

Lexicón trilingüe en línea
sobre *Information Technology*
para el Mercosur

**Grupo de investigación “Lexicón trilingüe”,
Universidad Nacional de La Matanza (UNLM),
Buenos Aires, Argentina.**

Oscar Domínguez Soler,
Julio Bossero,
Silvia Delayel,
Marcelo Estayno,
Alicia Inés García,
Graciela Malevini,
Inés Noguera,
Fernando Orthusteguy,
Ernesto Adolfo Solari.

Resumen

El desarrollo en el campo de la informática y la tecnología asociada conlleva el crecimiento de la terminología con la que se denominan los objetos creados por los mismos –publicados generalmente en inglés– y hacen necesaria su traducción a otros idiomas con una celeridad que excede las posibilidades de publicación de glosarios o diccionarios en papel. A su vez, la informática influye en el campo de la investigación terminológica, lo que se traduce en la posibilidad de trabajar con corpus extensos y la utilización de bancos de datos textuales. En el marco de la cooperación entre naciones, es fundamental contar con recursos adecuados a la realidad transnacional que ahorren tiempo en el campo de la interpretación para evitar el empobrecimiento de la lengua a través de la incorporación de terminología técnica directamente en inglés por la necesidad de ahorrar tiempo o por la extrañeza que causa utilizar terminología propia del español peninsular. Se evidencia así la necesidad de contar con un sistema que en el campo de la informática y su tecnología asocia-da permita viabilizar tridireccionalmente respuestas en línea a dudas terminológicas en inglés, español rioplatense y portugués, dada la importancia creciente de la cooperación en el marco del Mercosur.

Objetivo

El propósito de este trabajo es informar acerca de la tarea de investigación de docentes de la Universidad Nacional de La Matanza sobre la confección de un lexicón trilingüe (inglés, español rioplatense, portugués) en el área de Information Technology.

Una de las particularidades del trabajo es que se produce una interacción entre campos diversos, como son la Filosofía, la Lingüística, la Ingeniería en Sistemas, la Semántica y la Terminología, lo cual requiere una respuesta transdisciplinaria para acercarse al campo de estudio así como también la teorización que permita asirlo en forma práctica para producir un objeto –en este caso un lexicón– que sirva a los fines propuestos. Por lexicón entendemos una lista de palabras; pero en un sentido específico en el campo de la lingüística se refiere al cuerpo de unidades estructurales de las que se compone el lenguaje.

La elección misma del término lexicón para describir el objeto a producir es tema de discusión teórica, puesto que los investigadores entendemos que los diccionarios, glosarios o vocabularios existentes no dan respuesta a lo que se pretende encarar, que es satisfacer las necesidades de un usuario interesado en la comprensión de términos especializados en un campo dado.

Orígenes del proyecto

Como docentes universitarios, especializados en las áreas de Informática y de Idiomas de la Universidad Nacional de La Matanza, y como usuarios de medios, nos vemos afectados por la celeridad con que se producen a nuestro alrededor y por la necesidad que tenemos, compartida por nuestros alumnos, de contar con medios para disminuir el grado de incertidumbre que produce el crecimiento desmedido de términos con los que se describen objetos esenciales para comprender su interacción con otros, en redes sistémicas de complejidad creciente. A ello deben sumarse las condiciones de competitividad e inseguridad para intentar comprender un mundo globalizado, por lo que vimos la necesidad de elaborar un lexicón que satisficiera las demandas de lectores de español rioplatense, y luso-parlantes de textos en inglés del campo de la informática y la tecnología asociada en el ámbito del Mercosur.

Dato versus información

Es necesario recordar esta distinción entre dato e información: en la tecnología de la información se da un nuevo énfasis a la presencia de las telecomunicaciones y a la habilidad para interpretar 'información', mientras que se supuso que el procesamiento de datos manejaba datos 'en bruto'. Tan pronto como hablamos de la interpretación y transmisión de información en lugar de hablar de datos, estamos fundamentalmente interesados en hechos lingüísticos y en una industria del lenguaje multilingüe, como ya están empezando a reconocer los países de habla inglesa, defensores acérrimos del monolingüismo.

Vemos, entonces, que el crecimiento tecnológico lleva a nuevas actividades económicas, muchas de las cuales dependen de comunicaciones eficientes y de una información confiable. Esto no es solamente una cuestión de generación de riqueza o prestación de servicios, sino una actividad de pensamiento, de intento de liberar la información, la sustancia de la comunicación, de sus soportes poderosos que crean entornos discursivos novedosos y maneras de interpretación distintas: no es infrecuente escuchar 'clickear', u otro neologismo por el estilo, porque el término extranjero se ha incorporado con la misma velocidad con la que hizo su aparición el instrumento que incorpora esta manera de interacción. Si se tiene en cuenta que en un período relativamente breve los medios masivos de comunicación se han sumado a la *www*, vemos que estamos frente a una amplia distribución, donde un vasto público lego se enfrentará a información 'especializada', aunque no destinada a especialistas.

Forma de trabajo

Cabré (1993), sostiene que, aparentemente, los términos no parecen diferir mucho de las palabras, considerados desde una perspectiva formal o semántica; pero se diferencian notablemente de ellas si aportamos criterios pragmáticos y comunicativos.

En efecto, la peculiaridad más notable de la terminología, en contraste con el léxico común, consiste en que sirve para designar los conceptos propios de las disciplinas y actividades de especialidad. En consecuencia, los términos son conocidos fundamentalmente por los especialistas de cada una de esas materias y aparecen con una frecuencia muy elevada en los documentos especializados de cada disciplina.

Los términos, como las palabras del léxico general, son unidades signíficativas distintas y significativas al mismo tiempo, que se presentan de forma natural en el discurso especializado. Poseen pues una vertiente sistemática (formal, semántica y funcional), toda vez que son unidades de un código establecido, y manifiestan asimismo otra vertiente pragmática, puesto que son unidades usadas en la comunicación especializada para designar los 'objetos' de una realidad preexistente.

Según Lérat (1990), los términos son unidades de forma y contenido que pertenecen al sistema de una lengua determinada, en cuyo interior conviven distintos subsistemas específicos alternativos. Estos subsistemas no son, en su especificidad, subcódigos completos, sino que representan normalmente variaciones parciales respecto del código común.

Como cualquier otra unidad significativa de un sistema lingüístico, los términos forman parte de un sistema estructurado en el que ocupan un determinado nivel (el nivel de las unidades léxicas) y se relacionan, por un lado, con las demás unidades del mismo nivel y, por otro, con las unidades de los demás niveles, participando conjuntamente en la construcción del discurso.

Un paso imprescindible en la tarea terminográfica que se encare es "la apropiación, al menos en sus rasgos principales, del campo conceptual sobre el que se va a actuar" (Alpízar Castillo, 1997). Apoyándose en TERMCAT (1990), propone el árbol conceptual o árbol de campo, "una representación esquemática de la realidad del ámbito que se investiga, donde se recogen las principales esferas (ramas del árbol) que conforman la tarea en cuestión".

Así, se propone, tentativamente, adoptar un árbol compuesto por cuatro áreas:

- General
- Campos del área
- Áreas de aplicación
- Área de las ciencias relacionadas

General

Se trata del área terminológica que representa el campo de estudio en sí: Information Technology.

Campos del área

Se trata de las subdivisiones en las que puede descomponerse el campo de estudio general, y que reciben atención especializada.

Area fields

- knowledge-based systems
- database systems
- intelligent agents
- text processing
- natural language processing
- intelligent interfaces
- hypertext and information retrieval systems
- machine learning

- artificial life

Campos que componen el área

- sistemas basados en el conocimiento
- sistemas de bases de datos
- agentes inteligentes
- procesamiento de texto
- procesamiento de lenguaje natural
- interfaces inteligentes
- sistemas de hipertexto y de recuperación de información
- máquinas que pueden aprender (sistemas de aprendizaje para máquinas)
- vida artificial

Áreas de aplicación

Se trata de las áreas relacionadas con el campo, pero que no pertenecen directamente a él sino a través del desarrollo de aplicaciones.

Application areas

- information retrieval
- manufacturing planning and scheduling
- electronic commerce
- healthcare information systems
- electronic publishing

Áreas de aplicación

- recuperación de información
- ordenamiento y planificación de la producción
- comercio electrónico
- sistemas de información de la salud
- publicación electrónica

Área de las ciencias relacionadas

Se trata de las relacionadas pertenecientes a otras ciencias, lo que trae aparejado cambios de significado para ciertos términos.

Related sciences

- Computer Science
- Electrical Engineering
- Information Systems
- Law

Ciencias relacionadas

- Computación
- Ingeniería eléctrica
- Sistemas de información
- Legislación

De acuerdo con estos principios, el equipo transdisciplinario integrado por docentes del área de la Informática y de la Lingüística (inglés, español y portugués) se abocó a la tarea de diseñar el árbol conceptual sobre el que se iba a trabajar.

Luego, tomando como fuente textos de la especialidad que se encuentran disponibles en Internet, procedió a seleccionar términos, definirlos, traducirlos y almacenarlos en una base de datos creada para tal efecto para conformar un lexicón. También se contempló la inclusión en ella de figuras y diagramas.

A su vez, le fue encomendada a un grupo de alumnos avanzados de la carrera de Ingeniería en Sistemas de Información la tarea de crear una página Web, dentro de la página de la Universidad, donde se irán publicando los elementos que se encuentran en la base de datos mencionada. Dicha página se encuentra en una etapa de prueba y es la intención del equipo de investigación ir ampliándola no sólo con los aportes propios sino con los de los visitantes de la página.

Implementación

La naturaleza misma del campo de investigación, la terminología en el terreno de la informática hace que este proyecto requiera de ella no sólo como su base de alimentación lingüística, sino que se funda en ella para desarrollar su implementación. El proyecto ha incorporado los elementos destinados a ello:

- WIN NT 4.0 Service Pack 6
- SQL Server 7.0 Service Pack 2
- WIN NT 4 OPTION PACK
- Extensiones de FrontPage
- Internet Information Server
- Servicios de Web, FTP y SMTP
- Management Console
- Transaction Server
- Interdev RAD
- Script Debugging

Conclusión

El proyecto encarado pretende mostrar las bondades y beneficios de la aplicación de procedimientos tomados de campos diversos y aplicarlos al estudio de la terminología. La implementación de este modelo de lexicón intenta cubrir un espacio distinto para dar respuesta a una necesidad creciente en un campo en evolución y que potencialmente puede representar el riesgo de la pérdida de identidad lingüística por la incorporación de terminología. Además, la modalidad en línea en la que se comienza a trabajar para el desarrollo del modelo apunta a brindar una respuesta en tiempo real a la búsqueda y consulta que se plantee en el campo.

Bibliografía

- ALPÍZAR CASTILLO, RODOLFO, *¿Cómo hacer un diccionario científico técnico?*, Buenos Aires, Memphis, 1997.
- ARNTZ, R. Y PIGHT, H., *Introducción a la terminología*, Madrid, Pirámide, 1995.
- CABRÉ, MA. TERESA, *La terminología. Teoría, metodología, aplicaciones*, Barcelona, Antártida/Empúries, 1993.
- LÉRAT, P., "Les fondements théoriques de la terminologie", en *La Banque des Mots*, número spécial, 1990.
- SAGER, JUAN C., *Curso práctico sobre el procesamiento de la terminología*, Madrid, Pirámide, 1993.
- SAGER, JUAN C., *Language Engineering and Translation: Consequences of automation*, Amsterdam/Philadelphia, John Benjamins Publishing Company, 1994.
- Scientific and Technical Translation*, American Translators Association Scholarly Monograph Series, Volume VI, Amsterdam/Philadelphia, John Benjamins Publishing Company, 1993.
- ZAMPOLLI, ANTONIO, "Los bancos de datos léxicos: bases multifuncionales de datos léxicos", en *Las industrias de la lengua*, José Vidal Beneyto (ed.), Madrid, Pirámide, 1991.