

Estructura semántica del léxico en un diccionario onomasiológico práctico

Gerardo Sierra

Instituto de Ingeniería, UNAM

El diccionario onomasiológico es una herramienta práctica que asiste al usuario en la búsqueda del término relacionado o similar para cierto concepto que no sabe denominar de forma correcta. Gracias a la lingüística computacional es posible crear esta herramienta en un programa de cómputo que le permita a un usuario introducir todas las "ideas" que tenga sobre el término a buscar, con las palabras que quiera y en el orden que sea. Un aspecto fundamental es, por tanto, la previsión de las palabras que una persona usaría en esta búsqueda. Este artículo se enfoca al análisis del léxico que resulta significativo en la búsqueda onomasiológica, a la vez que se aplica a una terminología selecta.

Onomasiological dictionary is a practical tool that can help the user to search for related or similar terms of certain concepts he does not know how to denominate in a correct way. Computational Linguistics allows to develop this tool through a software. It eases the search for the desired term from an "idea" using any word in any order. An important aspect is foresee all the possible words to be used by user. The lexical analysis of pertinent words for onomasiological search is presented. Application for one selected terminology is given.

I. Introducción

Desde un punto de vista semántico, para llegar al concepto de significado, se pueden distinguir dos tipos de diccionarios (Baldinger 1970:141): el semasiológico y el onomasiológico. En el primero de ellos se pregunta por las “significaciones”, esto es, se inicia del nombre (significante) para buscar el sentido o los sentidos ligados a él. Por el contrario, en el segundo se pregunta por las “designaciones”, es decir, se parte del sentido y se busca el nombre o nombres conectados a éste.

Este segundo tipo de diccionario resulta de mucha utilidad cuando un autor creativo desea expresar una idea pero desconoce el término adecuado para ello. Sin embargo, en la praxis de la búsqueda onomasiológica, no se tiene a la fecha ningún diccionario *ad hoc*.

Un problema tradicional en los diccionarios de carácter onomasiológico es que la estructuración del léxico está definida con los criterios del lexicógrafo, los cuales en la mayoría de los casos no coinciden con los del usuario. Esto es perfectamente razonable si se entiende el postulado de la arbitrariedad del signo, de tal forma que los nombres no pueden designar biunívocamente las cosas, esto es, no pueden conducir directamente a lo que designan, debido a que en el camino se interpone el *sentido* que le da cada autor (Lara 1989:243).

En este sentido, ha faltado un diccionario pensado para el usuario, de tal forma que éste pueda realizar la búsqueda de términos con sus propias palabras, no obstante éstas no sean la más “adecuadas”.

Con este fin, en el Instituto de Ingeniería se está desarrollando un programa de cómputo, denominado DEBO (Diccionario Electrónico de Búsqueda Onomasiológica), que asiste al usuario en la búsqueda de términos de una área específica a partir del concepto, a través de las ideas con que cuenta, con las palabras que considere convenientes y en cualquier orden.

II. Estructura del DEBO

De manera general, el DEBO es una interfaz que le permite a un usuario introducir un concepto determinado, a través de palabras, y que le proporciona como resultado una lista de los términos usados para designarlo.

DEBO relaciona dos tipos de datos: los de entrada se denominan *palabras clave* y están integradas en una base de datos, llamada *PALCLV* (PALabra CLaVe); los de salida son los *términos*, integrados en una base de datos llamada *TERFIN* (TERminos FI-Nales).

Por cada entrada se puede obtener varios datos de salida, al igual que cada posible salida se puede obtener por uno o varios datos de entrada.

Para relacionar ambos tipos de datos, y a la vez hacer más eficiente el funcionamiento del DEBO, resulta conveniente y necesario formar asociaciones entre las palabras clave que presentan rasgos semánticos comunes y que pueden ser utilizadas con el mismo sentido en el contexto del término final al que corresponden. Dicha asociación de palabras se le ha denominado *paradigma*, aprovechando la concepción original de Saussure sobre *relaciones asociativas* (1982:175) y que posteriormente ha sido desarrollada por otros autores (Coseriu 1986).

Para diferenciar los distintos paradigmas, a cada uno se le da el nombre de cualquiera de las palabras clave que lo constituyen. A esta palabra se le distingue como *líder* del paradigma.

Así, en lugar de hacer corresponder un gran número de palabras clave con los términos finales, se utilizan los paradigmas. En total, se tienen dos tipos de relaciones: la primera para hacer corresponder a cada palabra clave uno o más líderes; la segunda para hacer corresponder cada líder con un conjunto de términos finales.

La forma en que DEBO opera se ha descrito en otros artículos, dentro del contexto de lingüística computacional (Sierra 1994, 1995). Este artículo, más de carácter semántico, expone la configuración de las distintas bases de datos y relaciones.

III. Estructura lexical del DEBO

Para determinar la configuración léxica que debe contener un diccionario de esta naturaleza, son dos tareas principales las que tienen que llevarse a cabo, que en los términos de Alpízar (1994:82) son: *tecnolexicología* y *teconoceptología*.

La primera consiste en la delimitación del *corpus* de los términos finales, que se construye sobre la base de las terminologías establecidas en cada campo del saber (Quiles y Zítara 1994:97).

La segunda tarea consiste en la identificación de las posibles palabras clave que resultan significativas para la búsqueda onomasiológica de cada uno de los términos finales. Esto es, palabras que el usuario escribe y que resultan determinantes para encontrar el término que se busca.

En cuanto a esta segunda tarea, existen diferentes herramientas tomadas de la lógica, la semántica, la documentación y la información para analizar las relaciones entre los conceptos y los términos que los denominan. En el proyecto del DEBO, entre diversos métodos alternativos y complementarios para definir las palabras clave asociadas a un término, se utilizaron:

a) *Marco conceptual*

Por la misma relación entre los términos, se consideró primordial definir estas relaciones y de agrupar los términos de acuerdo con la estructura conceptual del campo de conocimientos que está siendo analizado (Barbosa y Fernandes 1994:154), lo

que permite iniciar la identificación de las palabras clave por el lugar que se ocupa dentro de la asociación.

b) *Definiciones*

También se pueden encontrar las palabras clave utilizadas en un término dado a partir de las definiciones, que son descripciones de conceptos por medio de conceptos ya conocidos (Pozzi 1994:91).

c) *Expertos*

Una vez que ya se ha llegado a un buen número de selección de palabras clave, conviene realizar una consulta a expertos (Lara 1991), en la que se solicita que mencionen las palabras clave más relevantes que ellos consideran para representar un término dado. Se puede pedir que estas palabras se den aisladas o en conjunto.

d) *Contexto*

Finalmente, se puede consultar en el contexto en que se encuentre un término y de ahí seleccionar las palabras que se considerarían clave, con el riesgo de ser un método más subjetivo.

En cualquier caso, una vez identificadas las palabras clave en una primera aproximación, hay que ir sustituyendo cada una de ellas por otras palabras que pueden estar asociadas sin cambiar el sentido de la definición. De igual manera, hay que encontrar todas las variaciones morfológicas y derivaciones que puede utilizar el usuario en el proceso de búsqueda (Cardero 1993:33).

Hay que considerar que, por la relación entre términos y conceptos, es viable esperar que todos los términos que pertenecen al corpus del diccionario sean posible palabras clave para otros términos.

IV. Ejemplo de una estructura lexical

Como una aplicación concreta se tiene la terminología del área de desastres y protección civil (Sierra 1992), la cual se encuentra estructurada gracias al marco conceptual del Área de Investigación Interdisciplinaria de Desastres (IID), en la que se ha venido trabajando en el Instituto de Ingeniería.

Por restricciones de espacio, en este artículo se describe la configuración léxica para algunos términos selectos del glosario y se proporciona un ejemplo. Se escogieron las calamidades o fenómenos destructivos —en particular “maremoto”— porque, a diferencia de otros términos, normalmente cualquier usuario sabe distinguirlos. Esto es, cualquiera sabe lo que es un sismo, un incendio o una inundación, pero no cualquiera conoce sus características o las actividades necesarias para contrarrestarlas. Además,

dado que las calamidades son objetos que suelen expresarse funcionalmente por nombres (Cabré 1993:180), la ambigüedad de términos es menos frecuente, dándose solo casos de sinonimia.

Para escoger los paradigmas se tomaron como base los siguientes aspectos:

a) *Clasificación a la que pertenece*

Como es de esperarse, debe considerarse la relación del término en el conjunto de términos, lo cual se permite gracias al marco conceptual. En el caso específico de la terminología de desastres, en todas las definiciones deben aparecer primero dos palabras significativas, las cuales son necesarias considerar como clave en la búsqueda onomasiológica: *calamidad* y el tipo de calamidad.

Para el caso del primer paradigma, *calamidad*, se consideraron como palabras asociadas, en primera instancia, sus sinónimos (fenómeno-destructor y agente-perturbador) y, con ello, las variantes semánticas (por ejemplo, fenómeno=manifestación; destructor=destructor, devastador, desastroso) y morfológicas (género y número). En cuanto al tipo, en México existe una clasificación de calamidades en cinco grupos (Gelman *et al.* 1992:73), conforme se observa en la siguiente tabla.

GRUPO	CALAMIDADES
Geológicas	agrietamiento, colapso de suelos, deslave y deslizamiento de talud, erosión, flujo de lodo, hundimiento regional, maremoto, sismo, vulcanismo.
Hidrometeorológicas	avalancha de nieve, desforestación y desertificación, huracán, inundación, lluvia, nevada, sequía, helada, calor, tormenta de granizo, tormenta eléctrica, viento.
Físico-químicas	contaminación, envenenamiento, explosión, fuga y derrame de sustancias peligrosas, incendio, radiación.
Sanitarias	epidemia, plaga.
Socio-organizativas	accidente mayor, acto de locura, actos delictivos, sabotaje, crecimiento explosivo de la población, disturbios sociales, drogadicción y alcoholismo, efecto negativo por operar servicios, falla o error humano, interrupción de servicios, terrorismo, acción bélica.

En el caso de maremoto, esta se encuentra clasificada dentro de las calamidades geológicas; debido al ambiente de origen de la misma, por lo que el paradigma se forma por *geológica* (líder) y por *geofísica, terrestre*, etc.

b) *Acción o acciones en las que consiste la calamidad*

Cada uno de los fenómenos destructivos se puede caracterizar por una acción o forma en que se manifiestan, lo que define el paradigma. Por ejemplo, *derrumbe o reacomodo, aumento o disminución, acumulación o carencia*, etc.

Para conceptos que entran en la misma clasificación es frecuente observar que ciertos paradigmas presentan rasgos distintivos en común, de tal forma que contienen algunas palabras clave repetidas. Por ejemplo, todas las calamidades geológicas presuponen el sentido “movimiento”, el cual es considerado como palabra clave en diferentes paradigmas, como puede apreciarse en la siguiente tabla.

TERMINO	LIDER	PALABRAS CLAVE
Colapso de suelo	derrumbe	caída, movimiento
Deslizamiento de talud	derrumbe	ídem
Erosión	desgaste	desprendimiento, movimiento
Vulcanismo	expulsión	erupción, liberación, movimiento
Hundimiento	reacomodo	descenso, caída, movimiento
Sismo	desacomodo	separación, ruptura, movimiento
Maremoto	oleaje	oscilación, levantamiento, movimiento

En el caso de maremoto, la acción definida es el *oleaje* (líder), que también presupone el sentido de *movimiento*, además del de *oscilación* y *levantamiento*, las cuales son consideradas palabras clave asociadas al paradigma.

c) *Sujeto sobre el cual sucede la acción*

La clasificación sobre los fenómenos destructivos se basa en el origen de cada uno, esto es, el ambiente o agente sobre el que se generan, con lo cual se define el siguiente paradigma.

Así se tiene que puede moverse el *suelo* durante el sismo, que el material arrojado por el volcán es *magma*, que en el calor aumenta la *temperatura*, que lo ingerido en un envenenamiento son los *tóxicos*, etc.

Las palabras itálicas del párrafo anterior son los líderes de los paradigmas referidos al sujeto sobre el que se manifiesta el fenómeno, los cuales encierran otras palabras asociadas. En lo posible, es conveniente que las palabras clave de distinto significado pertenezcan al mismo paradigma cuando pueden ser agrupadas con el mismo sentido. Tal es el caso de *magma*, líder del paradigma que representa todas las cosas que también se pueden expulsar de un volcán: *lava*, *fuego*, *rocas*, *cenizas*, *gases*, *vapor* y *fumarolas*.

El agente sobre el que se manifiesta el maremoto es *ola* (líder), *onda*, *marea* o *agua*.

d) *Manera en que se desarrolla la acción*

Un aspecto significativo para los usuarios del DEBO y que debe considerarse es la característica de la acción o la manera en que ésta se desarrolla. Para las calamidades se utilizaron las siguientes características: *súbito*, *violento* o *gradual*; *extremoso* y *desproporcionado*; *extenso*, *masivo* o *concentrado*; *voluntario* o *involuntario*. La inclusión de la característica es muchas veces determinante para la diferencia-

ción de los conceptos, ya que mientras el hundimiento es paulatino, el colapso de suelos es súbito; la inundación es extensa, mientras la lluvia es concentrada; los accidentes son imprevistos, en tanto el sabotaje es premeditado.

La característica del oleaje en los maremotos es *súbito*, usado como líder del paradigma, aunque entre las palabras claves del mismo están: *rápido, instantáneo, repentino, inmediato, inesperado*, e incluso *violento, brusco y brutal*.

e) *Lugar donde se presenta la calamidad*

Un aspecto muy recordado por los usuarios en la búsqueda de una calamidad determinada, en particular para aquellos que tienen un campo de especialidad muy marcado, es el lugar en el que ocurre el fenómeno destructivo, es decir, donde se manifiestan los impactos. Entre los paradigmas referidos al lugar, se consideran: la *superficie* o la *corteza* (terrestre); el *clima* o la *atmósfera*; el *suelo*, el *mar* o el *agua*; las *personas* o las *ciudades*; las *instalaciones* o los *servicios*.

Para el ejemplo del maremoto, el lugar donde éste ocurre o se genera es en el *mar*, que es el líder de *océano y agua*, así como de *costa, golfo, playa e isla*.

f) *Causa o causas que originan la calamidad*

Continuando con el proceso, la definición de los subsiguientes paradigmas para las calamidades se realiza a partir de las causas que las originan.

La determinación de las causas requiere conocer el mecanismo de producción de cada fenómeno destructivo (Gelman y Macías 1983:8). Si bien son muy variadas las causas y cada una permite, por sí misma, en muchos casos distinguir a la calamidad, se procuró minimizar el número de paradigmas, tanto englobando en un paradigma una amplitud de palabras clave asociadas, como aplicando el mismo paradigma a varios términos. Por ejemplo, es bien conocido que el hundimiento de la ciudad de México es debido, en parte, por la extracción del agua de los mantos acuíferos para el abastecimiento de agua potable a la población. Por ello, dos paradigmas son *extracción y freático*; este último agrupa a *manto, acuífero, subterráneo, agua, intersicial y subsuelo*.

Debido a la interrelación que existe entre los fenómenos destructivos, de tal forma que uno es consecuencia del otro, se llegan a encontrar como paradigmas de las causas a los términos que designan calamidades. De hecho, en el caso del maremoto, las causas son *sismo y vulcanismo*, siendo ambos sus paradigmas.

Como ejemplo de la estructuración de los términos, lo que representa las posibilidades de entrada en su búsqueda onomasiológica, en la siguiente tabla se presentan diez calamidades y los líderes de los paradigmas, con base en los criterios mencionados (se omite la “clasificación”, ya que esta se presentó en el rubro correspondiente). Se puede dar el caso, sobre todo en “causa”, que se tengan dos o más paradigmas.

CALAMIDAD	ACCION	SUJETO	FORMA	LUGAR	CAUSA
vulcanismo	expulsión	magma	súbito	corteza	sismo
maremoto	oleaje	ola	súbito	mar	sismo/vulcanismo
sequía	carencia	humedad	extenso	suelo	naturaleza/clima/cañor
inundación	acumulación	agua	extenso	superficie	drenaje/lluvia
fuga de sust.pel.	derrame	sustancia	extenso	suelo/agua	negligencia/tecnología
incendio	expansión	fuego	súbito	combustible	negligencia/tecnología
epidemia	contagio	enfermedad	extenso	persona	higiene
plaga	crecimiento	insecto	despropor- cionado	suelo/agua/ ciudad	higiene
drogadicción	ingestión	drogas	voluntario	persona	sociedad
terrorismo	acto	destrucción	voluntario	instalación/ ciudad	desacuerdo

Para terminar, conviene advertir dos cosas. Primero, los líderes no son exhaustivos, esto es, el conjunto de ellos no proporcionan la definición completa de un término. Segundo, los líderes son subjetivos y no necesariamente representativos de las palabras clave asociadas, sino más que nada para seguir este proceso, dentro de las bases de datos. Se buscó tener el menor número de paradigmas porque, evidentemente, a mayor número de paradigmas, el sistema es más complejo y, por tanto, más lento.

IV. Observaciones finales

El DEBO pretende ayudar a un usuario en la búsqueda de términos de los que conoce su significado pero que en determinado momento no recuerda la palabra o palabras para designarlo. En este proceso de búsqueda, al usuario pueden venirsele a la mente desde palabras aisladas relacionadas con el significado del término que busca hasta la definición del mismo. En este sentido, el DEBO debe prever, en lo posible, las palabras que una persona usaría en esta búsqueda.

Sin embargo, en el proceso de búsqueda de términos a través del DEBO, el usuario puede encontrarse con tres limitantes principales que debe tomar en cuenta para obtener una mayor eficiencia.

La primera está relacionada con la información contenida en la base de datos *PALCLV*, ya que la estructura de las palabras clave está basada en el significado extenso de los términos —el cual es normalmente más amplio que los dados por los diccionarios o las terminologías—, pero no en toda la información contenida en una enciclopedia, razón por la cual es aconsejable que el usuario piense más en el significado del término a buscar.

La segunda, en cuanto al resultado de la búsqueda, el usuario puede encontrar términos que aparentemente no tienen relación con la información introducida, lo cual es debido a la misma estructuración del léxico en el DEBO, en donde se repiten palabras clave en diferentes paradigmas. Sin embargo, para evitar una lista extensa de términos finales, el DEBO prevé, además de la ordenación jerarquizada, la presentación por partes de los resultados, de tal forma que aparezca de la más a la menos probable, a petición del usuario.

La tercera y última es concerniente al procedimiento de búsqueda, en donde el usuario puede introducir una o varias palabras a la vez. Debido a que el DEBO realiza un análisis conjunto de todas las palabras captadas, es aconsejable que el usuario introduzca en un solo paso el mayor número de palabras por búsqueda.

El DEBO se ha desarrollado en una versión prototipo, cuya copia se puede solicitar al autor, para 33 fenómenos destructivos y ha sido probado con buenos resultados. No obstante a la fecha se ha tratado en particular con la terminología en el área de desastres, se ha comprobado su aplicabilidad a otras áreas (hidrometeorología) y campos semánticos específicos (vehículos de transporte).

Para el diseño de la versión final se ha contemplado que el DEBO permita indistintamente tanto la búsqueda onomasiológica como la semasiológica, ésta última proporcionando los sinónimos y la definición de los términos, así como permitiendo la posibilidad de manejar hipertextos. Además, con la inclusión del diccionario semasiológico se podrá verificar con la definición que el término proporcionado es el que se estaba buscando.

Se aprovecha la oportunidad de agradecer la participación de Ismael Noguérón Vázquez, quien programó la versión prototipo del DEBO, determinó gran parte de los paradigmas y proporcionó valiosas sugerencias durante el desarrollo del proyecto.

Bibliografía

- ALPÍZAR, R. (1994) "El léxico de la terminología. Algunas precisiones", *IV Simposio Iberoamericano de Terminología*, Tomo II. Buenos Aires, pp. 79-87.
- BALDINGER, K. (1970) *Teoría semántica: hacia una semántica moderna*. Madrid: Alcalá.
- CABRE, M.T. (1993) *La terminología: Teoría, metodología, aplicaciones*. Barcelona: Antártida/Empúries.
- COSERIU, E. (1986) *Principios de semántica estructural*. Madrid: Gredos.
- BARBOSA, A. y FERNANDES, D. (1994) "Desenvolvimento de um banco de termos estruturado na área de gestao da produgao: um tesouro e um glossaário", *IV Simposio Iberoamericano de Terminología*, Buenos Aires, pp 151-156.
- GELMAN, O. et al (1992) *Marco conceptual para el estudio, prevención y atención de desastres tecnológicos y ecológicos*. Tomo 2. México: UNAM.
- . y MACÍAS, S. (1983) "Desastres y su pronóstico". *Boletín IMPOS XIII*-69.
- LARA, F. (1991) *Expresiones lingüísticas de probabilidad en la prospectiva tecnológica*. México: Instituto de Ingeniería.
- LARA, L.F. (1989) *Investigaciones lingüísticas en lexicografía*. México: El Colegio de México.
- POZZI, M. (1994) "La normalización de la terminología en español", *IV Simposio Iberoamericano de Terminología*, Tomo II. Buenos Aires, pp. 89-95.
- QUILES, M.E. y ZÍTARA, E. (1994) "La gestión integral de los recursos hídricos en un tesouro: terminología propia y compartida", *IV Simposio Iberoamericano de Terminología*, Buenos Aires, pp. 95-100.
- SAUSSURE, F. (1965) *Cours de Linguistic/ue Generóle*. Paris: Payot.
- SIERRA, G. (1992) "El diccionario de desastres del GIID: un informe de su avance". *Estudios de Lingüística Aplicada* 15/16, 404-411.
- (1994) "Bases del diccionario onomasiológico computarizado en el área de desastres", *IV Simposio Iberoamericano de Terminología*, Tomo II. Buenos Aires, pp. 55-59.
- (1995) "Outline of an onomasiological dictionary software in the disaster area", *Terminology Science & Research* 6-2, pp. 18-30.