

1. INTRODUCCIÓN*

* Este trabajo se ha elaborado dentro del marco del proyecto de investigación *OncoTerm: Sistema bilingüe de información y recursos oncológicos*, financiado por la DGICYT (código PB98-1342). Su objetivo es el de poner al alcance de pacientes y profesionales de la medicina y la traducción especializada información relacionada con el cáncer. El sistema consta de una ontología y una base de datos terminológica y textual en inglés y español, recursos que estarán integrados entre sí en *Internet*.

1. Introducción

1.1. Presentación

En nuestra época, la fascinación por nuevos artefactos audiovisuales convive con un atenuado culto hacia la palabra escrita, encarnada en una infinidad de textos y géneros que las sociedades humanas han legitimado como vehículos de comunicación apropiados para un fin concreto. Si bien es verdad que la historia del hombre se ha construido sobre textos orales y escritos, sin embargo, no todas las disciplinas cuyo objeto de estudio son los textos han estudiado de una forma sistemática los distintos tipos textuales reconocidos por una comunidad lingüística.

De hecho, de las disciplinas que han llegado más lejos en su análisis del texto, la crítica literaria, la lingüística y la traducción, sólo la primera saltó de planteamientos programáticos a la elaboración de taxonomías textuales con anterioridad al siglo XX. Podemos encontrar una explicación de esto en el carácter poco tangible de esta unidad, el texto, que Brown y Yule (1983:6) definen como “the verbal record of a communicative act”.

Los estudios de traducción comenzaron a considerar al texto como unidad de traducción y de equivalencia con el desarrollo de la lingüística textual en la década de los setenta. No obstante sólo en la década de los noventa se ha producido el salto de modelos traductológicos programáticos y muy generales, a modelos que pueden explicar el ejercicio concreto de la traducción en un dominio determinado.

Esto se debe en parte a que, durante muchos años, los estudios de traducción no han sabido aprovechar las aportaciones de otras disciplinas que han explorado el área de la variación textual: la lingüística, el inglés para fines específicos, la psicolingüística, la inteligencia artificial, la retórica contrastiva, la sociología de la ciencia, la lingüística de corpus y la terminología. Estas disciplinas han aportado términos como el de *modo de discurso*, *superestructura*, *macroestructura*, *género*, *tipo textual*, *registro*, *discurso* y *sublenguaje* y han intentado identificar los parámetros que los definen. Buena prueba de esta escasa colaboración interdisciplinar lo constituye el hecho de que el concepto de *género* se ha introducido de un modo débil en el área de los estudios de traducción.

Por tanto, partimos de la premisa de que los estudios sobre género pueden ayudar al traductor a desarrollar estrategias para reconocer las opciones y las limitaciones que imponen determinados tipos textuales (Trosborg 1997: viii).

En la traducción de textos biomédicos, el traductor tiene que estar familiarizado con las opciones y limitaciones relacionadas con una serie de recursos morfosintácticos, con unas convenciones retóricas y ortotipográficas y con una terminología constantemente en expansión. En estas tres facetas subyace la identificación de tipos textuales biomédicos, sin la cual, el texto traducido perdería su función comunicativa y se podrían ocasionar consecuencias indeseables. Entre estas, podemos mencionar el rechazo de un artículo por el comité editorial de una revista prestigiosa o la creación de expectativas de curación falsas, ocasionadas por la traducción incorrecta de un verbo modal.

El traductor debe conocer los recursos morfosintácticos asociados a determinados géneros. Por ejemplo, para traducir un folleto explicativo de cómo realizar una autoexploración mamaria, es necesario que el traductor sepa que, en inglés, los folletos de instrucciones utilizan el imperativo mientras que la lengua española impone una mayor variación sintáctica, y así, es frecuente encontrar además del imperativo, el presente de indicativo, el impersonal con *se* o incluso un uso incorrecto del infinitivo. También es preciso que reproduzca las convenciones tipográficas y retóricas establecidas por la comunidad médica internacional. Entre estas se pueden citar las normas de publicación de manuscritos enviados a publicaciones científicas, que especifican las secciones de que consta un artículo experimental.

El conocimiento de la terminología biomédica comprende el conocimiento de las entidades y procesos implicados en la investigación biomédica y una competencia *documentalista* para acceder a bases de datos, glosarios y a organismos de normalización dependientes de la ISO (*International Standards Organization*) como AENOR, ANSI y BSI. Asimismo requiere un conocimiento de cómo la terminología cambia en distintos registros como el propio de una publicación periódica especializada o el de una revista de divulgación o el que utiliza un médico en su consulta.

La utilización de terminología médica supone la activación de redes conceptuales de significado que hacen del texto una unidad gracias al fenómeno que se

denomina *cohesión léxica*. La cohesión léxica contribuye a la textualidad y es un parámetro sensible al tipo textual en el que se establece. De ahí la necesidad de que el traductor tome consciencia de los mecanismos cohesivos presentes en el texto origen (en adelante, TO) y el texto término (TT), de las redes conceptuales activadas mediante las unidades léxicas y terminológicas del texto, al mismo tiempo que evalúa si se ha mantenido el valor discursivo de la cohesión en el TT.

1.2. Hipótesis de trabajo y objetivos

En este trabajo se parte de tres hipótesis. En primer lugar, existe una correlación entre tipo textual y los recursos cohesivos que lo materializan. En segundo lugar, existen diferencias con respecto a la cohesión entre las lenguas inglesa y española que tienen una importante repercusión en la traducción. Por último, la cohesión léxica activada en los textos médicos refleja en gran medida la compleja red conceptual que sustenta ese campo del saber.

Para validar estas hipótesis, se proponen una serie de objetivos que se materializarán en un corpus de textos sobre oncología:

- a) Revisar la literatura sobre tipología textual en traducción incidiendo en los aspectos cognitivos, lingüísticos y sociales implicados en la producción y comprensión de tipos textuales.
- b) Revisar los estudios sobre cohesión léxica en cuanto que esta puede ser un parámetro sensible al tipo textual en el que se establece
- c) Diseñar un método de análisis de la cohesión donde se computen las relaciones cohesivas en textos completos con la ayuda de recursos informáticos.
- d) Estudiar qué diferencias cohesivas se pueden encontrar al contrastar dos tipos textuales que difieren con respecto al destinatario y a los conocimientos que este tiene sobre medicina.
- e) Estudiar si existen diferencias con respecto a la cohesión entre un texto y su traducción a otra lengua y discernir cuáles son el resultado de las convenciones retóricas inherentes al inglés y al español.
- f) Establecer un marco teórico-práctico que ayude al traductor, principalmente, al que trabaja con el inglés y el español, a sistematizar los recursos cohesivos y asociarlos a

tipos textuales diferentes. El traductor debe ser consciente de las decisiones implicadas en la traducción de distintos tipos textuales relacionados con el ámbito de la medicina de manera que consiga traducir de una forma lo más comunicativa posible. También debe ser sensible al hecho de que en algunas ocasiones la selección de terminología médica en los textos no siempre responde a la organización conceptual de ese dominio. Conociendo cómo se articula la cohesión en los textos y los factores retóricos que inciden sobre la misma, el traductor puede discernir por qué se han activado unos términos, y no otros de esa organización conceptual.

- g) Aplicar la metodología de la lingüística de corpus y el concepto de cohesión en la detección de la representación conceptual subyacente a los textos y estudiar las implicaciones que esto puede tener en terminología, documentación y traducción.

1.3. Ubicación del trabajo en un marco interdisciplinar

Todos estos objetivos se van a contemplar desde una perspectiva interdisciplinar que aúne las aportaciones sobre tipologías textuales provenientes de los *estudios de traducción* y la *lingüística*. También recogerá los estudios sobre *género* realizados desde la *lingüística aplicada* y algunas de sus subdisciplinas como la lengua para fines específicos (Bhatia 1993; Eggins y Martin 1997; Nwogu 1997; Swales y Feak 1994), la retórica de contrastes (Connor 1996; Connor y Kaplan 1987) y la textología contrastiva. (Hartmann 1996).

El inglés para fines específicos arroja luz sobre cómo se estructuran los principales géneros médicos en inglés y sus rasgos lingüísticos. La retórica de contrastes identifica las diferencias lingüísticas y retóricas entre textos escritos en la lengua materna y en una segunda lengua. La textología contrastiva propone una integración de la lingüística de contrastes, el análisis de discurso y la lingüística de corpus en el estudio de textos paralelos en más de una lengua. Relacionada con estas disciplinas está también la *sociología de la ciencia* (Bazerman y Paradis 1991; Myers 1990), que indaga sobre el conocimiento cultural y social que determina y estructura retóricamente los textos científicos, a pesar de que estos aluden a una realidad objetiva independiente del investigador.

Asimismo, se revisarán algunas nociones de *psicolingüística* (Baddeley 1992, Bell 1991; Faber 1998a) e intentaremos explicar el papel de la cohesión en el procesamiento de la información y del lenguaje.

La *terminología* (Budin y Wright 1997; Cabré 1999a; Somers 1996; Sonneveld y Loening 1993) también será un claro referente porque es la disciplina que estudia los términos de una materia especializada, cómo recopilarlos, los mecanismos de normalización de los términos y cómo estructurar las relaciones conceptuales de ese campo de especialidad. Dentro de la terminología se adoptará una aproximación descriptiva en la que se gestione la información de una forma orientada a la traducción. Esto es lo que Wright (1999) denomina *translation-oriented terminology management*. Dentro de esta, no nos interesa tanto la terminología puntual que suscita la traducción de un texto especializado sino el intento de sistematizar recursos y aplicaciones útiles para el traductor. Esta aproximación es la que Cabré (1999b) denomina terminología *para* la traducción.

Por último, como se pretende la aplicación de recursos informáticos al estudio de la cohesión, nos vamos a beneficiar de la metodología de la *lingüística de corpus* y de algunas aportaciones provenientes de la *lingüística computacional* y la *ingeniería de conocimiento* (Morris y Hirst 1991, Ellman 1998). Esta última elabora ontologías, es decir, repositorios estructurados de conocimiento, en las que se organizan los conceptos que articulan un determinado campo del saber. La lingüística de corpus (Aijmer y Altenberg 1991; Baker *et al.* 1993; Biber 1993, 1998; Biber y Finegan 1991; Butler 1992; McEnery y Wilson 1996; Sinclair 1991; Stubbs 1996; Svartvik 1992) aporta respuestas a algunos de los dilemas que se plantean a la hora de recopilar un corpus en soporte electrónico *representativo*:

- criterios de selección del corpus (fecha de publicación, textos completos, variaciones dialectales, géneros, tema, etc.)
- tamaño
- formato de anotación más apropiado para que el corpus sea reutilizable
- programas informáticos que elaboren listas de frecuencia y líneas de concordancia y que etiqueten los corpórea para poder extraer de ellos información sintáctica, semántica, pragmática y suprasegmental.

- distinción entre *córpore comparables* y *córpore paralelos* en soporte electrónico

1.4. Breve resumen de los capítulos

1.4.1. Primera parte: revisión bibliográfica

La primera parte, dedicada a la revisión de bibliografía, girará en torno a las tipologías textuales, la cohesión y la lingüística de corpus. En el capítulo 2 se parte de un modelo cognitivo de la traducción para poder explicar los elementos de la competencia traductológica que intervienen en la comprensión y producción de tipos textuales. Se recalca la idea de que en la traducción participan los mismos sistemas neurológicos que intervienen en el procesamiento de la información. Estos ejecutan procesos interactivos y no-lineales que suponen una continua toma de decisiones. A continuación, se aborda el concepto de tipología textual, por una parte, desde aproximaciones funcionales y textuales como las de Reiss, House, la teoría del *skopos* (Vermeer), Sager, Nord y Hatim y Mason, y por otra, desde aproximaciones más eclécticas que, aún reconociendo la importancia de los aspectos pragmáticos en la traducción, hacen un gran hincapié en la cognición, como la de Werlich, Neubert, Snell- Hornby y Kussmaul.

En el capítulo 3, se aclara el significado de los términos *tipo textual*, *registro* y *género* y se incide en la idea de que el texto es un todo organizado. Se describen algunos de los tipos textuales médicos más usuales.

En capítulo 4, se revisan los estudios sobre cohesión y se presta una especial atención a la metodología propuesta por Hoey (1991, 1995) y, de todos los recursos cohesivos, a la *cohesión léxica*. También se incluye una aproximación proveniente del campo de la lingüística computacional, la de Morris y Hirst (1991), que hablan de la cohesión en términos de *cadena léxica*. Se comentan algunas herramientas informáticas para la detección automática de la cohesión y se destacan aquellas que reconocen relaciones conceptuales. Nuestro interés por el aspecto semántico de la cohesión nos llevará a revisar la propuesta de la *Teoría Sentido Texto* (Mel'cuk *et al.* 1984) en relación con las representaciones semánticas de los textos. Por último, se sugieren algunos elementos que interactúan con la cohesión, tal y como el tipo textual, el conocimiento que comparten el emisor y el receptor de un texto, las dificultades

comunicativas que impone el contexto situacional, el grado de explicitación y convenciones retóricas que dominan las lenguas, y el proceso traductor en sí.

El capítulo 5 trata sobre la relevancia que está adquiriendo la lingüística de corpus en la traducción y en terminología y de la necesidad de distinguir entre corpórea paralelos y comparables. Se citan algunas aproximaciones que se pueden tener en cuenta en la traducción biomédica.

1.4.2. Segunda parte: metodología

En la segunda parte se conjugan la mayoría de los modelos sobre cohesión presentados en la primera parte y se propone una doble metodología para el estudio de la cohesión que tiene en cuenta la naturaleza del fenómeno cohesivo, que sintetizamos en el capítulo 6. En este capítulo también se proponen categorías para describir las instancias de cohesión léxica y se sistematizan las posibles relaciones entre elementos cohesionados en una taxonomía de funciones léxicas cohesivas. La primera metodología, llamada *análisis intensivo*, es la que se sugiere en el capítulo 7 para el análisis de textos breves o fragmentos de texto y destaca la faceta interoracional de la cohesión. Se aplicará en el capítulo 10. El capítulo 8 se centra en el *análisis extensivo* de la cohesión, que parte de una concepción de la cohesión como un fenómeno tanto interoracional como intraoracional. En este capítulo se presentan los módulos de la herramienta informática que nos va a ayudar a detectar las *cadenas léxicas* del texto y cómo se plasmarán en papel los datos que se extraen del ordenador. Por último, en el capítulo 9 se detalla el corpus de textos sobre oncología que hemos manejado en esta investigación.

1.4.3. Tercera parte: resultados

Los resultados se exponen en la tercera parte de la tesis. Los obtenidos con un análisis intensivo van en el capítulo 10 y dejan entrever algunas diferencias cohesivas entre textos en inglés y español, y entre textos originales y traducidos. El capítulo 11 presenta los resultados del estudio de la cohesión léxica de forma extensiva en dos tipos textuales diferentes sobre exactamente el mismo tema: el tratamiento del cáncer de pulmón. El primer tipo textual lo constituyen textos escritos por oncólogos para profesionales de la salud y publicados en Internet; el segundo es la explicación del mismo tema al paciente y a sus familiares. En el análisis, después de describir los resultados obtenidos

automáticamente, se modifican estos para ajustarlos mejor a las relaciones cohesivas que detecta el analista humano. En todo caso, se comparan los resultados obtenidos de una y otra forma.

1.4.4. Cuarta parte: discusión y conclusiones

En el capítulo 12 comienza la parte de *Discusión y conclusiones* de la tesis y se establece la relación que existe entre las cadenas léxicas reconocibles en un texto y la activación de conceptos que se produce durante la lectura. Se comentan las diferencias conceptuales y terminológicas más significativas en las cadenas léxicas correspondientes a textos para especialistas y para pacientes. Las cadenas analizadas están vinculadas principalmente con los conceptos TREATMENT, PATHOLOGY, RESEARCH, OUTCOME OF TREATMENT, LOCATION IN THE HUMAN BODY, DIAGNOSIS, PATIENT, STAGING y SIZE / QUANTITY.

En el capítulo 13 se buscan las expresiones que contribuyen a crear cohesión en textos para lectores legos y, gracias a la cohesión, aumentan los conocimientos científicos de dichos lectores. Estas expresiones permiten que términos ininteligibles para el lector lego aparezcan en el texto cerca de lexemas de la lengua general que explican su significado.

En el capítulo 14 se indaga sobre el papel que la frecuencia de determinadas unidades léxicas en un texto o en un corpus puede jugar, por una parte, en la configuración de cadenas léxicas, y por otra, en el reconocimiento de unidades fraseológicas. Los datos que se comentan se obtienen mediante listas de frecuencia y líneas de concordancia. Se reflexiona sobre cómo la distribución de *palabras forma* (*function words*) en un corpus nos puede orientar sobre las características del mismo.

En el capítulo 15, se detectan las diferencias cohesivas entre textos originales en inglés y su traducción al español y se señalan aquellas que responden a las características del inglés y del español, y al proceso traductor, tal y como aparecen en traducciones al español de textos sobre oncología en inglés. Las conclusiones quedan reflejadas en el capítulo 16.