

Estrategias terminológicas para la resolución de traducciones de términos médicos

GRACIELA ROSEMBLAT

*Biblioteca Nacional de Medicina,
Institutos Nacionales de la Salud, Bethesda,
MD, EE.UU.*

INTRODUCCIÓN

La lengua inglesa ha venido a ocupar un lugar de privilegio en la globalización lingüística, y se ha convertido en una especie de *lingua franca* para la comunicación y transmisión de información científica. Es el idioma elegido para las conferencias científicas internacionales y las publicaciones técnicas de alto nivel. Académicos e investigadores de países no angloparlantes deben enfrentar una gran presión institucional y profesional para presentar sus escritos en publicaciones internacionales prestigiosas (en su gran mayoría en inglés y con un alto factor de impacto), y de esta manera diseminar sus teorías y dar a conocer sus trabajos, para obtener un mayor reconocimiento. Esta situación, que data de varias décadas y ha sido descrita en numerosos trabajos [1, 2, 3, entre otros], se observa en todos los países no angloparlantes y en casi todos los campos del conocimiento, incluso en temas de medicina o relacionados con la salud y se refleja en la Web. La economía de los países de habla inglesa permite destinar enormes presupuestos para lograr avances técnicos y científicos. De ahí que esos países lideren las investigaciones en esos campos y publiquen sus resultados, a pesar del sostenido aumento participativo de los demás países en la producción científica mundial.

En el campo de la salud en particular, y especialmente en los Estados Unidos, el creciente número de internautas de habla hispana en los últimos años ha dado lugar a la creación de sitios Web y recursos de información bilingües y en español. El objetivo es brindar apoyo a los hispanohablantes facilitando el acceso a la información en línea en su propio idioma. Esta bienvenida apertura trae aparejada otra serie de problemas. En muchos casos aunque los contenidos se presentan en español, la falta de una redacción pulida y coherente, el uso de frases forzadas o terminología sin traducir (y sin ninguna explicación al respecto) plantean serias dudas sobre la procedencia de la información y, en consecuencia, la fiabilidad de esos sitios Web. La terminología exacta en temas de salud es fundamental para las búsquedas y recuperación de la información e indización de documentos dentro del marco de la comunicación científica y especializada, en todos los idiomas. Con el advenimiento de nuevos tratamientos, compuestos farmacológicos y procedimientos terapéuticos, en su mayoría descubiertos o creados en los Estados Unidos y, por tanto, originados lingüística y conceptualmente en inglés, el rezago terminológico en otros idiomas es inevitable. En consecuencia, muchos términos médicos sólo están fácilmente disponibles en inglés, a pesar de los excelentes diccionarios y recursos multilingües que o no brindan suficiente cobertura en campos específicos, o no están actualizados. Esta ponencia presenta varios métodos para recuperar terminología en español utilizada por científicos y profesionales involucrados en el proceso de la comunicación científica y en los criterios para su validación.

MARCO CONTEXTUAL

Las estrategias que se describirán a continuación surgieron a partir de un proyecto de investigación de la Biblioteca Nacional de Medicina (parte de los Institutos Nacionales de la Salud de los Estados Unidos) para proporcionar información en español acerca de la investigación clínica, y facilitarle a los usuarios de habla hispana el acceso a información en línea sobre los ensayos clínicos asentados en los registros de dicha institución. Para ello se desarrolló un sistema de recuperación de información y visualización en español: las consultas en espa-

ñol eran traducidas al inglés para recuperar los ensayos clínicos pertinentes de esa base de datos, originalmente escritos en ese idioma. Como resultado de la consulta, los documentos o ensayos clínicos recuperados se presentaban en un formato abreviado, en el cual los títulos eran traducidos manualmente al español, al igual que el lenguaje controlado correspondiente a los tratamientos y trastornos médicos objeto de estudio en cada ensayo [4].

En ese proyecto de investigación, además de la traducción de terminología especializada referente a técnicas o terapias inéditas, la interpretación de las siglas o acrónimos representó un gran desafío, dado que los proveedores de datos no tienen la obligación de incluir las expresiones completas que corresponden a dichas abreviaturas. Es común que los acrónimos den lugar a ambigüedades al referirse a diferentes conceptos. En algunos casos se utilizan como una manera ingeniosa de acortar el título de un ensayo clínico, formando una palabra relacionada con el texto, simbolismo prácticamente irreproducible en la traducción a otros idiomas.

Un recurso incomparable que incluye tanto traducciones al español como expansiones de siglas es el Unified Medical Language System® (UMLS® [5]). La fuente terminológica comprendida en el Metathesaurus, parte del UMLS®, constituye el recurso léxico más exhaustivo para el campo médico. En esa base de datos, una sigla y su expansión en la expresión completa son representados como sinónimos o variantes de un mismo concepto, junto con las correspondientes entradas multilingües o términos traducidos. Los términos MeSH [6] y SNOMED CT son las principales fuentes de datos para las entradas en español del UMLS®. MeSH, o su equivalente en español, los Descriptores en el Campo de la Salud (DeCS), es considerado como un lenguaje controlado en el sentido que se utiliza para la indización de textos, lo cual facilita luego su recuperación. Además, es una forma de estandarizar el vocabulario técnico. No obstante, pese a su gran cobertura, muchas entradas en el UMLS® carecen de traducciones al español: en la versión 2007AB, de un total de 1.436.182 conceptos, sólo 342.123 (24%) tienen al menos un sinónimo en español (1.454.317 términos en español frente a 3.547.627 en inglés). Esta falta de cobertura es genera-

lizada: a pesar de los grandes avances informáticos, las bases de datos bilingües con glosarios especializados y las aplicaciones de software para la recuperación automática de textos multidisciplinares multi-lingües no son suficientemente exhaustivas para el ámbito biomédico.

Dentro del marco contextual de este proyecto de investigación, la metodología desarrollada para el tratamiento terminológico está basada en técnicas de búsqueda en Internet eficaces y principios lingüísticos para obtener traducciones correctas, no fácilmente disponibles, al español, de terminología médica especializada en inglés. La estrategia combinada que se describirá se presta a la validación y convalidación terminológica, y lleva a la recuperación de las mejores opciones de traducción o posibles traducciones. La estrategia está dirigida tanto a traductores especializados, lexicógrafos y bibliotecarios como a científicos y representantes de las ciencias de la comunicación que se dedican al desarrollo de contenidos de páginas Web, especialmente en el ámbito de la medicina y temas relacionados con la salud. La metodología puede aplicarse además en otras disciplinas y en otros idiomas ajustando los principios lingüísticos a éstos.

METODOLOGÍA

Se describirán tres métodos diseñados para su utilización de un modo iterativo y complementario, que pueden combinarse en una sola estrategia por etapas para obtener traducciones al español a partir de las expresiones en inglés. Estos métodos consisten en:

- 1) el aprovechamiento de textos ya alineados o *corpus* paralelos inglés-español, empleados con frecuencia en la traducción automática;
- 2) la utilización de señales semánticas y del contexto conceptual, semántico y sintáctico para cubrir los vacíos terminológicos en español;
- 3) la aplicación de conocimientos lingüísticos específicos para el español, ateniéndose a las características morfosintácticas propias de la lengua.

Antes de poner en práctica estos métodos, un paso preliminar de gran utilidad consiste en realizar búsquedas manuales de la expresión original en inglés en las fuentes primarias; es decir, en hacer búsquedas monolingües en textos en inglés, lo cual permite verificar:

- a) la inexistencia de posibles errores ortográficos, tipográficos o de uso: la ortografía correcta puede corresponder a una expresión para la cual ya existe una traducción al español y no procede continuar con la búsqueda.
- b) si la expresión en cuestión es una marca registrada y no se ha incluido el símbolo correspondiente. Un nombre propio o una marca registrada debe emplearse en la traducción tal cual aparece en el idioma original, con el agregado del símbolo correspondiente, y no procede continuar con la búsqueda.
- c) el contexto léxico-semántico de la expresión que da la pauta de su significado y uso correcto, además de la visualización de otros términos con los cuales se utiliza generalmente la expresión y que también constituyen un recurso en búsquedas puntuales para refinar los patrones lingüísticos en los que se utiliza, derivando contextos más limitados dentro del ámbito especializado;
- d) el concepto exacto denotado por la expresión o sinónimos de la misma, dado que el reconocimiento del significado es clave para una indización correcta [7]. Además de analizar el contexto semántico y el uso de una expresión, se puede también obtener el significado del concepto ingresando “define:xxx” en la casilla de búsqueda (buscador Google®), reemplazando “xxx” por la expresión deseada.
- e) el campo del conocimiento científico, que puede presentar variaciones en el uso de una misma expresión. Para determinar el campo apropiado se pueden limitar las búsquedas a ciertos encabezamientos de páginas Web, o recurrir a sistemas que permitan la detección automática o semiautomática de términos o candidatos a términos¹ [9].

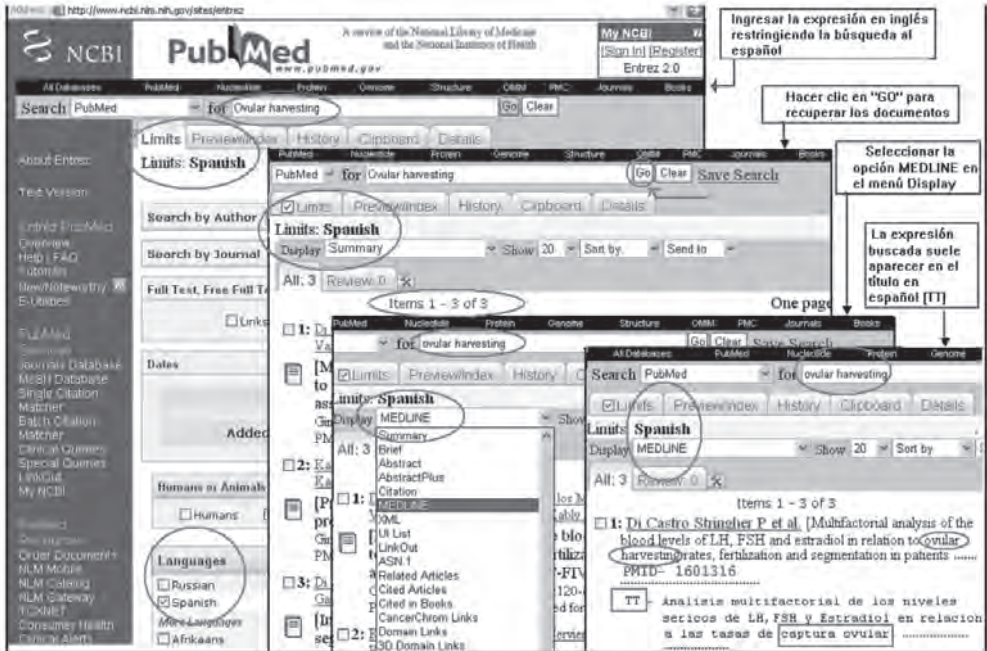
¹ Véase, por ejemplo, la herramienta YATE [8] que combina estrategias probabilísticas y lingüísticas.

Según los resultados de este paso preliminar, se procede o no a las técnicas de búsqueda, que se explicarán a continuación.

1) *Aprovechamiento de textos bilingües paralelos* existentes en línea, escritos en el idioma original (que puede ser tanto inglés como español) por especialistas en la materia, con su traducción correspondiente al otro idioma. Estos textos constituyen un inestimable recurso terminológico, aun cuando solamente una parte del documento presente traducción, que es generalmente la parte que corresponde al resumen o al título. El texto no necesariamente está alineado oración por oración o párrafo por párrafo (texto paralelo), método muy utilizado en la traducción automática, sino que puede presentar una alineación más local y conceptual, o temática [10], también llamada texto o *corpus* comparable [11].

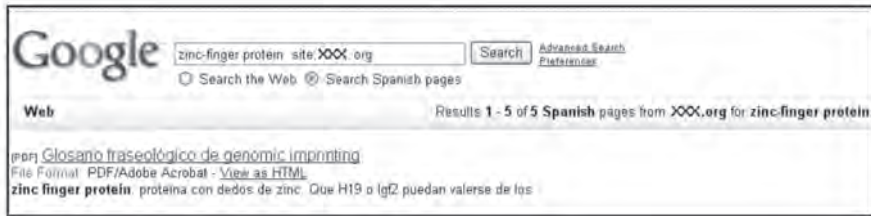
La búsqueda de estos corpus paralelos puede realizarse en sitios Web especializados u otros generales. En el primer caso, si existiera la posibilidad de que la traducción de una expresión en inglés pueda aparecer en publicaciones en español, se puede hacer una búsqueda en PubMed [12] para ver qué términos utilizan los investigadores de habla hispana en las fuentes primarias para referirse a la expresión original. La búsqueda en PubMed comienza ingresando la expresión en inglés en la casilla de búsqueda. En la pestaña o ficha denominada Limits (Límites), se puede restringir la búsqueda a los documentos originalmente escritos en español seleccionando "Spanish" en el menú de idiomas (Languages) (Fig. 1). Si hubiera textos recuperados como resultado de la búsqueda, al seleccionar la opción 'MEDLINE' en el menú de visualización (Display), se obtendrá la referencia completa de los artículos, con los títulos originales en español en la etiqueta [T], que a menudo incluyen la traducción de la expresión en inglés consultada, u otros términos relacionados con la misma (sinónimos o contexto léxico-semántico).

Figura 1. Búsqueda en PubMed de una expresión en inglés restringida al español



Además de las búsquedas avanzadas en PubMed restringidas a publicaciones en español, se puede aplicar el mismo principio para restringir una búsqueda avanzada en textos en español (*Search Spanish pages*) a algunos de los sitios Web especializados y de alta fiabilidad que ofrecen resúmenes o glosarios en ambos idiomas (inglés y español). Para ello, utilizando el buscador Google®, se indica el sitio Web deseado en la casilla de búsqueda a continuación de la expresión buscada, como en el siguiente ejemplo:

Figura 2. Búsqueda avanzada de una expresión en inglés en textos en español restringida a un sitio determinado, utilizando el buscador Google®



La expresión en inglés, ya sea completa o en partes, también puede buscarse en fuentes no específicas (más generales) disponibles en español, que proporcionan resúmenes o textos en ambos idiomas. Aun cuando no estén alineados, los resúmenes pueden dar lugar a la recuperación de posibles traducciones en español para su posterior validación en los textos monolingües. Las posibles traducciones pueden identificarse por la semejanza morfológica en ambos idiomas, por la correspondencia semántica, o por una comprensión clara del concepto denotado, junto con los correspondientes microcontextos léxico-semánticos (colocaciones típicas) y sintagmáticos en ambos textos.

2) Señales semánticas y contextuales en casos de vacíos terminológicos

Cuando las búsquedas para obtener traducciones o posibles traducciones no son exitosas y no se recuperan documentos en español, o no se puede determinar fehacientemente la existencia de una traducción estándar o aceptada por la comunidad científica, resulta útil recurrir a un análisis del contexto léxico-semántico de la expresión en inglés. Este nivel de fiabilidad no se determina por un número mínimo aceptable de apariciones en Google® establecido a priori, sino por una conjunción de factores: origen de las fuentes primarias en cuanto a país (hispanohablante) e instituciones confiables, número de apariciones en textos especializados (PubMed, Google®) y el sentido común y conocimientos del traductor o terminólogo. A partir de este análisis del contexto léxico-semántico pueden surgir varias opciones

de traducción al español. Una estrategia consiste en hacer búsquedas en las fuentes primarias en español de los términos que aparecen regularmente en el contexto de la expresión en inglés cuya traducción nos interesa. Aunque esta búsqueda recupera documentos en lugar de expresiones, la terminología utilizada para la traducción de la expresión en inglés puede identificarse si se tiene una comprensión del concepto incorporado en la semántica de la expresión original, de la mano de un minucioso análisis léxico y semántico del microcontexto, el cual difiere en cada caso y es determinado por el terminólogo-traductor, o las colocaciones típicas en que aparece la expresión, que incluye los términos vecinos, los modificadores directos o indirectos y los complementos, además de las relaciones semánticas y conceptuales entre todos estos. El microcontexto puede ir desde el sintagma nominal hasta el párrafo según la granularidad descriptiva del texto y la introducción de conceptos afines.

Por ejemplo, una búsqueda en PubMed de los títulos en español ([TT]) que contengan una traducción de la expresión en inglés ‘*second-look (surgery)*’, definida como “cirugía realizada después del tratamiento primario...”,² no obtuvo resultados. En cambio, una búsqueda avanzada en Google® obtuvo los siguientes resultados: (*cirugía secundaria*) (100 citas), (*cirugía de segunda exploración*) (15), (*de segunda inspección*) (6), (*de segunda intención*) (3), (*de segunda observación*) (1). A pesar de que el término correspondiente a ‘look’ no aparece en ninguna de las posibles traducciones para la expresión ‘second-look’ (surgery), los contextos de las expresiones recuperadas muestran una clara correspondencia semántica con el concepto en inglés. En este caso, las opciones de traducción incluyen:

- Utilizar una de las expresiones en español que más textos recuperan e incluir a continuación la expresión original en inglés entre paréntesis: *cirugía secundaria* (second-look surgery);
- Utilizar dos de las expresiones halladas en los textos recuperados en español, la segunda entre paréntesis a modo de aclaración

2 Dictionary of Cancer Terms, National Cancer Institute, NIH

de la primera: *cirugía secundaria (cirugía de segunda exploración)*;

- Utilizar sólo la expresión en español que más textos recupera, por ejemplo, *cirugía secundaria*, siempre que la fiabilidad de la fuente primaria pueda comprobarse mediante la institución y el país de procedencia, la calidad del resto del texto o la consonancia con expresiones similares que sean más conocidas, entre otros criterios.

Los posibles traducciones pueden validarse examinando su uso en textos monolingües en español. La recuperación de un número alto de textos de fuentes primarias fidedignas en español brinda buena validación para una traducción, especialmente cuando se utiliza como consulta de sitios Web principalmente originados en países latinoamericanos o en España. Para acotar la búsqueda en Google® a documentos de ciertos países, se debe incluir en la casilla de búsqueda el país de origen deseado, a continuación de la expresión que se necesita consultar, de modo similar a la restricción de la búsqueda a fuentes específicas, como en la Figura 2. En el siguiente ejemplo, la expresión “vía de señalización de Notch” tal como está escrita será buscada en sitios exclusivamente de España (.es):

Figura 3. Búsqueda avanzada de una expresión en inglés en textos en español restringida a sitios de España, utilizando el buscador Google®



Para los vacíos terminológicos de expresiones en inglés que no tienen traducción correspondiente en español, una explicación corta, puntual o una paráfrasis en español de la expresión buscada puede

utilizarse a modo de traducción con el fin de proporcionar el significado, dado que no es admisible incluir una expresión en inglés —que puede ser desconocida— en medio de una narrativa en español, sin antes introducir esta expresión mediante una descripción explicativa. Por ejemplo, actualmente no existe en español una traducción para el término ‘mindfulness’,³ comúnmente utilizado en psicología en la expresión ‘mindfulness-based therapy’. Muchos textos/documentos, aun los provenientes de países hispanohablantes, utilizan el término en inglés. Sería difícil que hubiera un cierto consenso entre diferentes profesionales acerca de cómo explicar el concepto en español, ya que puede hacerse con mayor o menor detalle. Algunas opciones son:⁴

- ‘Técnica terapéutica utilizada en psicología (llamada “mindfulness” en inglés)’
- ‘Técnica terapéutica que incluye observación activa, meditación y aceptación (conocida como “mindfulness” en inglés)’
- ‘Procedimiento terapéutico que consiste en tener plena conciencia de sí mismo de una forma deliberada y neutral, con observación activa, meditación y aceptación (conocido como “mindfulness” en inglés)’

3) Aplicación de conocimientos lingüísticos específicos para el español

De más está decir que los conocimientos lingüísticos y las características morfosintácticas propias de una lengua son una valiosa herramienta que ayuda a inferir posibles traducciones, especialmente en casos de variantes terminológicas que denotan un mismo concepto. Esta herramienta cumple un rol fundamental ante la falta de una traducción estándar (o estandarizada) que pueda ser validada mediante búsquedas en textos monolingües en fuentes primarias en español. Algunos campos lingüísticos en los que pueden aplicarse estos tipos

3 The practice whereby a person is intentionally aware of his/her thoughts and actions in the present moment applied to bodily actions and the mind's own thoughts and feelings. [Definición de ‘mindfulness’: www.wikipedia.org]

4 Opciones de traducción, G.R.

de conocimiento son: (3a) los procesos morfológicos que tienen lugar en la derivación terminológica, (3b) los sintagmas nominales en inglés compuestos de dos sustantivos y (3c) las cuestiones de uso de la lengua.

3a) Los sufijos utilizados en la morfología derivativa para la formación de las palabras y las categorías gramaticales a menudo le agregan al núcleo semántico de la raíz la información sintáctica, específica para cada idioma. Suele suceder que las correspondencias en diferentes idiomas de muchos términos médicos [13] y comunes sólo difieran en su terminación [14]. Por ejemplo, el sufijo adjetival inglés ‘-al’ puede corresponder a diversos sufijos adjetivales en español, uno de los cuales es también ‘-al’ (Cuadro 1). Esto último puede dar lugar a generalizaciones erróneas en las traducciones del inglés al español, a pesar de que en ciertos casos, con el tiempo y por imposición del uso, el anglicismo termina siendo aceptado como sinónimo de la expresión que corresponde en español, como es el caso del término inglés “viral” y las traducciones ‘vírico’ o ‘viral’ en español, o “aneurysmal” y ‘aneurismático’ o ‘aneurismal’ en español.

Cuadro 1. Terminaciones adjetivales en inglés y en español

INGLÉS	ESPAÑOL
-al	-al
distal, spinal, vaginal	<i>distal, espinal, vaginal</i>
mitral, transforaminal	<i>mitral, transforaminal</i>
monoclonal, hepatorenal	<i>monoclonal, hepatorenal</i>
	-ico
paroxysmal, surgical	<i>paroxístico, quirúrgico</i>
liposomal, menopausal	<i>liposómico, menopáusico</i>
	-ivo
educational, informational	<i>educativo, informativo</i>
	-oso
parenchymal, variceal	<i>parenquimatoso, varicoso</i>
mesenchymal	<i>mesenquimatoso</i>
	-ario
embryonal, diaphyseal	<i>embrionario, diafisario</i>
	-ano
carpal, tarsal	<i>carpiano, tarsiano</i>
retinal, cranial	<i>retiniano, craneano</i>
	-dor/-tor
explorational, reductional	<i>explorador, reductor</i>

3b) El proceso de formación de sintagmas nominales compuestos de dos sustantivos en inglés es inespecífico con respecto a la relación entre los sustantivos que lo componen, mientras que el español lexicaliza la relación entre ambos. Las traducciones de estos compuestos a menudo se expresan revirtiendo el orden de los sustantivos e insertando una preposición entre ellos. El traductor o lexicógrafo debe seleccionar la preposición que indica y especifica la relación implícita en el sintagma en inglés, y es idiomática en español (Cuadro 2). En algunos casos existe la opción entre dos preposiciones y ambas pueden ser igualmente válidas, como por ejemplo: *trastorno de ansiedad / trastorno por ansiedad* ('anxiety disorder');

Cuadro 2. Sintagmas nominales de dos sustantivos en inglés y en español

INGLÉS	ESPAÑOL
macrolide antibiotic	<i>antibioticoterapia con macrólidos</i>
interferon treatment	<i>tratamiento con interferón</i>
anodal stimulation	<i>estimulación por corriente anódica</i>
mercury poisoning	<i>intoxicación por mercurio</i>
magnesium oxide	<i>óxido de magnesio</i>
transmission condition	<i>condición para la transmisión</i>

Otros sintagmas nominales compuestos de dos sustantivos en inglés se traducen al español mediante un complemento adjetival pos-modificador del sustantivo núcleo:

Cuadro 3. Sintagmas nominales de dos sustantivos en inglés, sustantivo + sintagma adjetival en español

INGLÉS	ESPAÑOL
blood disorders	<i>trastornos hematológicos</i>
telephone intervention	<i>intervención telefónica</i>
telemedicine intervention	<i>intervención basada en telemedicina</i>

3c) Los idiomas difieren en el potencial léxico de las palabras cognadas, y estas diferencias afectan la traducción. Por ejemplo, el inglés

tiene pares de palabras relacionadas que se refieren a los procesos y sus resultados, como es el caso de ‘neoplasia’-‘neoplasm’. El español tiene las mismas palabras cognadas (*neoplasia* - *neoplasma*), pero la primera se usa generalmente tanto para el proceso como para el resultado. Por consiguiente, en español el término ‘*neoplasia*’ se usa como traducción del vocablo inglés ‘neoplasm’. Otros ejemplos son ‘sonogram’ - *ecografía*; ‘angiogram’ - *angiografía*; ‘arteriogram’ - *arteriografía*.

En el caso del español, las variaciones dialectales entre los diferentes países cobran mayor relevancia por cuanto atañen también a conceptos básicos en cuestiones de medicina. El ejemplo más común está dado por los diferentes vocablos para denotar las partes del cuerpo. Estas variaciones dialectales agregan otro factor de dificultad a las conocidas variantes terminológicas existentes en cada idioma (sinónimos, paráfrasis, polisemia, entre otras). El consenso terminológico está dado por el uso y la adopción. Cuando no existe consenso terminológico, dos o más variantes terminológicas pueden utilizarse para un mismo concepto, según el país o área geográfica de procedencia. Por ejemplo, ‘pelvic floor’ se traduce tanto como *suelo pelviano* (España) o como *piso pélvico* (América Latina) sin que ello implique una diferencia en cuanto a su significado.

ESTRATEGIA COMBINADA

La estrategia general que se propone consta de la aplicación de los tres métodos descriptos, que son complementarios y pueden utilizarse para la validación y convalidación terminológica. La secuencia para derivar la traducción más estándar existente en español se indica en los siguientes pasos, en el orden en el cual se presentan, a partir de la expresión en inglés:

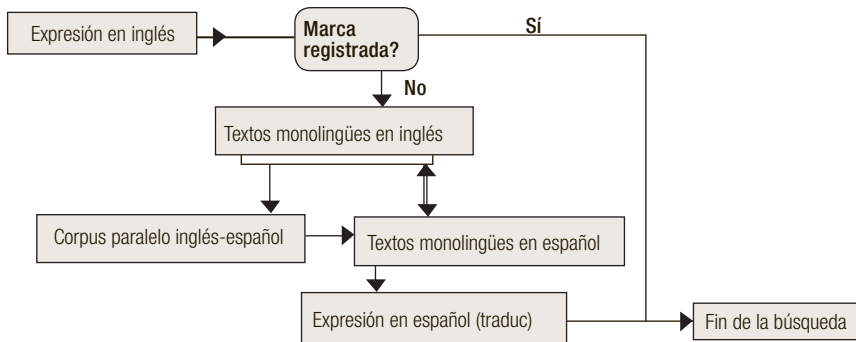
- Paso 1. Búsqueda en fuentes primarias en inglés (monolingües) para confirmar la ortografía, el significado y el contexto de uso de la expresión en inglés. Si la expresión es una marca registrada, la búsqueda concluye aquí. Este es el caso, por ejemplo, de ‘Lupron therapy’ [*Tratamiento con Lupron*®].
- Paso 2. Búsqueda de la expresión en inglés en fuentes en español, para identificar textos paralelos en los resúmenes o los títulos.

los, o corpus paralelos o equivalentes (comparables). Por ejemplo, 'maxillary sinus lift' [*elevación del seno maxilar*]⁵. Las expresiones recuperadas en español siempre deben validarse en textos monolingües en español (Paso 3).

- Paso 3. Por último, búsqueda de las posibles traducciones (expresiones recuperadas) en fuentes primarias en español (monolingües), para confirmar la ortografía, el significado y el contexto del uso en este idioma. Además del número de fuentes que utilizan una u otra posible traducción (frecuencia de uso), las fuentes confiables ya conocidas por el profesional y la procedencia de un país hispanohablante confieren una innegable señal de peso que sirve de guía para la decisión final del lexicógrafo o traductor.

Si el paso 2 no arroja ninguna traducción o posible traducción debido a la falta de corpus paralelos que contengan la expresión en cuestión, o por vacíos terminológicos en español, tendrán lugar varias reiteraciones de los pasos 1 y 3 hasta que los criterios lingüísticos sean satisfechos. Puede resultar útil recurrir también a la traducción automática para obtener algunas pautas de búsqueda, por paralelismo con otras traducciones en las cuales se emplea la expresión o expresiones objeto de búsqueda. El paso 3 siempre será el último, dado que en todos los casos las traducciones deben validarse en textos monolingües en español, como se muestra a continuación:

Figura 4. Secuencia de los pasos. La flecha doble representa una o más reiteraciones.



5 <http://www.scielo.isciii.es/pdf/medicor/v10n4/09.pdf> [02/01/2007]

EVALUACIÓN Y PRUEBA DE CONCEPTO

Para determinar la exactitud y eficacia de la estrategia descrita, se realizaron dos tipos de evaluaciones de las traducciones en el trabajo de investigación anteriormente mencionado, obtenidas mediante la aplicación de la metodología que depende de diferentes técnicas manuales utilizadas de manera complementaria, con validación y convalidación. Una evaluación preliminar de los resultados de esta metodología fue realizada por un traductor médico experto, inglés-español, sobre la base de 200 títulos de los ensayos clínicos, de un promedio de 25 palabras cada uno, y 300 expresiones de lenguaje controlado utilizadas en diferentes ensayos. Los resultados demostraron inequívocamente la validez de la estrategia, con la calificación más alta (traducción impecable) en el 95.4% de los casos. Otra evaluación de diferente carácter fue llevada a cabo por los participantes en un estudio formal de utilización [15] sobre el proyecto citado, quienes a través de las observaciones positivas sobre los títulos y expresiones traducidos al español brindaron validación externa de la eficacia de la metodología presentada, desde el punto de vista de los usuarios respecto de la información aportada.

CONCLUSIONES

Si bien el conocimiento del dominio es indudablemente un gran recurso, también se requiere competencia lingüística y del contexto de uso para obtener o derivar la mejor traducción posible. Dejando de lado la controversia acerca de si verdaderamente existe una traducción perfecta, los recursos metodológicos y lingüísticos presentados en este trabajo no constituyen una lista exhaustiva, sino que se han incluido para ilustrar un sistema que puede asimismo aplicarse en otros idiomas u otros campos del conocimiento para obtener los resultados deseados.

Cabe mencionar aquí las conocidas advertencias acerca de la falta de autoría oficial y revisión en Internet, para todos los idiomas. No toda la información en español es confiable y cada usuario debe determinar la fiabilidad de las fuentes. Cuando las búsquedas se realizan

de manera sistemática en un ámbito especializado, se llega a un buen grado de conocimiento de cuáles sitios Web son confiables y valederos, lo cual facilita tanto las búsquedas como la obtención de resultados. A pesar de que los métodos mencionados son manuales, para uso de traductores o terminólogos en el transcurso de su trabajo, sería posible automatizarlos hasta cierto punto, por ejemplo, creando barras de búsqueda para Internet Explorer que permitan realizar búsquedas simultáneas en 3 tipos de páginas Web (monolingües en inglés o en español, PubMed, páginas bilingües), con la opción de restringir las búsquedas a páginas Web procedentes de países hispanohablantes o de ciertas instituciones de reconocida fiabilidad por la experiencia del profesional.

Una característica importante de los recursos citados es la posibilidad de emplearlos como estrategia interdisciplinaria y multisectorial fuera del marco de la traducción, aplicable a la mejora de la indización y la recuperación de documentos, al suministro de un marco apropiado para la evaluación terminológica, a la normalización lingüística y estilística en medios multilingües, a campos como la bibliotecología, la elaboración de tesauros, la prensa hablada y escrita, la revisión editorial, la investigación científica y la docencia, así como a la creación de sitios Web idóneos y fidedignos, por citar sólo algunos. Estas aplicaciones prácticas permitirán atender la demanda de información multilingüe y en español, en particular.

Para finalizar, vale recordar que generalmente las traducciones van acompañadas de alguna limitación. La falta de especificación del contexto puede dar lugar a una interpretación equivocada (polisemia). Pueden existir varias traducciones estándar para una expresión dada (sinonimia), o ninguna (vacío terminológico). Sin embargo, esto se aplica a todos los idiomas y para todos los campos del conocimiento, y no menoscaba la validez de los métodos presentados.

AGRADECIMIENTOS

Deseo expresar mi agradecimiento a Martha Daza, Julia Aymerich y Olivier Bodenreider por la inestimable retroalimentación brindada. Este proyecto ha sido patrocinado por el Programa de Investigaciones

Intramurales de los Institutos Nacionales de la Salud de los Estados Unidos, Biblioteca Nacional de Medicina, Centro Nacional Lister Hill para las Comunicaciones Biomédicas (LHNCBC). Quedo también muy agradecida al Centro Universitario de Investigaciones Bibliotecológicas de la UNAM por la gentil invitación a esta conferencia.

REFERENCIAS BIBLIOGRÁFICAS

- [1] Carli A, Calaresu E., “Le lingue della comunicazione scientifica. La produzione e la diffusione del sapere specialistico in Italia”, en *Ecologia linguistica. XXXVI Congresso Internazionale di Studi della Società di Linguistica Italiana*, Roma, 27-74, 2003.
- [2] Curry MJ, Lillis T., “The global imperative to publish in English: Brokering multilingual scholars into English academic publishing”, en *NYS TESOL Idiom*, 34(2), 8-9, 2004.
- [3] Arunachalam S, Singh UN., “Where do Israeli scientists publish?” en *Journal of Information Science*, 14(4), 221-231, 1988.
- [4] Roseblat G, Tse T, Gemoets D, Gillen JE, Ide NC., Supporting access to consumer health information across languages, ICML 9/CRICS 7, Brazil, 2005.
- [5] *UMLS Knowledge Sources*, US National Library of Medicine, Documentation, 2007 AA, <http://www.nlm.nih.gov/research/umls/documentation.html>
- [6] Nelson SJ, Schopen M, Schulman JL, Arluk N., An Interlingual Database of MeSH Translations.
- [7] Fugmann R., Subject analysis and indexing: theoretical foundation and practical Advice, Frankfurt/Main, INDEKS Verlag, 1993, en <http://www.nlm.nih.gov/mesh/intlmesh.html>

- [8] Vivaldi J., Extracción de candidatos a término mediante combinación de estrategias heterogéneas, Barcelona, Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra, 2004, Serie Tesis 9. 2001.
- [9] Feliu, J., Relacions conceptuals i terminologia: anàlisi i proposta de detecció semiautomàtica, Barcelona, Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra, 2004.
- [10] Gómez González-Jover A, Vargas Sierra C., Córpora comparables y paralelos para la detección de terminología bilingüe: su explotación y uso con herramientas informáticas, VIII Simposio RiTerm, Cartagena de Indias, 2002.
- [11] Barzilay R, Elhadad N. Sentence, Alignment for Monolingual Comparable Corpora, EMNLP, Japan, 2003.
- [12] <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?DB=pubmed>
- [13] Zweigenbaum P, Grabar N., “Automatic acquisition of morphological knowledge for medical language processing”, en *Lecture Notes in Computer Sciences*, Vol. 1620, 416-420, 1999.
- [14] Baud RH, Nystrom M, Borin L, Evans R, Schulz S, Zweigenbaum P., “Sentence. Interchanging lexical information for a multilingual dictionary”, en *American Medical Informatics Association*, Simposio Anual 2005, 31-5
- [15] Roseblat G, Tse T., “Study of a Spanish Language ClinicalTrials.gov Prototype System”, en *American Medical Informatics Association*, Simposio Anual 2006, 659-664.