo a través de la página web. Esto quiere decir que, si pretende darse una imagen de empresa joven e informal, un usuario no puede acudir a tal empresa de forma presencial y ser tratado de usted porque resultaría incoherente con la imagen que pretendía mostrarse a través de la página.

Las dos formas personales mencionadas anteriormente no son las únicas que permiten a un emisor dirigirse a un receptor. El “nosotros” inclusivo también se ha utilizado en numerosas ocasiones para hacer sentir al lector la empatía del autor, como si conociera lo que este está pasando. Una última opción para dirigirse al público es a través de una tercera persona, ya sea en singular o en plural, que puede verse reflejada a partir de un término que no tiene porqué ser un pronombre personal; podría ser “el paciente”, “los familiares” o cualquier otro modo de dirigirse al lector. De todas estas opciones, Mayor Serrano (2008) indica que el “tú”, el “usted” y el “nosotros” son tres formas bastante aceptables para este tipo de textos puesto que ayudan a que el lector se implique en el contenido de los escritos. Por este mismo motivo, recomienda la elusión de cualquier tercera persona que suponga una distancia entre público y autor.

Para comprobar cuál es la forma más común entre los textos que componen el corpus agrupado para este estudio, se empezó por el subcorpus de textos originales y se buscaron las diferentes formas propuestas anteriormente, es decir, el “tú”, el “usted”, el “nosotros” y el “paciente”. El problema que supone el español para poder identificar cuál de estas opciones es la más utilizada es que permite la omisión del sujeto. Esto quiere decir que en numerosas ocasiones podrá verse el verbo conjugado en la persona correspondiente sin necesidad de un sujeto que lo acompañe. Esta cuestión dificultó la búsqueda de frecuencia de usos de estas formas, pero se consiguió una respuesta final. Para empezar, se comenzó introduciendo el “tú” en la herramienta “Concordance” y no se obtuvo ningún resultado. Del “usted”, por el contrario, se obtuvieron 41 ejemplos y de su abreviatura, “ud.”, otros 2. Es importante comprobar que estos no provengan todos de la misma fuente porque entonces sería un solo autor el que estaría utilizando esta opción. En este caso, se confirmó que procedían de 9 entidades distintas. Como el uso de un pronombre personal no confirmaba en español la frecuencia de usos de cada forma porque el español suele omitirlas, se buscó otro tipo de pronombres que ayudaran en la búsqueda de oraciones dirigidas a un “tú” o un “usted”. El “le” por ejemplo, del que aparecen 86 resultados vuelve a indicar la preferencia por el “usted”, ya que si se busca un “te” los ejemplos obtenidos se reducen a 15. El “su” muestra 323 usos, del que solo una parte son una forma relacionada con el “usted”. El “tu” aparece en 31 ocasiones y el “ti” proporciona otros 4 resultados más, lo que sigue siendo una cifra bastante reducida en comparación con la cantidad de ejemplos extraídos a partir del “usted”. Además de esto, se buscó en la lista de términos los verbos que se utilizaban con más frecuencia para ver la forma en que estaban conjugados. El “debe” era uno de los más usados, con 102 resultados. Esta es una de las formas que se utilizaría con el “usted”. Aun así, tras leer todos y cada uno de los

ejemplos, no todos ellos hacen referencia a un “usted” sino que en algunos casos hacen referencia a oraciones impersonales o pasivas reflejas. El verbo “consulte” aparece en 37 ocasiones e implica de nuevo el frecuente uso del “usted”. Por otro lado, el primer verbo que aparece en la lista de frecuencia de usos conjugado en segunda persona del singular para tú es “recibirás”, en el puesto 1.362, con solo 5 usos, lo que ya ayuda a concluir que el resto de verbos conjugados en esta misma persona y este mismo número serán iguales o menores a cinco, lo que implica de nuevo un empleo bastante reducido en comparación con el “usted”.

Mientras se buscaban verbos conjugados en segunda persona de singular (tú), se encontró el verbo “tengo” con una frecuencia de 16 usos. Esto permitió comprobar que la primera persona del singular también se utiliza, sobre todo en títulos a modo de pregunta, para resolver dudas que puedan tener los lectores. Utilizar una primera persona ayuda también a mostrar a la entidad cercana y confiable.

El verbo “podemos” es el primero en aparecer en las listas de frecuencia de usos conjugado en primera persona del plural, en el número 997, con unos resultados de 7 casos, mientras que el “nosotros” aparece en solo tres ocasiones. El “nuestro” muestra 13 ejemplos más, lo que implica la presencia de la primera persona del plural para referirse a la entidad u organización como tal y no con la intención de ponerse en el papel del paciente. Es importante resaltar esto puesto que así se descarta esta opción como modo de dirigirse al receptor.

Finalmente, se encuentra la opción de tercera persona tanto en singular como en plural. La palabra “paciente” aparece en 98 ocasiones y “pacientes”, en 162. Este término se utiliza para describir cualquier aspecto del cáncer colorrectal del mismo modo que se utilizaría una segunda persona de singular, pero en este caso en tercera. Tanto el “él” como el “ella” tienen también algunos resultados, pero ninguno de ellos se refiere al paciente como tal, sino al médico, a un procedimiento o a cualquier instrumental u objeto que tenga relación con el tema a tratar. Las cifras obtenidas de los verbos conjugados en tercera persona, tanto en singular como en plural, no son de gran ayuda puesto que se refieren a cualquier aspecto que se esté tratando (factores, cáncer, síndromes, tabaco, síntomas, entre otros) por lo que es muy difícil discernir cuáles de ellos afectan al lector como paciente.

De todo este análisis, se puede concluir que las formas más comunes son “usted” y “el paciente”, aunque también se combinan con otras formas como la primera persona del singular, es decir, el “yo”. Algunas entidades sustituyen ese “usted” por un “tú”, seguramente porque les parezca una forma mucho más cercana al paciente. Por último, la primera persona del plural, el “nosotros”, es otro pronombre personal que se utiliza a menudo para referirse a la empresa, la entidad o la organización que se encarga de emitir los escritos, pero nunca para hablar del paciente afectado por la enfermedad.

Si se procede a realizar el mismo análisis del subcorpus de textos traducidos, se puede comprobar que de nuevo el pronombre “tú” se utiliza en menor medida que el “usted”. Del primero solo aparecen dos muestras de la misma fuente mientras que del segundo se obtienen 155 resultados. Para acabar de comprobar si los 155 son los únicos, se introduce en la barra de búsqueda la versión acortada de dicho pronombre, “ud.”, ya que en el subcorpus anterior se habían extraído varios resultados. En este caso no se obtuvo

ninguno. El “tu” como determinante aparece en más ocasiones, 47 exactamente, pero siempre en los textos emitidos por la misma fuente. El “su” presenta 390 ejemplos, pero no todos están relacionados con el “usted”, por lo que no se podría usar dicha cifra para asegurar que aparece en más ocasiones que el “tú”. Si se analiza la lista de frecuencia de usos de los términos, se encuentran desde el principio varios verbos cuyo sujeto debería ser usted como el “consulte”, con 39 ejemplos, “refiérase”, con 27 y “hable” con 25, entre otros. Encontrar un verbo conjugado para un tú ha sido bastante más complicado ya que el primero aparece en el puesto 778 con 8 usos. Esto implica que el resto de verbos conjugados de esta misma forma tendrá un máximo de 8 ejemplos, por lo que se puede demostrar que el “usted” tiene una mayor presencia que el “tú” también en este subcorpus.

Si se busca el pronombre personal en primera persona del plural, es decir, el “nosotros”, se encuentran cuatro resultados procedentes de la misma fuente y con la misma intención que aparecían en el subcorpus anterior: la de hablar de la propia entidad. Eso significa que, aunque en los textos traducidos se utilice en alguna ocasión, no es muy frecuente su uso. Si se prueba a introducir el pronombre “nuestro”, aparecen cinco resultados nuevos, esta vez procedentes de tres fuentes distintas, lo que aumenta la frecuencia de uso de este pronombre personal. El “nuestra” se utiliza en una sola ocasión y el “nos” en dos más. Por otro lado, si se introduce el pronombre de la primera persona del singular, “yo”, se podrá apreciar un solo resultado en el que se trata de reproducir una frase emitida por una persona en un vídeo anterior al texto, por lo que se confirma que, en un primer momento, se podría creer que su uso tampoco es muy común en estos textos traducidos. No obstante, para confirmar este hecho, se introdujeron más pronombres, como se ha hecho en los casos anteriores como el “me”, “mi” o “mí”, extrayendo así 20, 24 y 4 ejemplos respectivamente. Además, más adelante, cuando se realizaba la búsqueda de los verbos con mayor frecuencia de uso, se pudieron encontrar “puedo” y “debo” en puestos bastante elevados, con 19 y 18 usos respectivamente. Esto significa que, aunque no se use siempre, sí que es una opción por la que suelen optar los redactores cuando quieren responder a ciertas dudas que puedan tener los pacientes. Esta es una técnica que se utiliza para parecer más cercanos y para atraer la atención del lector.

La tercera persona del singular en forma de “paciente” aparece también en varios casos, 43 para ser exactos, y otros 62 si se busca su forma en plural. Como ocurría en el subcorpus anterior, parece ser una de las opciones más comunes. El “él” aparece en cuatro ocasiones, pero solo dos de ellas hacen referencia a un paciente. Con el pronombre “ella” pasa lo mismo, de tres ejemplos, solo dos de ellos son válidos y hacen referencia a un paciente. Aun así, los verbos conjugados en tercera persona, tanto del plural como del singular, son los más comunes, con frecuencias de uso de hasta 99 casos. Estos verbos no siempre hacen referencia a los pacientes, por lo que no pueden considerarse todos los ejemplos como válidos. De todos modos, esta forma de dirigirse al lector, junto con el usted, son las opciones por las que suelen decantarse los autores tanto en lengua española como en lengua inglesa.

Antes de resumir las conclusiones que se han obtenido de este análisis y terminar esta sección, se quería hacer una breve mención a una cuestión relacionada con este aspecto que se considera de gran importancia. Mientras se analizaba la lista de frecuencia de usos en busca de verbos que mostrarán si el “tú” era un pronombre habitual en este tipo de textos, se pinchó sobre la palabra “consulta”, por si existía algún caso en el que se utilizase

dicho sustantivo también como verbo en imperativo. Entre esos mismos resultados se encontró que una misma fuente utilizaba dicha palabra como verbo, pero también como sustantivo. El problema de esto se halla en que, en algunas ocasiones en las que la usaba como sustantivo, se dirigía al lector como “usted”.

Solicite una Consulta en Mayo Clinic

Si notas cualquier síntoma persistente que te preocupe, *consulta* con el médico.

Ten en cuenta cualquier restricción previa a la consulta. Cuando *programes* la consulta, *asegúrate* de preguntar si hay algo que *debas* hacer con anticipación, como restringir tu dieta.

Aunque en la mayoría de los casos, el texto que contiene el primer ejemplo se dirige al público como “tú”, el verbo “solicite” aparece en repetidas ocasiones. El problema que esto presenta es que, si se tratase solamente de un caso aislado, sería simplemente una pequeña errata, algo que tiene fácil solución si la entidad programa tareas de revisión cada cierto tiempo. Sin embargo, se trata de una oración que se usa reiteradamente y que está creando una mala imagen de la empresa ya que, al final, se muestra una falta de coherencia textual importante que confunde al lector. A pesar de que las herramientas de traducción asistida son de gran ayuda, puede que en este caso hayan perjudicado al traductor. Está claro que un traductor debe leer el texto, analizarlo, identificar a quién va dirigido, cuál es su objetivo y con qué intención ha sido creado. A partir de ahí, debe tomar una serie de decisiones que se verán reflejadas en la traducción. Una de ellas debería haber sido de qué forma tratar a los receptores. Una posible razón que podría haber influido en esta falta de coherencia podría ser que dicho traductor tomase una decisión que cambiase en otro momento y, al introducir la herramienta de traducción esa oración que ya había sido traducida anteriormente del mismo modo, se mantuviese el error a lo largo de todo el texto. Aun así, esto podría haberse corregido con una revisión, ya que no es una falta que aparece en solo una ocasión, sino que se repite en tres casos distintos.

En conclusión, este análisis ha demostrado que cada traductor opta por el pronombre personal que le parece más adecuado para el texto. Más concretamente se podría concluir que el “usted” y la tercera persona tanto en singular como en plural a través del sustantivo “paciente” o “pacientes” son los más utilizados, pero eso no quiere decir que los demás pronombres tengan un uso mucho menor en comparación. De hecho, se ha observado que algunas entidades optan por una opción para el texto en general, pero utilizan otras con otras funciones. Por ejemplo, si pretenden resolver dudas que pueda tener un paciente, las redactan desde la primera persona del singular. Así, atraen la atención del lector y parecen más implicados en la causa. O también se hallan ejemplos de primera persona del plural para referirse a la entidad. Como se mencionaba anteriormente, esto no quiere decir que todos los textos estén redactados a partir de estas últimas personas ya que solo son técnicas que se aplican a ciertas partes del texto para obtener mejores resultados. También

se ha detectado un ejemplo de combinación de personas que, en ese caso, sí resultaría una falta de coherencia al no diferenciar entre el “tú” y el “usted”, dos opciones entre las que sí que hay que tomar una decisión clara para la redacción de un texto.

8. Conclusiones

El primer objetivo de este estudio era ser capaz de realizar un corpus comparable de textos médicos sobre cáncer colorrectal. El segundo consistía en confirmar si existía alguna diferencia entre textos redactados originalmente en español y textos con la misma función (informar) y misma temática (el cáncer colorrectal) traducidos del inglés a esta misma lengua, teniendo en cuenta que la hipótesis planteada sugería que los textos traducidos presentarían más problemas que los originales. Tras el análisis realizado anteriormente, se puede afirmar que sí que existen diferencias en algunos aspectos. Por un lado, los textos redactados originalmente en español poseen una terminología mucho más variada. Esto implica un uso de vocablos sencillos y asequibles y a la vez de términos bastante más complicados sin ir acompañados de ninguna explicación o definición. Sin embargo, los textos traducidos poseen una terminología mucho más comprensible, con un grado de especialidad bastante reducido, pero a su vez bastante específica para la temática que se está tratando. A pesar de que, en estos casos, los términos no son muy especializados, el redactor en lengua inglesa especifica a qué se refiere un concepto si cree que no va a ser comprendido por el lego, algo que aparece reflejado en las traducciones al español. Esta parece una técnica bastante efectiva si se tiene en cuenta que se están tratando textos a los que solamente accederán personas con pocos conocimientos sobre el tema cuyo objetivo sea informarse sobre la enfermedad. Un especialista que tenga como objetivo formarse más sobre un tema especializado en concreto acudirá a estudios mucho más elaborados y no a textos con información básica. Por tanto, hacer de esta lectura una experiencia sencilla y agradable ayudará al receptor a sentirse más cómodo y seguro a pesar de la situación que le está tocando experimentar.

En relación con la terminología, también se observó qué términos se utilizaban con mayor frecuencia. Ambos subcorpus obtuvieron los mismos resultados a pesar de que se mencionó que los textos traducidos tenían unas cifras de uso de dichos términos mucho más elevadas que los textos originales. La explicación más coherente podría ser que los textos traducidos imitan la costumbre anglosajona de utilizar los términos constantemente sin la necesidad de buscar sinónimos que confundan al lector y compliquen la comprensión de su lectura. Esto, en cambio, no parece una preocupación de los autores en lengua castellana que, como ya se comentaba anteriormente, hacen todo lo posible para conseguir un discurso elaborado, aunque esto suponga arriesgar el correcto uso de la lengua.

La diferencia que aparece entre ambos subcorpus a raíz del tema de las siglas se debe a que los textos redactados originalmente en español no tienen la necesidad de explicar de dónde provienen ciertas siglas puesto que hacen referencia a conceptos en español ya conocidos por el lector. Este hecho cambiaría si el autor se viese obligado a mencionar alguna noción proveniente de otro país en el que se hable otra lengua, pero no es el caso. Sin embargo, en el subcorpus de textos traducidos sí que se puede comprobar que el traductor siente la necesidad de explicar el significado o el origen de las siglas que utiliza ya que asume que no son conocidas por el receptor. La cuestión aquí es que no existe

ninguna técnica establecida para explicar el significado de las siglas. Cada traductor opta por una forma distinta, ninguna de las cuales podría considerarse incorrecta. No obstante, sí que debería decidir basándose en los puntos a favor que posea cada una de ellas.

Para empezar, si el texto en inglés presenta unas siglas, a excepción de aquellas bastante populares como OTAN o VIH que ya poseen sus propias traducciones oficiales al español, lo ideal es añadir su significado en esta lengua. Esta es una de las técnicas que más se ha repetido a lo largo del subcorpus. El problema que surge a raíz de utilizar solamente una traducción junto a las siglas es que el receptor observará que las palabras añadidas por el traductor no concuerdan con las letras de las siglas. Las traducciones están al alcance de cualquier persona y, por tanto, también de cualquier crítica. Por ello, tratar de evitar cualquier confusión beneficiará al proyecto final. Una posible solución sería añadir las palabras que conforman la sigla en inglés y su traducción. Así, el lector posee toda la información y no se ve obligado a tener que buscar en otra fuente las palabras originales que componen dicho término. Aunque esta parece una solución con numerosas ventajas, también posee puntos en contra. Es importante tener claro que se trata de un texto preparado para una página web, un medio donde en ocasiones el espacio está bastante restringido. Por tanto, añadir no solo una traducción sino también las palabras en el idioma original que la conforman puede llegar a ocupar muchos caracteres de los que no se dispone. Otra opción extraída de este subcorpus consistía en añadir la traducción de las palabras que componen las siglas y especificar que provienen del inglés. Esta es la estrategia que podría resultar menos adecuada ya que se está ofreciendo información innecesaria para el tipo de texto que se está tratando. Si se diera el caso de estar traduciendo un texto de lingüística o etimología, sería mucho más lógico incluir este tipo de datos. A continuación, se añadirá una última opción de traducción que, aunque no aparece en este subcorpus, resulta interesante mencionar. Esta se utilizará como argumento para explicar por qué los tres métodos anteriores son las opciones más frecuentes y aceptables. Esta técnica consiste en emplear solamente las palabras de donde provienen las siglas. Por ejemplo, en el caso de una entidad, se usaría Instituto Nacional de Cáncer y se omitiría NCI. Este método de traducción puede no ser el más adecuado ni aconsejable puesto que se puede comprobar con frecuencia en cualquier tipo de texto, estudio o investigación que los autores tratan de reducir conceptos formados a partir de dos o más palabras a siglas o acrónimos, creándolos si es necesario. Esto tiene dos motivos. El primero de ellos se desarrolla en base a que utilizar siglas agiliza el trabajo de redacción del emisor. Normalmente cualquier autor utiliza la primera vez que quiere mencionar un término formado a partir de dos o más palabras el nombre completo junto a sus siglas para, a partir de entonces, solo citar dichas siglas. De este modo, se ahorra escribir el mismo término una y otra vez. El segundo motivo que demuestra la utilidad de las siglas o acrónimos es que agiliza la lectura del receptor, que puede llegar a cansarse de leer el mismo término de cinco palabras veinte veces en una misma página. Por lo tanto, y a modo de conclusión, para estas cuatro opciones de traducción, se cree que la más adecuada por la clase de información que ofrece sería la de añadir a las siglas, las palabras que la conforman en inglés y su traducción al español, siempre y cuando no suponga un problema de espacio. Además, aunque añadir de qué idioma provienen dichas siglas no parece necesario, lo que está claro es que el uso de ellas es de gran ayuda y omitirlas podría entorpecer la lectura completa del texto. Asimismo, cabe recalcar que sea cual sea la opción escogida, es necesario mantener una coherencia a lo largo de todo

el texto, algo que no ocurría en ciertos textos que conformaban el subcorpus de traducciones.

Otro elemento que se ha analizado en esta investigación es el uso de nombres de medicamentos. En el marco teórico se veía que este aspecto suponía grandes problemas de traducción debido a que los autores españoles solían optar por los nombres comerciales mientras que los autores de habla inglesa preferían los nombres genéricos. Es de vital importancia tener en cuenta un aspecto cultural como este si se quiere elaborar una traducción lo más natural posible. Por esta misma razón, se quiso analizar qué nombres se usaban para apelar a los medicamentos en este corpus. Los resultados fueron bastante sorprendentes puesto que los textos originales utilizaban los nombres genéricos mientras que los traducidos, incluso en sus versiones originales, ofrecían ambas opciones. Esto se puede deber a que los emisores en lengua inglesa están más concienciados sobre lo positivo que resulta formar a los pacientes y sus familiares sobre ciertas nociones. De todos modos, parece una técnica que también utilizarían los traductores independientemente de cómo apareciese este aspecto en el texto original. Asimismo, de la búsqueda y análisis de estos elementos, se ha observado que tanto autores como traductores tienen bastantes conocimientos sobre nociones ortográficas relacionadas con medicamentos y fármacos ya que no se ha podido encontrar ningún error en ninguno de los dos subcorpus en cuanto al uso de mayúsculas y minúsculas. No obstante, sí que se han identificado algunas omisiones de artículos en el subcorpus de textos originales, lo que supone un error si se está hablando de nombres genéricos de medicamentos. Aun así, el hecho de que el número de errores sea tan reducido podría considerarse un hecho bastante positivo ya que demuestra que ambos profesionales se preocupan por la estética de sus textos y no solamente por el contenido o la información que proporcionan.

La forma de dirigirse al receptor fue otra de las cuestiones observadas y examinadas puesto que no existe ningún consenso que especifique cuál es la mejor forma de referirse a la persona que está leyendo un texto médico de carácter informativo. De ahí, se extrajeron diferentes opciones de las que se ha podido concluir que tratar a la persona que sufre cáncer colorrectal como “el paciente” o “los pacientes” es una de las opciones más utilizadas. Esto puede deberse a que la entidad no sabe si la lectura se debe a que esa persona sufre la enfermedad o a que simplemente quiere informarse porque está interesado en la temática. De este modo, si trata dicha enfermedad como algo ajeno al lector no está dando por sentado que este la padezca. No obstante, el “tú” y el “usted” juegan también un papel muy importante en la redacción de este tipo de textos ya que otras entidades las ven como una opción más clara. A pesar de que tampoco parecen ponerse de acuerdo en cuál de ellas es mejor, creo que cada una tiene sus puntos fuertes. El “usted” es una forma más alejada del lector, pero más respetuosa y formal, algo que protege a la entidad de ser grosero ya que no se sabe quién se encuentra tras la pantalla. Además, seguramente la forma en que se produce el mensaje sería el mismo modo en el que un médico se lo explicaría a su paciente en persona. Por otro lado, el “tú” proporciona más confianza y simpatía y si la página web que respalda al texto viene acompañada de un entorno más amigable, puede “restar” un poco de seriedad al asunto, teniendo en cuenta que se trata de un tema bastante serio. Si se consideran las recomendaciones que hacen los investigadores, se comentaba anteriormente que el “usted” es una mejor opción para ancianos y el “tú” para niños y adolescentes. La cuestión es que estos textos van dirigidos a adultos de mediana edad en su mayoría, aunque también podrían llegar a

incluir a algún anciano. Por tanto, y solo para asegurar que una persona entrada en edad no se encuentre en una situación incómoda, podría optarse por el “usted” como mejor opción de traducción entre las diferentes posibilidades que ofrece la segunda persona.

También se ha podido comprobar que la primera persona ocupa un papel importante en este tipo de textos. El hecho de que las entidades se refieran a ellos mismos como “nosotros” es algo frecuente en ambos tipos de textos, así como el uso del “yo” para elaborar preguntas que pueden hacerse los pacientes y que después se responden como si se tratara de una conversación en la vida real. De toda la información extraída a partir de la forma de dirigirse al lector, podría concluirse que ambos subcorpus emplean las mismas formas con las mismas funciones y en cuanto a frecuencia de usos, el “usted” y la tercera persona serían las opciones más utilizadas. Aun así, se sugiere lo mismo que en el párrafo anterior: prestar especial atención a la coherencia del texto. Si el autor o traductor decide un pronombre personal, debe mantenerlo a lo largo del texto completo, estrategia que no se seguía en una de las fuentes escogidas.

Como se ha podido comprobar con la forma de dirigirse al receptor, los resultados obtenidos de la búsqueda de pasivas y gerundios son similares en ambos subcorpus. De este análisis, se ha podido demostrar la presencia de pasivas, tanto perifrásticas como pronominales, y de gerundios. Aunque deberían tratar de evitarse en la medida de lo posible debido a su falta de naturalidad o a sus desaconsejados o incluso incorrectos usos, se utilizan tan a menudo que cualquier persona familiarizada con los textos médicos se habrá acostumbrado a ellos y los habrá interiorizado como formas correctas. No obstante, y como se mencionaba en el análisis, el subcorpus de textos originales posee un mayor número de gerundios, un hecho sorprendente pues se cree que esta mala costumbre se debe a las traducciones provenientes del inglés u otras lenguas que condicionan a los traductores a calcar inconscientemente sus estructuras. En el caso de la voz pasiva, ambos subcorpus poseen unas cifras bastante similares en cuanto al número de oraciones que utilizan esta estructura. Por un lado, la cantidad de pasivas perifrásticas es mucho menos frecuente de lo que se podía esperar puesto que se creía que la influencia del inglés era mucho mayor y que las traducciones se verían más afectadas por este aspecto. No obstante, son los textos originales los que, a pesar de poseer casi la misma cantidad de pasivas perifrásticas, se muestran distribuidas en un mayor número de fuentes. Dicho tipo de oraciones proviene prácticamente de una sola entidad en el subcorpus de traducciones. Este hecho indica que los autores se ven más influenciados por el calco de estructuras inglesas que los traductores, probablemente porque a estos últimos sí que se les exige tener una capacidad de redacción de calidad. Por otro lado, las pasivas reflejas son mucho más fáciles de identificar debido a que su presencia es descomunal en todos los textos. No obstante, no suponen ningún problema en cuanto a falta de naturalidad por lo que no se cree necesario prestarles tanta atención. En conclusión, la voz pasiva y el gerundio son dos conceptos a los que se les ha reservado un hueco dentro de la investigación en traducción, sobre todo médica, puesto que su presencia es abrumadora hasta el punto de apelar a un tipo de gerundio “gerundio médico” debido a lo mucho que aparece en los textos que tratan esta temática. Por ello, se creía que este corpus los contendría en abundancia. No obstante, parece que tanta relevancia en los estudios de traducción médica ha concienciado a los traductores de forma que la presencia de ambos aspectos se ha visto bastante reducida, un hecho muy positivo ya que significaría que este tipo de traducción se está acercando cada vez más a lo que se podría considerar una traducción ideal: aquella

que contiene una redacción tan natural que nadie puede imaginar que tal texto no es el original.

De los resultados que se han obtenido de cada uno de estos elementos por separado, se han sacado unas conclusiones generales. A grandes rasgos, se podría decir que estos subcorpus no poseen unas diferencias descomunales entre ellos. De hecho, resulta sorprendente que los resultados de ciertas búsquedas fueran tan escasos. Esto podría considerarse un rasgo positivo de este tipo de textos puesto que significaría que se le presta la suficiente atención tanto a la redacción de textos informativos sobre medicina en español como a su traducción a esta misma lengua. Todo esto puede deberse a que se está comprobando que emitir información sobre ciertas enfermedades a través de la red ayuda a concienciar al público sobre su peligro. Aunque no se puede asegurar que la calidad de redacción y traducción sea la misma para otro tipo de enfermedades menos comunes, sí que se puede afirmar que los textos que tratan el cáncer colorrectal en español peninsular son bastante adecuados para un lego. Las razones en este caso podrían ser más específicas que solo concienciar. Este tipo de cáncer puede curarse en un 90% de los casos si se detecta a tiempo. Esto ha provocado que, al menos en España, muchas comunidades y provincias hayan empezado a emitir mensajes que recuerdan a las personas de mediana edad que a partir de los 50 años sería recomendable hacerse una prueba que confirme que están totalmente sanas. El hecho de dedicar tanto tiempo a este tipo de campañas y a su propagación y difusión puede haber llevado también a controlar otro tipo de textos relacionados con esta enfermedad a los que pudiera acudir un paciente en busca de información: en qué consiste la prueba, por qué es tan peligrosa esta enfermedad, qué tipo de tratamientos existen, entre otros.

Otra razón por la que estos textos presentan cierto nivel de redacción puede deberse al hecho de que se consideran uno de los textos médicos más consultados. Los textos puramente científicos no interesan al público general porque lo más probable es que no los entiendan. En cambio, estos están al alcance de cualquier persona gracias a la red de Internet y presentan un vocabulario bastante accesible. Esto implica que, si una investigación sobre un nuevo método para combatir el VIH presenta algún error ortográfico o una falta de coherencia en la forma de dirigirse al lector, nadie lo tendrá en cuenta porque la importancia del texto reside en el contenido. Su función es dejar por escrito los avances obtenidos para encontrar una cura contra esta enfermedad. Sin embargo, los *Fact Sheet for Patients* son más que solamente un folleto informativo: su intención es que el público lea el texto al completo y si la redacción no está cuidada, probablemente el texto no cumpla con su función.

Existe un último motivo por el que estos escritos presentan una redacción de calidad. Aunque este afecta en mayor medida a los textos traducidos, también podría aplicarse a ciertos textos originales, como aquellos que tratan de convencer a una persona de 50 años de hacerse la prueba que diagnostica el cáncer colorrectal. Se trata de una cuestión de marketing. La mayoría de los textos que componen el subcorpus de traducciones han sido extraídos de fuentes norteamericanas. Allí, la sanidad es privada por lo que cualquier tratamiento supone un coste. El hecho de que muchos hospitales traten de causar buena impresión a través de la información que ofrecen no es solo casualidad. Si el paciente cree que la página de un hospital le ha transmitido confianza, es probable que trate de acudir allí para recibir su tratamiento o para hacerse la prueba de detección de este cáncer. No

obstante, si la información no es clara ni precisa y tampoco presenta una buena redacción, seguramente el lector opte por otra entidad a la que consultar. Por tanto, es lógico que un hospital quiera dar buena impresión a partir de unos textos que serán consultados por un gran número de usuarios en una sociedad donde se acude antes a Internet que a un médico.

Aunque se ha mencionado que no existía un gran número de diferencias entre los textos que conforman el corpus recopilado para este estudio, sí que pueden apreciarse las suficientes como para poder anteponer a uno antes que a otro. En un primer momento, antes de realizar el análisis de textos, pensé que el subcorpus de textos originales poseería unas características más adecuadas a la tipología que se está tratando por el simple hecho de que son personas que redactan en su lengua materna y sobre temas que son su especialidad. No obstante, se ha podido comprobar que aquellos que deben trasladar un texto redactado por otras personas en otras lenguas son capaces de adaptar mejor el texto a su idioma materno.

Existen varios motivos por los que el subcorpus formado a partir de textos traducidos presenta unas características más adecuadas que los textos originales. Para empezar, utilizar los nombres tanto genéricos como comerciales de medicamentos parece una buena estrategia de traducción puesto que el nombre genérico ayuda a comprender un poco la función del medicamento y el nombre comercial puede resultar más familiar al lector. Asimismo, los traductores parecen estar bien formados sobre las normas ortográficas que envuelven a los medicamentos ya que utilizan las mayúsculas y minúsculas de forma adecuada y coherente, aunque esta característica está también muy presente en los textos originales, pero no al mismo nivel. Otro motivo que me ayuda a decantarme por el subcorpus de textos traducidos sería el uso de la terminología. El análisis mostraba que la frecuencia de uso de los términos principales como “cáncer colorrectal” era mucho más elevada que en los textos originales. Esto parece un punto a favor de las traducciones puesto que indica la repetición de términos relevantes en lugar de utilizar sinónimos que no tienen por qué coincidir totalmente en su significado o usar pronombres que pueden confundir al lector si no sabe a qué término sustituyen exactamente. Además, la terminología empleada por estos era bastante sencilla y asequible, pero a su vez bastante exacta y precisa. Esto resulta otro punto a favor puesto que el constante uso de términos desconocidos por el lector le llevará a realizar búsquedas reiteradamente y perderá el ritmo de lectura o le hará directamente desconectar por la falta de comprensión. Asimismo, cuando el autor en lengua inglesa creía que un concepto podía ser demasiado especializado, añadía una pequeña definición o explicación, algo que se ve muy bien reflejado en las traducciones y que resulta de gran ayuda para que el lector comprenda el texto en su totalidad. Otra razón que demuestra que los textos traducidos están bastante bien cuidados es el escaso uso de gerundios y pasivas perifrásticas, sobre todo teniendo en cuenta que estas últimas provenían prácticamente de solo cuatro fuentes, tres de las cuales proporcionan solamente uno o dos ejemplos. Esto significa que las traducciones tratan de evitar la pasiva perifrástica en la medida de lo posible, una construcción que, como se comentaba, suele resultar artificial ya que se utiliza en muchas ocasiones por la influencia que tienen las construcciones y estructuras del texto original. Por otra parte, aunque es recomendable sustituir el gerundio por otra forma, muchos de los casos identificados en el texto se utilizaban de forma correcta.

Como se acaba de comprobar, existen varios motivos por los que el subcorpus de textos traducidos presenta unas características más adecuadas para la tipología y temática que se está tratando. El hecho de que sean los textos traducidos los más cuidados y no los originales también se debe a una serie de razones que se explicarán a continuación. Para empezar, al contrario de lo que ocurre con muchos autores de textos médicos, los traductores son profesionales de la lengua y conocen todos los elementos que deben evitar para conseguir una buena redacción. Además, aunque muchos de ellos no estén igual de formados que un autor de textos médicos en cuanto a conocimientos especializados se refiere, saben cómo documentarse sobre la temática que se trata en sus textos y si no consiguen entender una idea o concepto, acudirán a cualquier fuente, especialistas incluidos, que le ayuden a transmitir dicha idea de la mejor forma posible. El único problema que presenta esto es que, si se basan en fuentes que no tienen un alto nivel de redacción, podrán acoger estructuras procedentes de otras lenguas. El hecho de que el inglés sea de por sí un idioma con una forma de expresión más sencilla ayuda también a que las traducciones presenten ideas mucho más claras. Las oraciones en inglés se forman siempre del mismo modo: sujeto + verbo + complementos. Por el contrario, el español omite sujetos, puede colocar los complementos en cualquier lugar sin alterar el significado de la oración y tiende a variar sus estructuras para no sonar repetitivo. Por tanto, el hecho de que los textos originales provengan del inglés ayuda a que las traducciones posean un discurso mucho más claro y estructurado. Asimismo, la probable diferencia de conocimientos que puede existir entre traductores y profesionales de la medicina influye en la naturalidad y simplicidad del discurso hacia el lector. Esto quiere decir que, al leer el texto original, puede que, si el traductor se ha encontrado con ciertos problemas de comprensión al realizar su primera lectura, haya querido expresar la idea de forma más sencilla en la lengua meta para que no les ocurra lo mismo a los lectores. Y una última razón podría ser la siguiente. Un proceso de traducción llevado a cabo por una empresa que se precie no solamente consistirá en trasladar el texto de una lengua a otra. Incluirá tareas de revisión y edición para evitar cualquier tipo de error ortográfico, terminológico, de sentido o de coherencia. Por tanto, el hecho de haber sido leído por un mínimo de tres personas influye también en la calidad del texto.

9. Traducción

Introduction

The emergence of the Internet has been a turning point in human history. It is true that before anyone had Internet access, other forms of communication existed: the telephone, the ordinary mail, the radio and television. However, the immediacy offered by the Internet cannot be compared to any other previous invention. It is as simple as comparing the loss of users of other devices or services with the number of children and even elderly people who want to use a 4G mobile phone. It is true that people still listen to the radio, watch the news or certain gossip shows on television or use the ordinary mail for certain issues. Nevertheless, this cannot be compared to the number of series that are seen through the different audiovisual content services such as Netflix, HBO or Amazon Prime, the messages sent through applications such as WhatsApp or Instagram, the number of times a song is listened on Spotify or the number of times any information is entered into the Google search bar to find out about any news. Also, the Internet is not only at the top of the communication pyramid, but it is also a key means for many jobs, such as those that require GPS, instant information search or web-connected programmes. In short, the Internet has earned a very important place in people's lives, but why is it so important for medicine or, more specifically, for colorectal cancer?

As it was mentioned above, the Internet is used to communicate and, in many cases, to communicate means to inform. When the word “inform” is used, it does not only refer to what is going on around the world. There are so many sciences and disciplines nowadays that it is impossible for a person to know everything that is involved in each of them. The Internet makes it easy to store information on any subject. The fact that anyone can access to the network and is able to add any piece of information without any kind of filter makes it difficult to be reliable. Nevertheless, the current knowledge about the Internet is so wide that most users already know certain strategies to recognize when a website can be trusted or not.

If the Internet is one of the most complete, accessible and consulted sources of information, any company, entity or organization will want to be part of the network, either through advertisements in other websites or through the creation of their own. Although this may seem more of a marketing strategy through which companies promote their products or services, some entities do not intend to receive a monetary benefit in return. That is where the subject of this study comes in: colorectal cancer.

Many people suffer from cancer, either because of its different types, or because no definitive cure has been found yet for any of them. The fact that it is so aggressive, since it involves a high mortality rate, implies two things: on the one hand, that any person belonging to a developed country knows about it; and, on the other hand, that health entities want to inform and provide knowledge to the general public about this disease so that it can be prevented in the greatest number of cases.

Currently, hospitals, health organizations and insurance companies around the world offer information on certain diseases and conditions with the aim of informing potential patients and helping them to act. Although professionals working for these companies have a high level of knowledge about all these notions, they need to adapt the language to a layperson. To achieve that, this information is grouped into a text genre characterised by its accessible vocabulary, its functionality, and its simplicity. However, health workers in charge of this writings may have great difficulties in summarising a large amount of knowledge that they have acquired throughout their professional career into a basic and reduced message. This problem is probably caused by its lack of training in writing. On the other hand, translation professionals also face certain terminological and decision-making problems that may affect the way a text is received. As they are dealing with an audience from another country, they may need to adapt certain cultural aspects.

Objectives

The main objective of my Master's thesis is to discover the differences, if any, between texts originally written in peninsular Spanish and texts translated into this variety. These translations will always come from English, since combining several source languages could alter the results of the research. In order to make the study more specific, I decided to focus on a text type and a specialised subject: medical texts aimed at a non-specialised audience on colorectal cancer. In order to achieve my goal, it was necessary to collect a large number of texts of both types and to carry out a thorough analysis of each of them. So, I thought that using a corpus of translated texts and original texts could be an appropriate methodology. Therefore, I had a second objective, which was carried out before the main objective, which was to be able to create a comparable corpus following some basic criteria so that the results of this study were not contaminated.

Motivation

Since linguists began to pay special attention to translation and to consider it a discipline worth studying, different fields have been assigned to each of its sciences. One of the broadest and most extensive ones is specialised translation. Specialised translation is formed by a wide variety of subfields, which are sometimes not completely defined. It is quite obvious that texts whose main subject is medicine contain specialised language. However, this statement does not imply that such documents are only intended for an audience that has been previously trained in this science. In fact, since the Internet became accessible to everyone, this type of texts has been increasingly adapted to a general public. This is because medical specialists have observed that training an ordinary person in this discipline can prevent certain diseases or conditions. This is why it is now very common to find texts containing general information on many diseases with a fairly accessible vocabulary and discourse. The problem that lies in this statement is that foreign organizations and entities take their time to produce simple texts that inform an increasingly wide audience, but they do not show the same concern when it comes to their

translation. For this reason, comparing texts that have only been worked on in Spanish with others that have been produced in another language and translated into Spanish will show whether this difference is a real problem as it might seem at first.

In terms of medical translation, this research should not give us any kind of problem since, as medicine is one of the oldest fields of human knowledge, it is one of the most studied branches of translation. However, corpus linguistics and the different tools available to researchers for their analysis are two much more recent fields that have therefore been studied to a lesser extent. Although it is true that there is a large amount of programmes that helps in the handling of corpora in several aspects (making lists of their forms (types), adding labels or providing lines of concordance), the biggest problem lies in the criteria that must be followed for the realization of a corpus, especially when establishing a threshold of representativeness. The fact that this may seem to be an obstacle when carrying out this research is another reason to continue with the project.

Another reason why I chose this type of study is the fact that it is related to one of the main objectives of a good translation: to give the recipient the feeling that he or she is reading an original production and not a translation. Any professional translator will always try to produce a work that contains the most natural language possible in the target language, even making the necessary changes to the original text so that the reading of his or her translation is as fluent as possible. So, I think that a study that tries to analyse the “flaws” of a text that prevents it from being natural in the medical field could be very interesting and necessary since it tries to find a way to get to more and more adequate and realistic productions.

Hypothesis

As it was mentioned in one of the previous sections, this work has two objectives. The first one is to build a comparable corpus. This implies the compilation of a first group of informative texts on colorectal cancer originally written in Spanish and a second group of translations from English into Spanish with its same text type and subject. Also, it must be taken into account that both groups must follow specific guidelines so that they can be considered suitable for this research. The second objective is to find the differences that may exist between the translated texts and the original texts and to conclude whether any of them presents general characteristics that are more appropriate for this type of texts.

My initial hypothesis is that the group of texts written in Spanish will present a much more natural and consistent discourse, as well as a more specific vocabulary for each concept. One of the reasons on which I have based my hypothesis is that, when collecting samples, it could be verified that the authors of these texts were healthcare professionals. However, in the case of the other group, it was practically impossible to check its authorship. This could mean that they had not been carried out by professional translators or by native speakers of the target language. Furthermore, the cost of a translation will increase to a greater or lesser extent depending on the number of words of the text, the source and target languages and the level of specialisation. Although it is true that English

and Spanish are two of the languages with the greatest number of speakers and the level of specialisation is not as high as it could be, it must be taken into account that the vast majority of texts appear on web pages, which means a higher price for the entity as it not only needs a translation, but also a localisation. This could be a reason why the entity has its translation done by a person who is not aware of this price increase in order to save some money. Finally, one more reason why original texts in Spanish are more consistent is the author's level of knowledge. Although the translator will carry out a prior analysis of the text and will document himself or herself, if he or she is not completely careful, he or she may make some mistakes when using certain words or expressions, use some calques or not understand certain ideas and represent them in a wrong way. In any case, these are simple suppositions and this research has been carried out to check whether this hypothesis is totally or partially true or totally wrong.

Methodology

I decided to use a mixed methodology based on the use of corpora, more specifically a comparable corpus made from original texts in Spanish and texts translated into this same language. This was made in order to check if there is any difference between a specific type of text written in one language and this type of text translated into the same language. For this purpose, it was needed to create a corpus. This was due to the fact that there are so many types of corpus that it would be practically impossible to find a large number of texts previously grouped by another researcher that followed the same criteria that I wanted for my study.

Although I will talk about the different typologies and classifications of a corpus later on, I wanted to highlight some of the characteristics that the corpus for this study has in order to explain the process of collecting the texts. After reading several studies that used a corpus, I verified that it was very common to use virtual corpora in translation studies. As I thought that the Internet would surely contain a larger selection of texts than any other physical resource and its compilation would be much simpler than having to scan or copy any paper document, I decided that this corpus would be formed by texts downloaded from the Internet. After making this decision, it was quite clear to me that, among the different specialities we had studied in the master's degree, medical translation had been the one that had attracted my attention the most. For this reason, I decided to look for the different types of documents that doctors, nurses and other health professionals deal with. Then, I began to explore websites of medical entities and I found a quite interesting modality aimed at the general public: the Fact Sheet for Patients. This type can be found on the website of any medical entity, insurance company or hospital, among others. So, I think it is important to give them its importance as they may be one of the most easily accessible medical texts and, therefore, they are probably the most consulted ones by the general public.

Since I was asked to write my thesis, I knew that I wanted it to focus on medical texts because I think that it is a subject that is widely used on the Internet and that offers many

options. This is why, after doing some research on the different concepts involved, such as conditions, diseases, entities, treatments or drugs, among many others, I chose colorectal cancer as the main topic of my texts. The reason why I wanted to focus on this disease is that cancer is one of the medical terms to which society is more used since it is one of the main causes of morbidity in the world (Sociedad Española de Oncología Médica, 2018). Among the more than 200 types of diseases it involves (Asociación Española Contra el Cáncer, 2018), I chose colorectal cancer, including its prevention, causes, symptoms, diagnosis, stages and treatment, among other issues, because it was one of the most diagnosed tumors in Spain in 2017 in both men and women (Sociedad Española de Oncología Médica, 2018). However, it is a disease that can be cured in more than 90% of cases if it is diagnosed in time. This has led the different Spanish communities to carry out early detection programmes to reduce this incidence and to produce texts with basic information on this disease to raise awareness among the population. These reasons assured me the possibility of finding enough texts originally written in Spanish, and most importantly, with a high level of quality and with the subject and function I was looking for. Furthermore, colorectal cancer is also one of the main causes of morbidity worldwide, which also assured me that there would be texts with the same characteristics that were written in English and translated into Spanish.

Once I had decided the type of study and material I was going to use, I began to collect texts for my corpus, which would be divided into two parts, one consisting of texts originally written in Spanish and the other one of texts translated into Spanish. I decided to start collecting the original texts because, despite the fact that Spanish is one of the most widely spoken languages in the world, I was not accepting any variety of Spanish. My goal was to find texts in peninsular Spanish in order to reduce the number of samples and thus obtain much more specific results. This implied relying only on sources from Spanish entities that had a certain degree of reliability. So, it would be a waste of time collecting texts written in English that would have been translated into a neutral variety of Spanish and then find out that the number of samples written directly into peninsular Spanish was much lower. Once I had collected all the possible texts on colorectal cancer produced by Spanish health organizations, I focused on searching for translated texts until I got a similar number of words. In this way, a corpus divided into two parts composed of about 70,000 words each was created.

In order to start analysing both parts of the corpus, I needed a computer programme or application that could help me examine all the texts simultaneously. Searching for certain elements in a group of 83 texts and then in another one of 67 individually would not only be a never-ending task, but it would also prevent me from identifying each example I was going to focus on, which would condition the results of the study. To do this, I chose AntConc, a free and easily accessible programme. I had worked with it on several occasions before and it had proved to be very useful. Before starting with the analysis, different aspects that could give some problems were chosen. I got some of these elements from the theoretical notions consulted at the beginning of this study. The bibliography that was looked up was based on medical texts, especially on the chosen genre, the Fact Sheet for Patients, web pages, and their respective translations. These elements were also

chosen based on the kind of tasks allowed by the AntConc application. This means that, although some elements are quite usual in the text type that was chosen for this research paper, they can be very difficult to identify. An example that explains this issue is the presence of spelling mistakes. Although this is an important aspect in translation, especially in web pages, it is quite complicated to look for them if we take into account that these errors range from accents to alterations of letters, including the lack of concordance between article and word or subject and verb, among many others. Thus, I concluded that the gerund, the passive voice, the terminology, acronyms, the names of drugs and the way of addressing the reader would be appropriate elements for two reasons. On the one hand, these are aspects that have already been identified as problematic in other research papers. On the other hand, the chosen computer programme does not present any problem to look for them in the corpus.

Once the elements to be examined have been chosen, it was necessary to look for them and, then, analyse them in order to be able to make a comparison between both parts of the corpus. In this way, it will be possible to confirm the hypothesis mentioned at the beginning of the study. This hypothesis stated that there would be differences between both parts of the corpus and more specifically that the original texts would be more careful, natural and fluid than the translated texts. Likewise, the decisions made by both the writers and the translators would be also observed and it would be pointed out which strategy is the most appropriate one for each of the elements to be analysed, according to everything learned in the theoretical framework.

Theoretical framework

Corpus

Definition

According to the Centro Virtual Cervantes' Dictionary of Key Terms, the word corpus has two meanings from a specialised point of view. The first one “*designa la recopilación de material lingüístico hecha con un propósito de investigación concreto, ya sean muestras de oraciones, de enunciados o de textos*” (Centro Virtual Cervantes, online) while the second one refers to “*una recopilación extensa de textos (escritos, orales o de ambos tipos) recogidos con el fin de servir como muestra representativa de una lengua, como conjunto de datos lingüísticos reales que reflejen el uso de la lengua (o del tipo de lengua específico) del cual quieren ser representativos*” (Centro Virtual Cervantes, online). This last definition is closer to the type of corpus that will be referred to throughout this paper.

Pérez defines a corpus as “*cualquier colección que contenga más de un texto*” (Pérez, 2002, online). However, she is aware that there are several characteristics that this collection of texts has to possess in order to be considered a corpus. This is why she relies on the definitions of different authors such as Francis (1982) or Sinclair (1991) to show

three implications which any corpus currently possesses: the magnetic format of its texts, an adequate representativeness and a series of criteria that will be discussed later on. A definition that follows these guidelines and that seems to be quite in line with the term we are going to deal with would be the one proposed by Sinclair: “a corpus is a collection of pieces of language text in electronic form, selected according to external criteria to represent, as far as possible, a language or language variety as a source of data for linguistic research” (Sinclair, 2004 in Navarro Colorado, 2015).

General criteria for the design of a corpus

Many authors have classified the different types of corpora that are available nowadays. This distinction is based on a great variety of criteria that lead the way of the researcher when designing a corpus. Before starting the collection of texts, it is important to be clear about the type of corpus that is going to be used for the research. That is why some criteria that have been collected from the studies of different authors will be explained later on.

Although none of the studies consulted fully agrees on the exact number of existing criteria or their denomination, they all have something in common: determining the size of the corpus, which refers to the number of texts or, sometimes, to the number of words that it will contain. Some of them, such as Torruella and Llisterra (1999) or Tolchinsky (2014), include in this same category the distribution of the texts. Another level that is quite present in a large part of the classifications is the specificity of the texts, that is, whether the textual subject is going to be general or specialised. Determining whether the samples will be written, oral or transcribed is another guideline that should be taken into account before collecting the texts. Pérez included in her classification of criteria for specialised corpora the division between original texts, translations and texts produced by native or non-native speakers in her category “condición lingüística de los textos” (Pérez, 2002, online), where the decision of the means of production of the text also appeared. The number of languages and which ones will be used is also important to select the texts. This means that a corpus can be monolingual or multilingual. The latter category includes the subcategories of parallel, comparable or texts in different languages according to Professors Sierra, Bel and Lazaro (2018, online). These authors also pointed out the importance of determining spontaneity, authorship, time, accessibility, purpose, quality, coding or annotation and documentation. These categories also appear in other studies independently. Flores Acuña (2014) wanted to include in her classification the subject, text type and genre, which is an aspect also supported by Navarro Colorado (2015). He inserted in his classification the division between the use of full texts or fragments, in the same way that Tolchinsky (2014) did. Finally, we have the fact that a corpus contains only virtual samples, paper samples or a combination of both and the distinction between fictitious or real texts. In the case of dealing only with a specialised corpus, Pérez (2002) points out the importance of the text type, as it has been mentioned previously, the level of technicality and the recipient of the text.

As a conclusion to this section, it can be seen that, in order to compile a corpus, it is necessary to make several decisions beforehand: to use only online resources, paper samples or a combination of both? To focus on a specialised field or to create a corpus that represents speech in a general context? To collect complete texts or only the fragments that are of interest? Among these, there will be many more questions that a researcher should ask himself before collecting a corpus. All this will be done in order to collect a number of texts that present the same characteristics so that the corpus is representative of speech in a specific context.

Representativeness

Some authors consider that representativeness is another criterion that must be taken into account when compiling texts for an adequate corpus (Sierra, Bel & Lázaro, 2018). In fact, it could be said that this is one of the most important aspects when it comes to designing a corpus. This is due to the fact that a study carried out from a corpus that is not representative enough could contaminate it completely. However, other researchers believe that representativeness is the reason why the rest of criteria that need to be considered to elaborate a corpus were created (Pérez, 2002). The truth is that this term has always showed a great controversy in corpus linguistics. It was already mentioned by Pérez in *Explotación de los corpórea textuales informatizados para la creación de bases de datos terminológicas basadas en el conocimiento*: “representatividad, estandarización y tipología de los corpórea han sido tres de los temas más debatidos entre la comunidad científica, con opiniones diversas recogidas en varios artículos y propuestas, algunas de ellas hechas en el seno de importantes proyectos europeos” (Pérez, 2002, online). For this reason, a space has been reserved to debate the meaning of representativeness in corpus design.

In the world of corpus linguistics, there has always been a controversy between a quality corpus versus a quantity corpus. This separated those who defended the creation of a large corpus from those who were in favour of a representative corpus that would ensure enough samples to represent an area, which did not imply a minimum number of texts. This is a subject that has been well studied and which authors such as Corpas and Seghiri have written about:

la pregunta surgiría en este punto pues, a pesar de que los adelantos tecnológicos nos permiten gestionar grandes volúmenes de información, cabe plantearnos éstos si son verdaderamente necesarios o si corpus más reducidos pueden ser igualmente representativos —al menos en un campo concreto de especialidad (Corpas & Seghiri, 2006, p.5).

As explained above, there has always been this conflict between quantity, a category that represents large corpora, and quality, which embodies the smaller ones, but with enough

samples to be considered representative. This confrontation involves a great problem and it is that, after all, both options have their advantages and disadvantages. In the case of a corpus of a general subject, the compilation of a very large number of documents could be a simple task, since there are many texts representing general speech in both digital and physical resources. Also, the more documents obtained, the more samples that will represent a language in either written and audio. However, when a specialised corpus is chosen, where it is required to be as much specific as possible, it may be quite complicated to approach even the number of words or texts that a general corpus may contain. Moreover, specialised corpora are not the only ones that have problems to achieve an adequate number of documents. A corpus that requires a translation, such as comparable or parallel corpora, may make it more difficult for the researcher to collect enough samples. This is due to several factors that “*dificultan el equilibrio de la muestra, como la escasez de traducciones en ciertos países, el estatus de la traducción en un período determinado, la reputación del traductor, el sexo del autor de los originales y de los traductores, etc*”. (Rojo, 2013, p.116), which does not mean that, because of the fact that the number of samples is not equal to that of a general corpus, it will not be representative, as there are probably not such a number of data in specialised areas.

On the other hand, although it will rely on one's own criteria depending on the needs of the corpus, it is important to obtain enough samples to prove that the data that has been extracted from the analysis are valid and acceptable. If the number is not high enough to have a wide variety of texts showing all the possibilities that exist within the same text type and subject, the results may not really represent that text type and subject. This is why, in some cases where a text type or subject has enough samples to compile a corpus with a considerable number of writings, it is necessary to find a way to know when to stop collecting texts. This is because sometimes corpora are collected for a specific use, as it is the case of ad hoc corpora, where samples are needed to help in a translation. However, in this type of corpora, the translator cannot spend all the time given for the assignment collecting texts.

Other problems in designing a corpus

Representativeness is one of the aspects that gives more problems to the use of corpus in research. It is necessary that the chosen corpus has enough texts that follow the same patterns for the results to be reliable. Finding corpora on the Internet is not difficult. In fact, Sierra, Bel and Lazaro offer a list of easily accessible corpora such as the Corpus Diacrónico del Español (Diachronic Spanish Corpus), the Corpus de Referencia del Español Actual (Reference Corpus of Current Spanish), the British National Corpus, the Corpus del Español de Mark Davies (Mark Davies' Spanish Corpus), the Archivo de textos hispánicos (Hispanic Text Archive), the ETDEWEB Database, the Corpus del Derecho Ambiental Mexicano (Mexican Environmental Law Corpus), among others (2018, online). The Institute of Applied Linguistics of the University of Cadiz (online) provides different multilingual, special and parallel corpus samples: the Hansard Corpus, the ACTRES parallel Corpus, the Fono.ELE and the CHILDES. It also offers other

monolingual options for specific languages such as English or French. Martin Weisser (2016) created a large list of specialised corpus in English, including the Wolverhampton Business English, the Saarbruecken Corpus of Spoken English, the Oxford Psycholinguistic Database, the Business Letters Corpus or the Air Traffic Control Corpus, among many others. The problem that arises from this is that, although it has been proven that there is a large number of quality and easily accessible corpora, it is quite difficult for these to accomplish every feature required for a specific research. The solution lies in creating our own corpus. Although this might seem a good solution at first, we must consider the difficulties that arise from this alternative.

At the beginning of this section, it was explained that in order to carry out a study based on the use of a corpus, it was necessary for the corpus to have enough quality texts that followed the same patterns. The term quality implies that the texts must come from a reliable source. However, sometimes none of the corpora that has already been compiled by other professionals meets the requirements of the subject. This means that we must design our own corpus. In these cases, it is important that the documents that make up the corpus have been written by professionals or have been selected from sources that provide a certain degree of reliability. This can sometimes be a problem because some entities entrust the writings to different professionals, both internal and external to the entity. So, these texts are not signed, but simply belong to the entity as such, which makes it difficult to know who the author is. Thus, it is essential to focus on the professionalism of the entity. However, this is not the only obstacle for a researcher who wants to use a methodology based on the use of a corpus created by himself or herself. If translations are used, for example, it is necessary to prove that the documents have been previously written in another language. So, the researcher must look for their authors in the original language, the main language of the entity or a translator's signature.

It is very difficult to achieve an adequate size for a corpus since every sample will have to comply with some previously established features. It is true that there are both physical and digital sources and that, with the spread of the Internet, the number of texts on the net is increasing. However, depending on the topic, it is still quite difficult to find enough texts that have a certain level of quality or reliability. This means that, if the researcher decides to base his or her corpus on texts dealing exclusively with congenital insensitivity to pain, a very rare condition, even though it is difficult to find documents dealing with it, he or she will not be able to use Wikipedia as a source in order to have one more sample, since any user has the possibility of editing the information provided. In terms of corpus size, the only problem could be not being able to find an appropriate amount of texts but the fact that the exact amount to establish whether a corpus is representative or not is not stipulated yet makes it a little bit easier.

Another task that complicates, to a certain extent, the creation of a corpus is the fact of having to convert the texts to an electronic format if they have been selected in physical format. This process will be necessary in order to use certain tools that help to analyse the different texts that make up the corpus. Although this small problem has a quick solution, it involves a long period of scanning and saving. If the texts have been saved in

digital format, this step can be omitted. Nevertheless, it will still be necessary to eliminate all kind of format, which also implies a good use of time.

Comparable corpora, which are the type of corpus that will be used in this research, also involve complications that must be taken into account. The first of these is that both parts of the corpus will never have the same texts as it might happen with a parallel corpus. This implies that the researcher may find translations of five hundred words on a subject and find out that the original texts dealing with that subject are made up of five thousand words. The problem that this implies is that, if it has been decided to use the complete texts, even if it is solved by equalizing the number of words in each group to obtain a balanced corpus, the group that has a smaller number of texts will surely present a smaller number of differences in the analysis. This could be due to the fact that the author or translator, depending on the case, will have tried to maintain a cohesion and coherence throughout the text, using the same terms or the same kind of discourse. On the other hand, the second problem that this type of corpus may give is that it may not be entirely clear when a text is an original one or a translation. It is true that, in many cases, the names of the authors or the translators appear at the beginning or at the end of the text. However, in the case of the texts that appear on web pages, it is difficult to find their authorship. This is why it is necessary to find the origins of the page through the country where it comes from. Other strategies can be to find out if the option of changing the language exists to check if the authors do appear, to observe if a part of the text has been translated, but more fragments can be found in other languages, or if some elements are written in another language with an explanation that helps to know the origins of the website.

Medical translation

Translation began as an activity whose objective was to allow communication between two groups of people. It surely emerged with the first commercial transactions (Montalt & González, 2014). Since then, this activity has persisted and been introduced into every discipline and science that we have nowadays. One of the basic characteristics of human beings is their need to live in community. What at first may have involved the coexistence of a small group of people has evolved to the current situation, where globalization is considered a fundamental base for our society. This means that all fields require a translator in order to be able to disseminate their news and discoveries to the rest of the world.

Medicine, like translation, is a very old science. Living beings have always had this tendency to fight for the permanence and survival of their species. Hence, one of its basic functions is reproduction. However, it is not enough to reproduce themselves in order to prevent their extinction. It is necessary to combat several threats if humans want to survive and prosper. This is why medicine is one of the disciplines that has evolved the most over the years, allowing the life expectancy of a human being to be around 70-80 years, when a few decades ago it was no more than 40-50. This makes medical translation

one of the oldest types of scientific translation. However, this does not mean that it is the one to which most attention is paid. In fact, it is quite the opposite: the volume of research on medical translation is much smaller than what we could expect. Also, the offer of training programmes on this translation type is very reduced (Muñoz Miquel, 2009).

All the medical knowledge that has been acquired over the years has been recorded. Currently, there are many writings related to medicine. This great variety of documents is due to the fact that each one has a different objective which means that, for example, a pathological anatomy manual for medical students will not contain the same information as the medical history of a patient who has come to the hospital due to a sprain. This means that the field of medicine is so vast that it seems necessary to delimit the aspects that this study will cover.

To begin with, we will briefly explain some notions about medicine that every medical translator must know. In the book *Medical translation step by step* (2014), Montalt and González offer eight aspects, also mentioned by other authors, that differentiate medical translation from other types of translation and that should therefore be taken into account before starting any medical assignment. The first point to consider when receiving a medical translation assignment is that there are different specialties that comprise this discipline (Montalt & González, 2014; Muñoz Miquel, 2009). A hospital is divided by a certain number of departments that deal with the different situations that may lead a person to come to this centre. It is important to have enough knowledge about this since there may be the same term within more than one of these specialties. This term may have different connotations depending on the field being treated. As it was mentioned before, colorectal cancer is the main subject of the texts collected for this corpus. So, these documents could be classified within internal medicine, “*una especialidad médica con una larga tradición, que aporta, una atención global al enfermo adulto de carácter preferentemente hospitalario*” (Sociedad Española de Medicina Interna, online), a specialty that, therefore, includes the diagnosis and treatment of any type of cancer, among other issues.

Another skill that a medical translator must have is the ability to understand any medical aspect, including terminology (Montalt & González, 2014; Muñoz Miquel, 2009), since “whereas the literary translator's main focus is normally on aspects such as register, rhythm, puns, character's attitude, or cultural references, the medical translator's priority is to deal adequately with factual complexity and accuracy” (Montalt & González, 2014, p.20). Other authors, such as Congost Maestre, have also used this comparison between medical and literary translation to highlight certain issues in medical translation that are being underestimated. In this case, the author notes that, “*suele aceptarse que la traducción técnica es más fácil que la literaria solo por el hecho de que los términos son más exactos, sin connotaciones*” (Congost Maestre, 1994, p.22). However, she rejects this statement through two arguments: the first one refers to the fact that the translator, before starting the translation, must understand perfectly a text full of technicalities of which perhaps he or she has not heard before. This will be greater or lesser depending on whether he or she has already dealt with that subject or not. The second reasoning is based on the

high probability of finding specialised terminology for which a translation may not exist. This implies that, in order to work as a medical translator, it is extremely important to understand all the ideas that are being expressed in a text and to be able to translate them into the target language clearly and appropriately. The translator must bear in mind that he or she is working with the most precious thing a person has: his or her health. It is practically impossible to know all the concepts that cover each of the branches of medicine, and even more so if we take into account that only a few medical translators have university degrees on health sciences. That is why it is very important to know reliable sources that provide accurate and correct information, both when translating and when doing research.

In addition to the two aspects mentioned above, it is also essential for translators to be aware of the communicative situations in which medical messages may occur and the genres involved (Montalt & González, 2014; Muñoz Miquel, 2009). It is also necessary to be aware of the poor quality of writing that these texts may present, and to maintain confidentiality, privacy, and respect for patients, among other actions involving any ethical or moral issue (Montalt & González, 2014). Marsh (1999) also points out the importance of considering the reader, since it will have consequences on the strategies to be adopted.

Translation of websites

Nowadays, the Internet is one of the most popular sources of information worldwide. It is fast, easily accessible for both entering and extracting data, and quite affordable. In fact, it is practically impossible to live without Internet access in a developed country, since more and more jobs require this tool to work, students use it daily to gather information or to solve doubts, and most people even carry a 4G device with them everywhere. This allows them to keep in touch with any individual, to buy, sell or entertain themselves through social networks, among many other activities.

The Internet is one of the most influential factors in globalization. Thanks to it, anyone can find out what is going on more than 10,000 kilometers away without having to travel. You can get access to texts, videos or images regardless of their author, origin or date of publication. Nowadays, if any writing is spread on the Internet, the intention is to reach as many people as possible, since the net allows it. This is why many entities take advantage of it to try to reach a wider audience with what they intend to transmit on their web pages. As they already have the channel, they only need to be able to transfer their texts to all those languages that allow them to reach the recipients they are interested in.

The translation of a website can be a great advantage for the entity or organization that it represents since it reaches a greater number of users. It is true that, currently, most people have a certain level of English, either to a greater or lesser extent. However, readers prefer texts written in their mother tongue for two reasons. On the one hand, understanding a text in another language requires a greater effort by the user and, therefore, more time; on the other hand, the texts that appear on these websites may contain some terminology

whose equivalent in their second language is unknown. This means that, in the case of having two options, one written in English and another one in their mother tongue, visitors will choose the one that makes them easier to get the information. This is why large international companies or entities offer their content in different languages.

The problem with this type of translation is that it is usually quite expensive and involves a long process because it implies techniques that do not consist of simply transferring the message from one language into another. The translation of web pages goes much further. The cultural elements and all the technical specifications involved in the correct functioning of a website are two of the factors to which the translator must pay special attention. If this is not the case, the webpage could have the opposite effect to the one intended. However, the managers of these entities are not aware of this problem and entrust these tasks to people or even machines that are not prepared to do it correctly. That is why I think it is important to emphasize that *“economizar durante el proceso de traducción por no comprar nuevos nombres de dominio sería mala decisión de marketing: los gastos que se hacen traduciendo la página profesionalmente al final se amortizan muy rápido”* (Berends, 2007, p.25).

Conclusions

The first goal of this study was to be able to produce a comparable corpus of medical texts on colorectal cancer, which was achieved successfully. The next goal was to confirm whether there were any differences between texts originally written in Spanish and texts with the same function (informing) and subject (colorectal cancer) translated from English into Spanish. It had to be taken into account that the hypothesis of this study suggested that translated texts would be less adequate than the original ones. After the analysis, it could be stated that there were indeed differences in some aspects. On the one hand, the texts originally written in Spanish had a much more varied terminology. This means that they used simple and accessible terms and at the same time highly specialised terms without any explanation or definition. On the other hand, the translated texts had a much more comprehensible terminology, with a reduced degree of specialisation, but at the same time quite specific to the subject being dealt with. Despite the fact that, in these cases, the terms were not very specialised, the English-speaking writer specified the idea to which a concept referred if he or she thought that it was not going to be understood by the layperson. This strategy is reflected in the Spanish translations. This seems to be a fairly effective technique if we take into account that we are dealing with texts that will be consulted by people with little knowledge on the subject and who want to learn more about it because it affects them personally. A specialist who wants to learn more about the subject will turn to much more elaborated studies and not to texts with basic information. Therefore, making this reading a simple and pleasant experience will help the receiver to feel more comfortable and secure despite the situation he or she is experiencing.

In terms of terminology, it was also observed the terms that were most frequently used. Both parts of the corpus obtained the same results although the translated texts had a higher usage of some terms than the original texts. The most consistent explanation could be that the translated texts imitate the English-speaking habit of using the same terms over and over again instead of using synonyms that may confuse the reader and complicate the understanding of the reading. On the other hand, this does not seem to be a concern of the Spanish-speaking authors who do anything to achieve an elaborate discourse, even if this means risking the correct use of the language.

In terms of the acronyms, the difference that arises between both parts of the corpus is due to the fact that texts originally written in Spanish do not have the need to explain where certain acronyms come from since they refer to concepts that are already known by the reader. This would be different if the author felt the need to mention some terms coming from another country where another language is spoken, but it is not the case. However, it can be seen in the translated texts that the translator is forced to explain the meaning or origin of the acronyms he or she uses, since he or she assumes that they are not known by the recipient. The point here is that there is no established technique for explaining the meaning of acronyms. Each translator chooses a different form, none of which could be considered incorrect. However, he or she should decide based on the strengths of each of them.

To begin with, if the English text has an acronym, excepting for the most popular ones such as NATO or HIV that already have their own official translations into Spanish, their meaning should be added in the target language. This is one of the strategies that has been repeated most often throughout this part of the corpus. The problem that arises from using only one translation next to the acronym is that the recipient will see that the words added by the translator do not match the letters of the acronym. Translations are available to anyone, and therefore, criticism too. Therefore, trying to avoid any confusion will benefit the final project. One possible solution would be to add the words that make up the English acronym and its translation. It is a way to give the reader all the information without having to look for it in another source. Although this seems to be the perfect solution, it also has its disadvantages. The translator must think that these texts are prepared for a website, where space is sometimes quite restricted. Therefore, adding not only a translation but also the words in the original language that make it up can take up many characters. Another option extracted from these translated texts was to add the translation of the words that make up the acronym and specify that they come from English. This is the option that could be less appropriate since it is providing unnecessary information for the type of text being dealt with. If it were a text on linguistics or etymology, it would be more logical to include this type of information. Finally, a new translation option that has not been taken from the corpus will be added. This strategy will be used as an argument to explain why the three previous methods are the most frequent and acceptable options. This technique consists in using only the words where the acronym comes from. For example, in the case of an entity, you would use National Cancer Institute and omit NCI. This method of translation may not be the most appropriate or advisable since it can often be seen in any type of text that authors try to

reduce concepts of two or more words to acronyms or abbreviations, creating them if necessary. There are two reasons for this: the first is that using acronyms speeds up the writer's work. Normally, authors would use, the first time they want to mention a term formed by two or more words, the full name together with its acronym and, from then on, they would only use the acronym. In that way, they will not have to write the same term repeatedly. The second reason that proves the usefulness of acronyms is that it makes it easier for the recipient to read. Otherwise, he or she may get tired of reading the same five-word term twenty times on the same page. Therefore, and as a conclusion, considering these four translation options, the most appropriate one could be to add to the acronym the words that make it up in English and its translation into Spanish, as long as it does not give space problems. Furthermore, adding the language from which the acronym comes does not seem necessary. What is obvious is that the use of such acronyms is very helpful and omitting them could hinder the complete reading of the text. It should also be stressed that whatever option is chosen, it is necessary to maintain it throughout the text. This did not happen in certain texts that made up the group of translations of the corpus.

Another element that has been examined in this research is the use of drug names. In the theoretical framework, it was seen that this aspect gave great problems of translation because Spanish authors used to opt for brand names while English-speaking authors preferred generic names. It is very important to take into account a cultural aspect like this one if we want to make a translation as natural as possible. For this same reason, we wanted to analyse which names were used to talk about drugs in this corpus. The results were quite surprising since the original texts used the generic names while the translated ones, even in their original versions, offered both options. This may be because English-speakers are more aware of the benefits of educating patients and their families about certain aspects. In any case, it seems to be a technique that translators would also use regardless of how it appeared in the original text. Likewise, from the search and analysis of these elements, it has been observed that both authors and translators have a lot of knowledge about orthographic rules related to drugs and medicines since there were no mistakes regarding the use of upper or lower case letters. However, some omissions of articles have been identified in the original texts, which is a mistake if we are referring to generic drug names. Still, the fact that the number of errors is so small could be considered a positive development as it shows that both professionals are concerned about the aesthetics of their texts and not only about the content or information they provide.

The way of addressing the reader was another issue since there is no consensus on which is the best way to do it. Different options were extracted, and it has been proved that talking about the person suffering from colorectal cancer as “the patient” or “the patients” is one of the most used options. This may be because the entity does not know if the readers are suffering from the disease or if they are simply interested on the topic. Thus, if the author treats the disease as something alien to the reader, it is not taking for granted that he or she is ill. However, both Spanish personal pronouns “*tú*” and “*usted*” also play a very important role in this type of texts as they are preferred by some other authors. Although they do not seem to agree on which option is better between “*tú*” and “*usted*”,

I think each one has its advantages. The “*usted*” is a more distant form from the reader, but at the same time it is more respectful and formal. This protects the entity from being rude since they do not know who is behind the screen. Also, a message produced in the “*usted*” form represents the way in which a doctor would explain it to his or her patient personally. On the other hand, the “*tú*” provides more confidence and sympathy. In addition, if the website where the text appears has friendly features, it can “take a little bit away” the seriousness of the subject, considering that it is a serious issue. If we take into account the recommendations made by researchers, the “*usted*” is a better option for the elderly and the “*tú*” is more adequate for children and the youth. The point is that these texts are mostly aimed at middle-aged adults, although they could also include some elderly people. Therefore, to ensure that an older person does not feel uncomfortable, the “*usted*” could be chosen as the best translation option.

This research has also helped to verify that the first person has a role in this type of texts. The fact that entities refer to themselves as “we” is common in both types of texts, as well as the use of “I” to make questions about common doubts that patients may have. In this way, they are answered as if it was a real-life conversation. From all the information extracted from the way of addressing the reader, it could be concluded that both parts of the corpus use the same forms with the same functions. In terms of the frequency of use, the “*usted*” and the third person would be the forms that appear more often. In addition to this, authors and translators must pay special attention to the coherence and cohesion of the text. If the author or translator chooses a personal pronoun, it cannot be switched with other forms, as it was the case of one of the chosen sources.

As it has been seen with the form of addressing the recipient, the results obtained from the passives and gerunds are similar in both parts of the corpus. From this analysis, it has been proven the presence of both forms of Spanish passives (*pronominal* or *refleja* and *perifrástica*) and gerunds. Although they should be avoided where possible because they do not sound natural in Spanish or because their use is grammatically wrong, they are used so often that anyone familiar with medical texts will be so used to them that they will think they are correct. Nevertheless, original texts have a greater number of gerunds. This is very surprising since it is believed that this bad habit comes from translations from English or even other languages. So, it is thought that translators unconsciously copied the structures of these other languages and established these calqued structures into Spanish. In the case of the passive voice, both parts of the corpus have similar numbers. On the one hand, the number of *pasivas perifrásticas* is much lower than what was expected. It was believed that the influence of English was much greater in Spanish translations. However, the original texts are the ones that, despite having almost the same amount of *pasivas perifrásticas*, appear in a greater number of sources. These types of sentences are just used in the texts provided by one of the sources taken to make the translations part of the corpus. This indicates that authors are more influenced by the English structure than translators, probably because the latter are required to be good writers. On the other hand, the *pasivas reflejas* are much easier to identify because their presence is greater. However, they do not give problems in terms of lack of naturality, so it is not necessary to pay that much attention to them. In conclusion, the passive voice

and the gerund are two concepts that are given a lot of importance within translation, medical translation especially, since their presence is so overwhelming to the point that a type of gerund has been called in Spanish “medical gerund” because of how much it appears in texts that deal with this subject. Therefore, it was thought that this corpus would have plenty of them. However, it seems that giving them so much relevance in medical translation studies has made translators aware of the presence of both aspects. This awareness has reduced the presence of both elements, which is very positive since it would mean that this translation type is getting closer to what could be considered the perfect translation: a translation that contains such a natural speech that nobody can even imagine that it has been written in another language.

From the results obtained from each of these elements separately, some general conclusions have been drawn. Broadly speaking, it could be said that both parts of the corpus do not differ greatly from each other. In fact, I was surprised that the results of certain searches were so scarce. This could be considered a positive feature of this text type since it would mean that authors and translators are paying enough attention to the writing of Spanish medical texts and to their translations from English into Spanish. This may be because it is being proven that informing about certain diseases through the Internet helps to raise public awareness of their danger. Although the quality of the writing and the translation cannot be assured for other less common diseases, it can be said that texts dealing with colorectal cancer in peninsular Spanish are quite adequate for a layperson. The reasons in this case could be more specific than just awareness. This type of cancer can be cured in 90% of the cases if it is detected in time. This has caused that, at least in Spain, many communities and provinces have started to send messages reminding middle-aged people that from the age of 50 onwards it would be advisable to have a colorectal cancer detection test. The fact of dedicating so much time to this type of campaigns and its diffusion may have also led to monitoring other types of texts related to this disease to which patients may recall in search of information: what the test consists of, why this disease is so dangerous, the different types of treatments, among others.

Another reason why these texts are so well prepared may be the fact that they are considered one of the most widely consulted medical texts. Purely scientific texts are not of interest to the general public because they are unlikely to be understood. Instead, informative texts are readily available on the Internet and they have an accessible vocabulary. This means that, for example, if a research into a new method of ending with HIV has misspelled words or it is inconsistent in its way of addressing the reader, no one will take it into account because the importance of the text lies in the content. Its function is to write down the progress made in finding a cure for this disease. However, Fact Sheet for Patients are more than just an information leaflet: their goal is that the recipient reads the text in its entirety. So, if it does not have a careful wording, the text will not fulfil its function.

There is one final reason why these documents are really well written. Although this mostly involves translated texts, it may also apply to certain original texts, such as those that try to convince a 50-year-old to get tested for colorectal cancer. This reason refers to

marketing. Most of the translated texts have been taken from American sources. There, healthcare is still private, so any treatment costs money. The fact that many hospitals try to make a good impression through the information they offer is not just a coincidence. If a patient believes that a hospital's website has given them confidence, they may try to go there for treatment or testing. However, if the information is not clear, accurate or well written, the reader may choose another entity to consult. Therefore, it is logical that an entity wants to give a good impression through their informative texts since, nowadays, people surf the Internet before referring to a doctor.

Although it has been said that there were not many differences between both parts of the corpus, there are some aspects that makes me prefer one before the other. At first, before carrying out the text analysis, I thought that the part of original texts would have some characteristics more appropriate to the typology being dealt with, simply because they are people who write in their mother tongue and on subjects that are their specialty. However, it has been found that those who must transfer a text already written in other languages are better able to adapt the text to their mother tongue.

There are several reasons why the translated texts have more suitable characteristics than the original texts. To begin with, using both generic and brand names of drugs seems to be a good translation strategy. This is because the generic name helps to understand a little bit about the function of the drug and the brand name may be more familiar to the reader. Also, translators seem to be quite well trained in the spelling rules involving drugs as they use upper and lower case letters appropriately and consistently. Another reason why I prefer the strategies used in translated texts would be the use of terminology. The analysis showed that the frequency of use of terms such as “colorectal cancer” was much higher than in the original texts. This is a point in favor of translations since it proves the repetition of relevant terms instead of using synonyms that do not completely have the same meaning or using pronouns that can confuse the reader if he or she does not know exactly which term they replace. In addition, the terminology used by these was quite simple and affordable, but at the same time accurate and precise. This is another advantage if we consider that the constant use of unknown terms will lead the reader to search them constantly or make him or her switch off from the reading. Also, when English-speaking authors thought that a concept could be too specialised, they added small definitions or explanations, which is very helpful for the reader to understand the text in its entirety. This was reflected in the translations. Another reason why the translated texts are well written is the limited use of gerunds and *pasivas perifrásticas*, especially considering that the latter came from practically only four sources, three of which provide only one or two examples. This means that translations are quite free of *pasivas perifrásticas*, which are unnatural and often used because of the influence of the original text. On the other hand, although the use of the gerund should be avoided, many of the cases identified in the text were correctly used.

As we have just seen, there are several reasons why the translated texts have more appropriate features for the text type and subject being dealt with. The fact that the translated texts are better written than the original ones is also due to a series of reasons

that will be explained down below. To begin with, unlike many authors of medical texts, translators are language professionals and they know all the elements to avoid in order to achieve good writing. Furthermore, although many of them are not as well trained as healthcare professionals in terms of medical knowledge, they know how to document themselves about the subject of their texts and if they cannot understand an idea or concept on their own, they will turn to any source to help them convey that idea in the best possible way. The only problem that arises from this is that, if they rely on sources that do not have quality writing, they may fall into the same bad habits. The fact that English is a language with a simpler form of expression also helps translations to present much clearer structures. English sentences are always formed in the same way: subject + verb + complement. In contrast, Spanish omits subjects, it can place complements anywhere in the sentence, and tends to vary its structures so as not to sound repetitive. Therefore, the fact that the original texts come from English helps translations to have a clearer and more structured discourse. Also, the fact that these translators may not have the same level of knowledge as a medical professional who writes this type of texts makes it closer to the reader. This means that, when reading the original text, the translator may have encountered certain comprehension problems when doing his or her first reading. So, he or she may have wanted to express the idea in a simpler way in the target language so that the reader does not have the same difficulties. A final reason could be the following. A translation process carried out by a reputable company will not only consist of transferring the text from one language into another. It will include proofreading and editing to avoid any spelling, terminology, meaning or consistency errors. The fact that it has been read by at least three people also improves the quality of the text.

10. Bibliografía

10.1. Fuentes consultadas y utilizadas

Acantho Ideas & Culturas. (2016). *La importancia de la localización web*. Recuperado de: <https://acantho.eu/es/la-importancia-de-la-localizacion-web>

Asociación Española Contra el Cáncer (s.f.). *¿Qué es el cáncer?* Recuperado de: <https://www.aecc.es/es/todo-sobre-cancer/que-es-cancer>

Berends, A. (2007). *La traducción de páginas web* (Trabajo de Fin de Máster). Utrecht: Universidad de Utrecht. Recuperado de: <https://pdfslide.net/documents/la-traduccion-de-paginas-web-a-berends.html>

Calpe Berdiel, L. (s.f.). *Haz de tu web médica un sitio confiable: sellos de calidad*. Laura+Calpe. Recuperado de: <https://www.lauracalpe.com/web-medica-confiable-sellos-de-calidad/>

Chemocare. (s.f.). *Medicinas para Quimioterapia – Nombres de Medicina*. Recuperado de: <http://chemocare.com/es/default.aspx>

Congost Maestre, N. (1994). *Problemas de la traducción técnica: los textos médicos en inglés*. Alicante: Universidad de Alicante.

Corpas Pastor, G. y Seghiri Domínguez, M. (2006). *El concepto de representatividad en la lingüística del corpus: aproximaciones teóricas y metodológicas*. Málaga: Servicio de Publicaciones de la Universidad. Recuperado de: https://www.uma.es/hum892/Publicaciones/Corpas_Seghiri_2006i.pdf

Corpas Pastor, G. y Seghiri Domínguez, M. (2007). *Determinación del umbral de representatividad de un corpus mediante el algoritmo N-Cor*. Procesamiento del Lenguaje Natural (39), 165-172. Recuperado de: <http://journal.sepln.org/sepln/ojs/ojs/index.php/pln/article/view/2670/1188>

Corte, N. (2002). *Localización e internacionalización de sitios web*. Revista Tradumàtica – Traducció i Tecnologies de la Informació i la Comunicació (1), 1-8. Recuperado de: <https://ddd.uab.cat/pub/tradumatica/15787559n1/15787559n1a3corte.pdf>

De Felipe Boto, M. R. (2007). *El enfoque basado en corpus como metodología para investigar rasgos de normalización en la lengua traducida*. Interlingüística (17), 261-267. Recuperado de: <https://dialnet.unirioja.es/servlet/articulo?codigo=2317058>

De La Cova Morillo-Velarde, M. E. (2017). *La localización de la ayuda online: Categorización de problemas para la traducción* (Tesis Doctoral). Sevilla:

Universidad de Sevilla. Recuperado de:
<https://dialnet.unirioja.es/servlet/tesis?codigo=146748>

Diccionario de término clave de ELE del Centro Virtual Cervantes. (s.f.).
Lingüística de corpus. Recuperado de:
https://cvc.cervantes.es/ensenanza/biblioteca_ele/diccio_ele/diccionario/linguisticacorporus.htm

Educalingo (s.f.). Diccionario español. Recuperado de:
<https://educalingo.com/es/about>

EUPATI Academia Europea de Pacientes. (s.f.) *Denominación común internacional*. Recuperado de:
<http://www.patientsacademy.eu/es/glossary/denominacion-comun-internacional/>

Félix Fernández, L. y Alarcón Navío, E. (1998). La terminología especializada y el léxico común en el marco de la traducción de textos médicos de divulgación. En L. Félix Fernández y E. Ortega Arjonilla (Coords.), *Traducción e Interpretación en el ámbito biosanitario* (pp. 167-187). Granada: Editorial Comares.

Flores Acuña, E. (2014). *El corpus como herramienta para la traducción especializada italiano/español: una experiencia con textos de la industria cosmética*. *Philologia Hispalensis* 28 (3-4), 155-182. Recuperado de:
<https://idus.us.es/xmlui/bitstream/handle/11441/36991/EI%20corpus%20como%20herramienta%20para%20la%20traduccion%20especializada.pdf?sequence=1>

Froehlich, H. (2018). *Análisis de corpus con AntConc*. The Programming Historian. Recuperado de:
<https://programminghistorian.org/es/lecciones/analisis-de-corpus-con-antconc>

Infomed. (2014). *El gerundio en el texto médico*. Traducción y Terminología Médicas. Recuperado de: <https://temas.sld.cu/traduccion/2014/11/07/el-gerundio-en-el-texto-medico/>

Infomed. (2016). *Exceso de construcciones pasivas*. Traducción y Terminología Médicas. Recuperado de: <https://temas.sld.cu/traduccion/2016/05/05/exceso-de-construcciones-pasivas/>

Instituto de Lingüística Aplicada de la Universidad de Cádiz. (s.f.) *Corpus multilingües, especiales y paralelos*. Recuperado de: <https://ila.uca.es/corpus-multilingues-especiales-y-paralelos/>

Liroz, F. (2010). *Los tipos de se*. Fernando.Liroz.es. Recuperado de:
<http://fernando.liroz.es/s/tiposse.htm>

Marsh, M. (1999). Algunas consideraciones sobre la traducción médica. En A. Gil de Carrasco y L. Hickey (Coords.), *Aproximaciones a la traducción*. Recuperado de: <https://cvc.cervantes.es/lengua/aproximaciones/marsh.htm>

Mayor Serrano, M. B. (2008). *Cómo elaborar folletos de salud destinados a los pacientes*. Fundación Dr. Antonio Esteve (14), 1-75. Recuperado de: <https://www.esteve.org/capitulos/documento-completo-18/>

Mediluce Cabrera, G. (2002). *El gerundio médico*. Panacea 3 (7), 74-79. Recuperado de: <https://dialnet.unirioja.es/servlet/articulo?codigo=6555093>

Montalt Resurrecció, V. y González Davies, M. (2014). *Medical Translation step by step: learning by drafting*. New York: Routledge.

Muñoz Miquel, A. (2009). *El perfil del traductor médico: diseño de un estudio de corte socioprofesional*. Panacea 10 (30), 157-167. Recuperado de: https://www.tremedica.org/wp-content/uploads/n30_tribuna-Miquel.pdf

Muñoz Miquel, A. (2016) *La traducción médica como especialidad académica: algunos rasgos definitorios*. Hermeneus (18), 235-267. Recuperado de: <https://dialnet.unirioja.es/>

Navarro Colorado, B. (2014). *Diseño y compilación de corpus*. Alicante, Repositorio Institucional de la Universidad de Alicante. Recuperado de: <https://rua.ua.es/dspace/bitstream/10045/43959/3/3.DisenyoCompilacionCorpus.pdf>

Navarro, F. A. (2002) *La nomenclatura de los fármacos (II)*. El Trujamán. Recuperado de: <https://esteve.org/wp-content/uploads/2018/01/137015.pdf>

Navarro, F. A. (2003). *La traducción de los medicamentos (III)*. El Trujamán. Recuperado de: <https://cvc.cervantes.es/trujaman/busqueda/resultadosbusqueda.asp?Ver=50&Pagina=1&Titulo=La%20traducci%F3n%20de%20los%20medicamentos%20&OrdenResultados=2>

Navarro, F. A. (2003). *La traducción de los medicamentos (IV)*. El Trujamán. Recuperado de: <https://cvc.cervantes.es/trujaman/busqueda/resultadosbusqueda.asp?Ver=50&Pagina=1&Titulo=La%20traducci%F3n%20de%20los%20medicamentos%20&OrdenResultados=2>

Navarro, F. A. (2003). *La traducción de los medicamentos (VI)*. El Trujamán. Recuperado de: <https://cvc.cervantes.es/trujaman/busqueda/resultadosbusqueda.asp?Ver=50&Pa>

[gina=1&Titulo=La%20traducci%F3n%20de%20los%20medicamentos%20&OrdenResultados=2](https://cvc.cervantes.es/trujaman/busqueda/resultadosbusqueda.asp?Ver=50&Pagina=1&Titulo=La%20traducci%F3n%20de%20los%20medicamentos%20&OrdenResultados=2)

Navarro, F. A. (2003). *La traducción de los medicamentos (VII)*. El Trujamán. Recuperado de: <https://cvc.cervantes.es/trujaman/busqueda/resultadosbusqueda.asp?Ver=50&Pagina=1&Titulo=La%20traducci%F3n%20de%20los%20medicamentos%20&OrdenResultados=2>

Navarro, F. A. (2003). *La traducción de los medicamentos (VIII)*. El Trujamán. Recuperado de: <https://cvc.cervantes.es/trujaman/busqueda/resultadosbusqueda.asp?Ver=50&Pagina=1&Titulo=La%20traducci%F3n%20de%20los%20medicamentos%20&OrdenResultados=2>

Navarro, F. A., Hernández, F. y Rodríguez-Villanueva, L. (1994) *Uso y abuso de la voz pasiva en el lenguaje médico escrito*. Medicina Clínica (103), 461-464. Recuperado de: <https://www.esteve.org/capitulos/8-uso-y-abuso-de-la-voz-pasiva-en-el-lenguaje-medico-escrito/>

Olvera-Lobo, M. D. y Castillo-Rodríguez, C. (2018). *Recomendaciones para la localización de sitios web en el contexto corporativo: las pymes españolas*. Transibérica 2018: 1º Congreso Internacional de Traducción e Interpretación de lenguas ibéricas, en línea. Recuperado de: <http://locweb.aulaint.es/wp-content/uploads/2019/01/RECOMENDACIONES-PARA-LA-LOCALIZACION-DE-SITIOS-WEB-EN-EL-CONTEXTO.pdf>

Ortega Arjonilla, E. y Félix Fernández, L. (1998). *Traducción e interpretación en el ámbito biosanitario*. Granada: Editorial Comares.

Outbook. (2010). *Listado de palabras vacías en español*. Recuperado de: https://blog.outbook.es/biblioteconomia_y_documentacion/documentacion/listado-de-palabras-vacias-en-espanol

Pereda, V. (2012). *¿Cómo crear un Corpus Lingüístico?* PeredaValentinaTIC. Recuperado de: <https://peredavalentinatic.wordpress.com/2012/03/22/como-crear-un-corpus-linguistico/>

Pérez Hernández, M. C. (2002). *Explotación de los corpóra textuales informatizados para la creación de bases de datos terminológicas basadas en el conocimiento*. Estudios de Lingüística del Español (18). Recuperado de: <http://elies.rediris.es/elies18/index.html>

Sánchez Castany, R. (2013). *Los géneros de Información para pacientes en el contexto español: una primera aproximación*. Fórum de recerca (18), 833-846. Recuperado de: <http://repositori.uji.es/xmlui/handle/10234/132808>

Santamaría Pérez, M. I. (2009). *Análisis textual con el programa de concordancias Antconc*. Recuperado de: https://rua.ua.es/dspace/bitstream/10045/12768/8/Tutorial_Programa_de_Concordancias.pdf

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), 12-30. Recuperado de: https://scielo.conicyt.cl/scielo.php?pid=S0718-48832011000200002&script=sci_arttext&tlng=pt

Sierra, G., Bel, G. y Lázaro Hernández, J. A. (2018). *Lingüística de corpus: Procesamiento de corpus textuales y orales*. Recuperado de: <http://www.corpus.unam.mx/cursocorpus/index.html>

Sociedad Española de Medicina Interna. (2007). *Guía de Formación de Especialistas en Medicina Interna*. Recuperado de: <https://www.fesemi.org/grupos/formacion/publicaciones/guia-de-formacion-de-especialistas-en-medicina-interna>

Sociedad Española de Oncología Médica. (2018). *Las cifras del cáncer en España 2018*. Recuperado de: https://seom.org/seomcms/images/stories/recursos/Las_Cifras_del_cancer_en_Espana2018.pdf

Tolchinsky, L. (2014). *El uso de corpus lingüísticos como herramienta pedagógica*. Textos de Didáctica de la Lengua y de la Literatura (65), 9-17. Recuperado de: <http://diposit.ub.edu/dspace/bitstream/2445/66774/1/653407.pdf>

Tomasena, M. (s.f.) *Cómo decidir si escribir de tú o de usted en la web de tu negocio*. Maider Tomasena. Donde el copywriting genera resultados. Recuperado de: <https://www.maidertomasena.com/como-decidir-si-escribir-de-tu-o-de-usted-en-la-web-de-tu-negocio/>

Torruella, J. y Llisterri, J. (1999). Diseño de corpus textuales y orales. En J.M. Blecua, G. Clavería, C. Sánchez y J. Torruella (Eds.), *Filología e informática. Nuevas tecnologías en los estudios lingüísticos* (pp.45-77). Barcelona: Seminari de Filologia i Informàtica, Departament de Filologia Espanyola, Universitat Autònoma de Barcelona – Editorial Milenio. Recuperado de: http://liceu.uab.es/~joaquim/publicacions/Torruella_Llisterri_99.pdf

Trusted Translations. (2010). *Traducción médica: medicamentos y fármacos*. Recuperado de: <https://blog-de-traduccion.trustedtranslations.com/traduccion-medica-medicamentos-y-farmacos-2010-07-23.html>

Villoria, V. (s.f.) *Abreviaturas, siglas y acrónimos. Teoría*. Proyecto Aula. Lengua y Literatura en Internet. Recuperado de:

<https://lenguayliteratura.org/proyectoaula/abreviaturas-siglas-y-acronimos-teoria/>

Weisser, M. (2016). *Specialised Corpora*. Recuperado de: http://martinweisser.org/corpora_site/spec_corpora.html

WordReference. (s.f.). *Online Language Dictionaries*. Recuperado de: <https://www.wordreference.com/>

10.2. Fuentes de los textos que conforman el corpus

10.2.1. Subcorpus de textos originales

Adeslas Seguros Médicos:
https://adeslas.isalud.com/?utm_medium=cpc&utm_source=google&utm_campaign=280415573&utm_term=adeslas_e&account_id=3536760232&adgroup=19825660493&keyword=adeslas&matchtype=e&device=c&network=g&creative_id=272290310770&campaign_name=Search%20-%20ES/es%20-%20Provider%20Adeslas%20-%20Exact&campaign_type=search&campaign=GN-AD-S&geoloc=1005487&gclid=Cj0KCQjwpNr4BRDYARIsAADIx9x6AsUFZv_6Tu0fhMCT6eFal57aCQvck3sLr-UKRAT90aLRT0HbJm0aArWDEALw_wcB

Amgen: <https://www.amgen.es/>

Asociación Española Contra el Cáncer: <https://www.aecc.es/es>

Astursalud: <https://www.astursalud.es/astursalud>

Barnaclínic: <https://www.barnaclinic.com/es/>

Centro Médico-Quirúrgico de Enfermedades Digestivas:
<https://www.cmed.es/>

Cinfasalud: <https://cinfasalud.cinfa.com/>

Clínica Universidad de Navarra: <https://www.cun.es/>

EuropaColon España: <https://europacolonespana.org/>

Fisterra: <https://www.fisterra.com/>

Fundación Española del Aparato Digestivo: <https://www.saludigestivo.es/>

Gobierno de Cantabria: <http://saludcantabria.es/>

Gobierno de Castilla y León: <https://www.saludcastillayleon.es/es>

Gobierno de la Comunidad de Madrid: <https://www.comunidad.madrid/>

Gobierno de la Rioja: <https://www.riojasalud.es/>

Gobierno de Navarra: https://www.navarra.es/home_es/

Gobierno del País Vasco: <https://www.euskadi.eus/inicio/>

Grupo Recoletas Red Hospitalaria: <https://www.gruporecoletas.com/>

Hospital Clínic Barcelona: <https://www.clinicbarcelona.org/>

Hospital Fuensanta: <https://hospitalfuensanta.com/>

Imegen: <https://imegen.es/>

Instituto Quirúrgico Lacy: <https://www.iqlacy.com/>

La vanguardia: <https://www.lavanguardia.com/>

Mejor Sin Cáncer: <https://mejorsincancer.org/>

Ministerio de Sanidad, Consumo y Bienestar: <https://www.mscbs.gob.es/>

Quirón Salud: https://www.quironsalud.es/blogs/es/conexion-cancer.categorias/ELEM_ENTRY_LABELS/CRISP-CAS

Seguros Mapfre: <https://www.mapfre.es/seguros/particulares/>

Sociedad Andaluza de Patología Digestiva: <https://www.sapd.es/>

Sociedad Española de Medicina Interna: <https://www.fesemi.org/>

Sociedad Española de Oncología Médica: <https://seom.org/>

10.2.2. Subcorpus de textos traducidos

ADAM: <https://www.adam.com/about>

American Academy of Family Physicians: <https://es.familydoctor.org/>

American Cancer Society: <https://www.cancer.org/es.html>

American College of Gastroenterology: <https://gi.org/>

American Society of Clinical Oncology: <https://www.asco.org/>

American Society of Gastrointestinal Endoscopy: <https://www.asge.org/>

Centers for Disease Control and Prevention: <https://www.cdc.gov/>

Colon Cancer Coalition: <https://coloncancercoalition.org/>

Mayo Clinic: <https://www.mayoclinic.org/es-es>

MD Anderson: <https://www.mdanderson.org/>

MedlinePlus: <https://medlineplus.gov>

MSD Manuals: <https://www.msmanuals.com/es-es/hogar>

National Cancer Institute: <https://www.cancer.gov/>

RadiologyInfo: <https://www.radiologyinfo.org>

Roche: https://www.roche.es/es_es/roche-pacientes.html

UI Health: <https://hospital.uillinois.edu/es>