

BTMEX, el banco de terminología de México

María Pozzi
El Colegio de México

Conscientes de la necesidad que existe en varios sectores del país de establecer una comunicación técnica y científica multilingüe eficiente y conociendo al mismo tiempo los problemas de orden lingüístico que enfrentan los profesionales de las diferentes industrias, debido a la falta de disponibilidad de terminología de alta calidad, El Colegio de México está desarrollando un programa de investigación, "BTMEX, el banco de terminología de México", por medio del cual se intenta resolver por lo menos una parte de estos problemas. El proyecto BTMEX consiste en registrar los conceptos de un campo del conocimiento a través de los diferentes términos que se usan para denotarlo, señalando el más usual como el término principal y los demás sinónimos como formas secundarias. Asimismo se indican también las características lingüísticas de cada término y se proporcionan una definición y un contexto en donde se muestran los términos en su ambiente lingüístico. Finalmente se indican las referencias bibliográficas que se utilizaron durante la investigación del concepto.

As a response to the increasing need to establish technical and scientific communication effectively at national and international levels, the project "BTMEX, the term bank of Mexico" was set up at El Colegio de México. BTMEX's main objective is to register each concept of a specific subject field and the different terms used to denote it, indicating the recommended term, acceptable synonyms and terms that should be avoided. It also includes grammatical information on each term as well as a definition of the concept and a context to show each term in its own linguistic environment. Bibliographic references used during research are recorded as well. Since BTMEX was designed as a working tool, a flexible computational system was implemented in which a wide range of search strategies are available as to allow the end-user to find it if not all the information he is looking for, at least a useful part of it. These strategies include searching for full terms; left, right or left and right truncated terms; any string of contiguous characters, individual words on an arbitrary subset of the full term. It is also possible to consult an alphabetic index of terms contained in the bank from which one or several terms may be selected to retrieve the required information. BTMEX includes information in Spanish, English and French, and during the first stage it will only contain information on the environment.

1. Introducción

La terminología es una actividad tan antigua como el conocimiento mismo. De hecho, en la antigüedad, la terminología consistía en asociar sistemáticamente nombres a las cosas y regular el uso de los lenguajes técnicos y las reglas de nomenclatura. Sin embargo, no es sino hasta fines del siglo pasado, cuando empezó el desarrollo vertiginoso de la ciencia y la tecnología que el estudio de la terminología cobró importancia, estableciéndose hace relativamente poco tiempo como una disciplina independiente que está relacionada a su vez con muchas otras disciplinas y cuyas bases teóricas son tanto lingüísticas como cognoscitivas.

En las últimas décadas, la comunicación instantánea ha llegado a ser una parte indispensable de nuestra forma de vivir. La transferencia de conocimientos y de tecnología se lleva a cabo día a día, permitiendo así una verdadera globalización de los mercados. Ahora bien, para que esta transferencia se logre de una manera eficiente, es necesario transmitir tanto los nuevos conocimientos como la tecnología a través de una terminología especializada en varias lenguas.

Es un hecho que en términos generales, en México, al igual que en los demás países de habla hispana, no se producen desarrollos substanciales en el campo de la ciencia y la tecnología como sucede en los países de habla inglesa, alemana y francesa; por esta razón tampoco se produce terminología original en español, lo que quiere decir que al importar nuevas tecnologías hay que traducir la terminología correspondiente.

Esto presenta varios problemas de índole práctico, ya que en ocasiones el concepto que se importa es nuevo y por lo tanto no existe un término equivalente que lo denota en la lengua de llegada. Es necesario crear entonces un neologismo que tarda varios meses si no es que años en aparecer en los diccionarios técnicos especializados. Mientras tanto, si ese mismo concepto tuvo que ser traducido por alguien más, es muy probable que haya encontrado otro término para denotarlo. Y así sucesivamente, hasta que en un momento dado existen n términos equivalentes que denotan un concepto. Esta proliferación innecesaria de términos dificulta la comunicación.

Por otro lado, cuando ya existen varios términos que denotan un sólo concepto en la lengua de llegada, ¿cómo puede el traductor asegurarse de que el término escogido es el que mejor representa el concepto que está traduciendo? Esto nos lleva a pensar que al igual que la sinonimia es la fuente principal de riqueza y creatividad individual en la lengua natural, en los lenguajes especializados, esta situación no sólo es inconveniente sino que es altamente indeseable, ya que, en la mayoría de los casos, puede crear una seria confusión al lector.

Consciente de la necesidad que existe en varios sectores del país de establecer una comunicación técnica y científica bi- o multilingüe eficiente, y conociendo al mismo tiempo los problemas de orden lingüístico que enfrentan los profesionales de las diferentes industrias, debido a la falta de acceso a terminología de alta calidad específica de cada área, El Colegio de México, con el apoyo del CONACYT está de-

sarrollando un programa de investigación “BTMEX, el banco de terminología de México”, por medio del cual se intenta ayudar a solucionar una parte importante de estos problemas.

2. BTMEX, el banco de terminología de México

BTMEX se estableció en 1994 con el objeto de proporcionar información terminológica de calidad controlada al mayor número de traductores, intérpretes, escritores técnicos, científicos, investigadores, técnicos, etc. quienes en virtud de sus actividades profesionales necesitan tener un acceso fácil y eficiente a la terminología de su especialidad en español, inglés y/o francés. En su primera etapa, incluirá únicamente cuatro áreas: medio ambiente, economía, finanzas y comercio internacional. Otras áreas se incluirán gradualmente.

Ya que BTMEX ha sido diseñado como una herramienta de trabajo, se consideró que era necesario implementar un sistema que incluyera la posibilidad de realizar búsquedas correspondientes a un concepto, que es la unidad básica de la investigación terminológica, sin importar si existen uno o más sinónimos que lo denotan. Por ejemplo, el concepto *exchange rate mechanism* tiene cuatro equivalentes aceptados en español, que son: *mecanismo de tasa cambiaria*, *mecanismo europeo de tasa cambiaria*, *mecanismo de tipo cambiado* y *mecanismo de tipo de cambio*. Si el usuario hace la búsqueda sobre cualquiera de estos cuatro términos, el resultado aparece desplegado en la pantalla, en donde se indica que el término *mecanismo de tasa cambiaria* es el preferido, y por lo tanto se recomienda su uso, y los demás corresponden a las formas secundarias.

2.1 Contenido

La información contenida en BTMEX está organizada en registros, cada uno de los cuales corresponde a un concepto determinado; ésta puede agruparse en cuatro categorías:

Datos de control, que se utilizan para identificar el registro en sí, a cada una de las personas que intervinieron en la investigación del concepto; las fechas de entrada y actualización de la información, etc. Existe también un campo en donde se indica el status del registro y la confiabilidad de la información contenida.

Datos lingüísticos, que incluyen el término principal, sinónimos, abreviaturas, variantes ortográficas, información gramatical correspondiente a cada término, el código de lenguaje y/o región, y uno o más contextos en donde se muestran estos términos en su ambiente lingüístico.

Datos conceptuales, que proporcionan la definición del concepto en cuestión, el(los) códigos de área a que pertenece el concepto y un conjunto de términos con los que el concepto está relacionado, indicando también el tipo de relación que los une.

Datos bibliográficos, que indican las referencias bibliográficas que se utilizaron durante la investigación terminológica. Estos datos son de suma importancia, ya que la calidad de la información depende en gran medida de las fuentes utilizadas.

40t'Gut cvgi kc uf g'dAus wgf c'qht gelf cu't qt 'DVO GZ

El objetivo principal de BTMEX consiste en proporcionar a sus usuarios una herramienta de trabajo que realmente funciona, de manera que hay que darles toda la oportunidad posible de encontrar la información que están buscando, o en caso de no encontrarla, obtener por lo menos alguna información útil.

Para asegurar la mejor manera de recuperar la información deseada, BTMEX ofrece al usuario varias estrategias de búsqueda, que a continuación se describen:

2.2.7 Búsqueda de términos completos

Un término completo es cualquier elemento recuperable definido en BTMEX y que corresponde ya sea a la entrada principal, o bien, a cualquiera de sus formas secundarias expresada en forma de sinónimos, abreviaturas o variantes ortográficas en la lengua de partida seleccionada. Estos términos pueden estar compuestos de una, dos o más palabras. Cuando un usuario desea obtener información por este método, debe preguntar por el término en la misma manera en que éste fue almacenado en la base de datos, ya que la menor discrepancia en, por ejemplo, género y/o número resulta en que no se encuentre una respuesta satisfactoria. Sin embargo, los términos cuyas palabras están unidas por un guión, o términos que contienen acentos u otros diacríticos pueden ser recuperados independientemente de la forma en que se busquen, ya que el sistema antes de ejecutar la búsqueda convierte la consulta hecha por el usuario a mayúsculas, removiendo al mismo tiempo todos los guiones, acentos y diacríticos. Por ejemplo, la consulta *relación capital- producto* producirá el mismo resultado que *relación capital-producto* ya que la conversión de ambas resultará en la búsqueda del término *RELACION CAPITAL PRODUCTO*.

2.2.2 Búsqueda de términos truncados por la derecha

Cuando el término completo no se encontró, se pueden especificar términos truncados por la derecha. Esta técnica permite al usuario buscar cadenas de caracteres al principio del término. De esta manera pueden encontrarse términos relacionados con el término original que pueden proporcionar información relevante al caso. Por ejemplo, si el término *income account* no fue encontrado en el banco, el usuario pueden encontrar algunos términos relacionados al buscar el término *income\$*, que producirá la siguiente lista de términos contenidos en el banco:

- *income*
- *income bracket*
- *income effect*
- *income tax*
- *in comes policy*

2.2.3 Búsqueda de términos truncados por la izquierda

Esta técnica permite al usuario buscar cadenas de caracteres al final del término, siendo particularmente útil cuando habiendo buscado un término completo en inglés, éste no se encontró. Por ejemplo, la búsqueda de \$tax producirá los siguientes términos contenidos en el banco:

- *cid valorem tax*
- *excise tax*
- *income tax*
- *sales tax*
- *shifting of tax*
- *valué added tax, etc.*

2.2.4 Búsqueda de términos que contienen una cadena predefinida de caracteres

Esta técnica permite al usuario encontrar términos que contienen una cadena predefinida de caracteres que puede localizarse en cualquier lugar del término. Por ejemplo, al consultar \$income\$, los términos que contienen esa cadena son los siguientes:

- *factor income payments*
- *income*
- *income bracket*
- *income effect*
- *income tax*
- *bicornes policy*
- *national income*
- *proportional income tax*
- *real income*

2.2.5 Búsqueda parcial de términos de dos o más palabras

Cuando un término compuesto de dos o más palabras no se encuentra en BTMEX, el sistema ofrece una técnica que permite al usuario buscar toda clase de combinaciones de las palabras relevantes del término. Esta técnica fue implementada con el uso de operadores de proximidad, que permite la recuperación de términos que contienen palabras específicas en la posición definida por el usuario. Por ejemplo, si la consulta orí-

ginal *proportional tax* no es encontrada, el usuario puede tratar de buscar un término que contenga las palabras *proportional* y *tax*, encontrando el término *proportional income tax*, que probablemente le será de utilidad.

2.2.6 Búsqueda de texto libre

Para el usuario más avanzado, el sistema ofrece la posibilidad de encontrar una cadena de caracteres cualquiera en los campos que no fueron indexados, como por ejemplo en las definiciones, contextos, etc.

2.2.7 Conversión ortográfica rudimentaria

En inglés, principalmente, existen variantes ortográficas que corresponden a diferentes regiones geográficas. Esto se refleja en la existencia de términos sinónimos cuya única diferencia se encuentra en su forma ortográfica. Idealmente estas variantes deberían quedar registradas; sin embargo, desde el punto de vista práctico, ya que en México se recibe la mayor parte de la información terminológica de Estados Unidos y no en inglés británico, es usual registrar tan sólo los términos en inglés americano. Pero, si un usuario desea encontrar un término específico del inglés británico, debe poder obtener esta información con un mínimo de trabajo. Entonces, para los casos en que existe una regla general de conversión entre el inglés británico y el americano, se implemento un procedimiento que genera una nueva consulta que contiene ambas variantes ortográficas. Un ejemplo de esta situación son los términos que terminan en *ise/ize*, *our/or*, *tre/ter*, etc. Por ejemplo, si la consulta es *commercialise*, el sistema la convertirá en *commercialise* + *commercialize*, por lo que la búsqueda se llevará a cabo sobre las dos formas.

2.2.8 Despliegue de términos contenidos en BTMEX

Algunas veces, cuando un término no se encuentra en la base de datos, el usuario puede decidir desplegar en la pantalla los términos que están “cerca” del que no se encontró. En este caso el sistema ofrece la opción de desplegar alfabéticamente hasta 35 términos, algunos de los cuales pueden ser útiles, ya que por lo menos pueden estar relacionados en su forma lingüística.

2.2.9 Despliegue de términos que contienen una cadena predefinida de caracteres

En el caso de que el usuario no encuentre un término en el banco, se le ofrece también la posibilidad de buscar en el índice de términos aquellos que contienen una cadena de caracteres definida por él mismo. Esta cadena puede ser una palabra clave o la parte relevante de un término, pero que no necesariamente corresponde a una palabra entera.

De la descripción de las estrategias de búsqueda ofrecidas por el sistema BTMEX puede deducirse que si existe alguna información que pueda ser útil al usuario, existe por lo menos una manera de recuperarla, aunque es el usuario quien tiene la última palabra sobre el método que debe utilizar para obtener la información que requiere en un determinado momento.

2.3 Interfaz del usuario

La interfaz del usuario fue diseñada para proporcionar un método integrado, eficiente y fácil de usar y aprender para interactuar con el BTMEX. Esta interacción puede ser en forma de los procedimientos de entrada de la información, incluyendo la creación, modificación y eliminación de registros; en forma de recuperación de la información, incluyendo la elección de estrategia de búsqueda, o en forma de producción de glosarios e informes impresos, a través de varios formatos de salida. Finalmente, esta interacción puede ser en forma del control y mantenimiento de la base de datos misma.

El componente básico de esta interfaz es una estructura jerárquica de menús que permiten al usuario escoger entre los servicios disponibles. Desde el punto de vista del usuario, los servicios de recuperación de información constituyen la parte fundamental del BTMEX, ya que éstos le permiten ejecutar una búsqueda, desplegar alfabéticamente los términos incluidos y desplegar y/o salvar para su posterior impresión los resultados de una búsqueda. También es posible re-ejecutar búsquedas previamente ejecutadas, seleccionar el formato de salida apropiado para una o más consultas y desplegar secciones del archivo maestro de acuerdo ya sea secuencialmente por el número de registro, o los registros que contienen una cadena de caracteres definida por el propio usuario. Finalmente, el usuario puede seleccionar las lenguas de partida y llegada y, opcionalmente, el área del conocimiento que consultará en la presente sesión.

BTMEX, por otro lado, ofrece facilidades de diálogo multilingüe en español, inglés y francés. Esto significa que el usuario puede seleccionar el lenguaje en que desea comunicarse con el sistema, y así, todos los menús, submenús y mensajes se despliegan en dicha lengua.

2.4 Formatos de salida

Ya que BTMEX fue diseñado para satisfacer las necesidades de sus usuarios potenciales, se tomó en consideración que no todos ellos requieren el mismo patrón de presentación de la salida final. Por este motivo se ofrece la posibilidad de seleccionar en cualquier momento el formato de salida en que desean obtener la información recuperada.

Con el propósito de generar la salida final, la información contenida en cada registro se dividió en cuatro bloques de acuerdo con su contenido. Estos bloques son los siguientes:

bloque de control (BC), que contiene los datos de control,

bloque lingüístico (BL), que se refiere a toda la información sobre los términos que denotan el concepto en cuestión,

bloque de apoyo textual (BT), que contiene información que facilita la comprensión del concepto y el(los) término(s) que lo denota(n), bloque de referencia (BR), que contiene las fuentes de referencia utilizadas durante la investigación terminológica.

Los formatos de salida definidos son los siguientes:

- BL
- BL + BT
- BL + BR
- BL + BC
- BL + BT + BR
- BL + BR + BC
- BL + BT + BC
- BL + BT + BR + BC

La información se presenta al usuario en la lengua de partida seguida por la información en la lengua de llegada.