

Clasificación semántica de colocaciones verbales para la adquisición y codificación de conocimiento experto

El caso de los riesgos naturales

Silvia Montero Martínez y Míriam Buendía Castro

Universidad de Granada / Universidad de Castilla-La Mancha

Los trabajos lexicográficos y terminográficos, por lo general, han favorecido el estudio de las unidades nominales en detrimento de las unidades verbales, a pesar de que los verbos son elementos claves en la transmisión de conocimiento general y experto (L'Homme, 1998). Este artículo propone una metodología de clasificación y descripción de las colocaciones verbales, de acuerdo con su significado, con la finalidad de facilitar la adquisición y codificación de conocimiento experto. La hipótesis de trabajo es que la clasificación de verbos, en dominios y subdominios léxicos, de acuerdo con el Modelo de la Gramática Léxica (Faber & Mairal Usón, 1999), y la identificación de los requisitos y restricciones de cada subdominio, de acuerdo con el tipo de argumento que activan los predicados, permitirá predecir el significado de las colocaciones verbales resultantes y establecer generalizaciones dentro de cada subdominio a través de patrones fraseológicos (Montero Martínez, 2008). Para ello, se propone una descripción de los argumentos que atienda a su categoría semántica, a los roles que activan (roles temáticos y macro-roles de la Gramática del Papel y la Referencia) (Van Valin & LaPolla, 1997; Van Valin, 2005) y a su estructura morfosintáctica. Las descripciones resultantes, ilustradas para colocaciones verbales en español del área del medio ambiente y representadas en la base de conocimiento *EcoLexicon*, resultarían muy valiosas en recursos lexicográficos y terminográficos destinados a la producción textual y la adquisición de conocimiento por parte del usuario.

Palabras clave: colocaciones, verbos, fraseología, Modelo de la Gramática Léxica, Terminología basada en Marcos, *EcoLexicon*

1. La fraseología y las colocaciones verbales

La *fraseología* se encarga del estudio de las frases o unidades fraseológicas. No obstante, al igual que ocurre con otros muchos fenómenos lingüísticos, aún no hay consenso en cuanto a la denominación de dichas unidades, sus propiedades y los límites de la disciplina (Cowie, 2012, p. 389). Esta situación permite trazar un *continuum* desde una concepción restringida de la fraseología (Rossenbeck, 1989; Zuluaga, 1997; García-Page, 2008) hasta una concepción amplia (Gläser, 1994/1995; Roberts, 1998; Heid, 2001; Corpas Pastor, 2003, 2008; Montero Martínez, 2003, 2008).

Partiendo de una mayor a menor restricción, hay estudios que consideran como unidades fraseológicas únicamente a las expresiones fijas, tales como las locuciones idiomáticas (*idioms*), otros incluyen los proverbios y dichos, algunos las colocaciones y, finalmente, otros defienden que todos los tipos de unidades pluri-lexemáticas frecuentes en el discurso, incluidos los compuestos, son parte de la fraseología (Pawley, 2001, p. 122). El desacuerdo surge, fundamentalmente, a raíz de los distintos factores que se consideran determinantes en la identificación de estas unidades; por ejemplo, el grado de opacidad semántica, la restricción sintáctica y la frecuencia de aparición. Estos aspectos permean la mayor parte de las clasificaciones de unidades pertenecientes al fenómeno fraseológico. En consecuencia, una definición genérica de estas unidades que recoja el amplio espectro existente en su estudio sería la siguiente: unidad léxica formada por dos o más elementos, de uso frecuente en el discurso, que presenta un grado variable de lexicalización, una estabilidad sintáctica y semántica y un posible carácter idiomático, connotativo y enfático (adaptado de Gläser, 1998, p. 125).

Este artículo se ubica dentro de los enfoques menos restrictivos, en donde tiene cabida el estudio de las *colocaciones*. Tradicionalmente, estas unidades se han abordado con un enfoque semántico y/o estadístico. La perspectiva semántica atribuye a los componentes de la colocación cierta *polaridad* (Firth, 1968; Mel'čuk et al., 1984–1999; Hausmann, 1989; Benson, Benson, & Ilson, 1986, 2009), ya que aparece una *base* con un significado léxico pleno y un *colocado* con un significado modificado o reducido (Heid, 1994, p. 234). El componente estadístico implica que es necesaria una alta frecuencia de *co-aparición* de los componentes de una unidad léxica en un corpus para que esta tenga la categoría de colocación (Sinclair, 1991, 2007; Cowie, 2012). Además, a estas unidades se les atribuyen por lo general patrones morfosintácticos concretos, que están sancionados por los usos lingüísticos de una comunidad de hablantes. Por tanto, la forma léxica de una colocación está restringida en el nivel pragmático; la norma de una lengua establece ciertas combinaciones que se adquieren y usan como bloques por parte de los hablantes mientras que bloquea otras, gramaticalmente correctas, que

responderían a combinaciones libres (Heid, 1992, p. 534) de la lengua y no a fenómenos fraseológicos.

Así pues, la noción de colocación es difusa en los estudios actuales y parece que solamente se puede precisar su definición teniendo en cuenta una aplicación determinada y un marco teórico concreto (Williams, 2013, p. 91). La única observación indiscutible es que determinadas unidades léxicas tienden a unirse y formar asociaciones más o menos fuertes y que las colocaciones son centrales en la transmisión de significado, ya que las palabras no funcionan de forma aislada sino que adquieren su significado en contexto y, por tanto, a través de la fraseología (Moon, 2008, p. 243, Williams, 2013, pp. 91–94). Esta perspectiva implica la necesidad de una aproximación empírica al estudio de las colocaciones que tenga en cuenta el estudio de los patrones léxicos y el significado de las palabras dentro de estos (Sinclair, 2007, p. 156).

En este trabajo, se entiende por colocación la combinación de dos o más unidades léxicas, formada por un sustantivo + verbo o verbo + sustantivo, cuya semántica responde a un patrón conceptual propio de un dominio de especialidad. En estos ámbitos restringidos, por ejemplo el medio ambiente, el significado del colocado (el predicado verbal) está condicionado y, en muchos casos, impuesto por el significado de la base (el sustantivo). No obstante, a su vez, el predicado restringe la selección de posibles argumentos (los sustantivos) (Montero Martínez, 2008, p. 1019). Por ejemplo, en la colocación ‘el fuego arde’, el predicado ARDER sólo admite argumentos que puedan estar, literalmente, en combustión (ej. el fuego), y el argumento ‘fuego’ precisa de un verbo que designe un proceso de combustión (ej. ARDER). Así pues, esta aproximación al concepto de colocación presenta premisas estadísticas y semántico-conceptuales que precisan la descripción tanto del comportamiento de los argumentos como de los predicados verbales de las colocaciones.

Este estudio se basa en los postulados teóricos de la Terminología basada en Marcos (Faber, 2009, 2011, 2012) y en su aproximación a la descripción de las construcciones fraseológicas (Montero Martínez, 2003, 2008; Buendía Castro, 2013) en la base de conocimiento *EcoLexicon* (especializada en el medio ambiente). Desde esta perspectiva, se ilustra una metodología para la identificación y descripción de las colocaciones verbales que pueda hacerse extensible a la elaboración de recursos terminográficos y lexicográficos destinados a facilitar la adquisición y codificación de conocimiento experto. Dicha metodología se aplica en el área del medio ambiente, específicamente en el ámbito del EVENTO EXTREMO¹ (entendido como ‘desastre natural’) y en particular al concepto INCENDIO.

1. Siguiendo las convenciones ortotipográficas propias de la conceptología y la terminología, en este artículo se usan sistemáticamente las versales para referirse al plano conceptual y ontológico

2. El dominio conceptual del medio ambiente en EcoLexicon

De acuerdo con la Terminología basada en Marcos (TBM) (Faber, 2009, 2011, 2012), los distintos *dominios de especialidad* se configuran en torno a una serie de macro-estructuras denominadas *eventos*. Estos son independientes de la lengua y sirven para llevar a cabo la representación prototípica del conocimiento experto de cada dominio (Grinev & Klepalchenko, 1999).

Cada evento está conformado, a su vez, por *subeventos* y *categorías conceptuales*, tanto genéricas como específicas, que facilitan la estructuración de los distintos tipos de información (conceptual, léxica, multimedia, etc.) relevantes en la configuración del conocimiento experto. Dichas estructuras son dinámicas, flexibles y se encuentran interconectadas de forma parecida al sistema de conceptualización humano (Faber, León Araúz, & Prieto Velasco, 2009). Para la creación de los eventos, la TBM combina principios de la Lingüística de Corpus, del Modelo de la Gramática Léxica (Faber & Mairal Usón, 1999) y de la Semántica de Marcos (Fillmore & Atkins, 1992).

Por ejemplo, el EVENTO MEDIOAMBIENTAL (EM) (Figura 1) permite representar el dominio del medio ambiente a través de una estructura basada en un enfoque *top-down* (análisis de diccionarios y consulta a expertos) y *bottom up* (análisis de corpus). El resultado es un macro-evento prototípico de dicho dominio que está conformado por las siguientes subestructuras: AGENTE, PROCESO/ESTADO, PACIENTE/RESULTADO, LUGAR y DESCRIPCIÓN. Estas a su vez incluyen las categorías conceptuales genéricas que están presentes en el dominio: AGENTE NATURAL, AGENTE HUMANO, PROCESO NATURAL, etc. Dicho de otro modo, el EVENTO MEDIOAMBIENTAL proporciona un marco general en el que se interrelacionan conceptos más específicos (pertenecientes a las categorías genéricas) a través de relaciones jerárquicas (ej. AGUA *es-un* AGENTE NATURAL) y no jerárquicas (ej. AGUA *causa* MOVIMIENTO). El establecimiento de las categorías conceptuales, más o menos específicas, sigue el siguiente proceso: (i) identificación de las relaciones semánticas de los conceptos expresadas por las realizaciones lingüísticas en un corpus, (ii) verificación a través de test basados en *las clases d'objets* de Gross (1994).

Concretamente, según la Figura 1, el EVENTO MEDIOAMBIENTAL tiene dos tipos de AGENTES que pueden iniciar un proceso, agentes inanimados (ej. las fuerzas naturales) y agentes animados (ej. los seres humanos). Los AGENTES NATURALES, como los movimientos de agua y los fenómenos atmosféricos causan PROCESOS en

de las realidades mencionadas y las minúsculas para referirse a sus realizaciones lingüísticas en una lengua concreta.

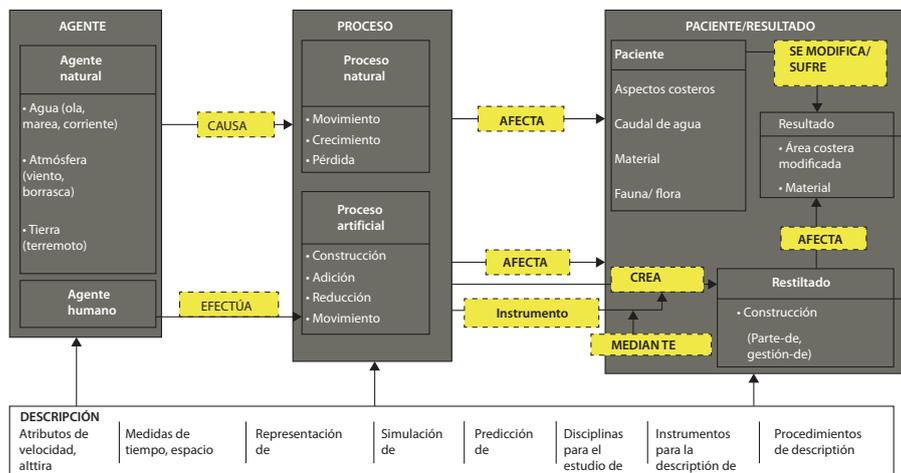


Figura 1. El EVENTO MEDIOAMBIENTAL

un área geográfica. Estos procesos, a su vez, afectan a otras entidades o PACIENTES que, como resultado, pueden sufrir cambios. Los AGENTES HUMANOS con la ayuda de INSTRUMENTOS también pueden llevar a cabo PROCESOS ARTIFICIALES (ej. construcción), que pueden generar o evitar los EFECTOS producidos por procesos naturales. Así pues, AGENTE, PACIENTE, RESULTADO e INSTRUMENTO son los roles semánticos característicos de este dominio de especialidad y el EM representa sus interrelaciones. No obstante, conviene señalar que también hay categorías y roles periféricos que incluyen conceptos que se usan para medir, analizar y describir los procesos en el evento.

Esta estructura conceptual del EM se materializa en una base de conocimiento especializado llamada *EcoLexicon*.² A través de un tesoro visual, todas las entradas se vinculan a su correspondiente (sub)evento y categoría conceptual. Dicho de otro modo, la información gráfica, conceptual y lingüística relativa a cada entrada (incluidas las colocaciones verbales) se estructura en torno al sub-evento conceptual subyacente.

3. Los predicados verbales en el ámbito de los desastres naturales

La especificación conceptual del EM mediante la creación de sub-eventos como el EVENTO EXTREMO (entendido en el sentido de ‘desastre natural’ o ‘riesgo natural’) sigue también un enfoque *top-down* y *bottom-up*. No obstante, para la descripción de las unidades léxicas (en este caso de las colocaciones verbales) que activan el

2. Disponible en <http://ecolexicon.ugr.es>

área semántica del EVENTO EXTREMO es también necesario el uso de corpus, la factorización de información procedente de definiciones, la aplicación de los dominios léxicos del Modelo de la Gramática Léxica y el análisis pormenorizado de los argumentos y predicados de las construcciones verbales (Montero Martínez & Buendía Castro, 2010, 2012; Buendía Castro, 2013).

3.1 Descripción del corpus

Para llevar a cabo la descripción de las colocaciones del ámbito de los desastres naturales, se ha compilado un corpus temático siguiendo la metodología *web for corpus (WfC)* (De Schryver, 2002). Así pues, se ha utilizado internet como fuente de textos que, posteriormente, se han gestionado como un corpus *off-line*, almacenado en una aplicación al efecto. No obstante, dada la limitación de fuentes electrónicas en español, el corpus se completó con textos escaneados.

Para asegurar la fiabilidad de todos los textos, se aplicó un protocolo de evaluación de recursos que se apoya en los siguientes parámetros (Buendía Castro & Ureña, 2009): (i) la *autoría*, es decir, la reputación y pericia de los autores responsables de los contenidos del sitio web; (ii) el *contenido*, que incluye la cobertura, la precisión, la objetividad, la actualización y el nivel de especialización; (iii) el *diseño*, que se refiere a las ayudas de navegación, la accesibilidad, la presentación y la gestión de la información. Como resultado, el corpus compilado responde a las características descritas en la Tabla 1.

Tabla 1. Características generales del corpus

Datos del corpus	
Número de <i>tokens</i>	449 416
Número de <i>types</i>	31 230
<i>Type/token ratio</i>	7,24
<i>Standardized TTR</i>	41,28
Número de textos	44
Medio	escrito
Subdominio	EVENTO EXTREMO
Cronología	1996 a 2014
Lengua	español (peninsular)
Fuente de los textos	actas de congresos, monografías, revistas divulgativas

Se trata de un corpus *monolingüe* en español peninsular de carácter *sincrónico* y *plano*, ya que incluye textos desde el año 1996 hasta el año 2014. No contiene ningún tipo de anotación lingüística, si bien es cierto que se han añadido algunas

etiquetas para mejorar las búsquedas y la categorización de la información. Incluye textos de carácter *especializado*, *semiespecializado* y *divulgativo* sobre el EVENTO EXTREMO, ya que es preciso reflejar los distintos grados de especialidad presentes en este ámbito. En concreto, contiene 44 textos y cuatro de ellos son monografías.³ El corpus asciende a un total de 449 416 *tokens* y 31 230 *types*. El *standardized type/token ratio* (STTR), obtenido con *WordSmith Tools*, señala que el corpus tiene una media de 41,28 palabras diferentes por cada secuencia de 1 000 palabras, lo que da cuenta de su gran diversidad léxica.

Es preciso mencionar que el objetivo inicial de este trabajo contemplaba la compilación de un corpus más voluminoso, pero esto no fue posible dada la escasa disponibilidad de textos especializados y semiespecializados en español, sobre todo en formato electrónico. Así pues, manuales como Olcina (2006a, 2006b), originalmente redactados en español, tuvieron que escanearse y convertirse a formato *.txt* con un programa de reconocimiento óptico de caracteres, con la posterior revisión manual por la cantidad de errores que contenían. Obviamente, esta escasez de materiales obliga a los expertos y semiexpertos a consultar fuentes en inglés y materiales traducidos.

Dado que los traductores están contribuyendo a la formación de términos y a su interiorización por parte de los receptores (Montero Martínez, Fuertes Olivera, & García de Quesada, 2001), es necesario que los corpus incluyan material traducido fiable y de calidad para que reflejen las unidades terminológicas y fraseológicas *in vivo*, que son las que recogen recursos descriptivos como *EcoLexicon*. Por ejemplo, la monografía de Keller y Blodgett (2007), incluida en el corpus, se ha traducido al español peninsular por una traductora experta (Pilar Gil Ruiz) y se ha revisado por varios expertos en la materia. Se trata de un libro de texto que aparece como referencia obligada en España, tanto en los programas de grado como de posgrado, en materias como la geología, la geofísica o las ciencias ambientales. También se han seleccionado artículos divulgativos traducidos de la revista *National Geographic*, que cuenta con unos estándares de calidad reconocidos públicamente.

Por último, además de utilizar el enfoque *WfC*, en este estudio se ha hecho uso del enfoque *web as corpus* (*WaC*), que explota la red como si fuera un *corpus* propiamente dicho (Baroni & Bernardini, 2006; Fletcher, 2007). Este sistema ha servido para encontrar ejemplos de uso concretos cuando los contextos del corpus primigenio resultaban insuficientes y para comprobar o refutar una hipótesis a través del estudio de un mayor número de casos. Con este fin, se ha hecho uso de

3. En el anexo1 se puede encontrar la relación de los textos que componen el corpus.

programas de pre- y post-procesamiento de datos (*WebCorp*)⁴ y de programas de gestión de corpus (*Sketch Engine*).⁵

3.2 Análisis y estructuración del EVENTO EXTREMO y la categoría DESASTRES NATURALES

La estructuración conceptual del EVENTO EXTREMO se ha delimitado mediante la descripción preliminar de la situación activada por dicho evento y tras un proceso de análisis y *factorización* de las definiciones de los conceptos más salientes, consultadas en diccionarios, tesauros y fuentes de referencia. Entre otros, fueron de especial ayuda el *Tesaurus Multilingüe de Medio Ambiente (GEMET)* de la Agencia Europea de Medio Ambiente,⁶ el *Tesaurus Multilingüe de la Unión Europea EuroVoc*,⁷ el compendio *UNISDR Terminología sobre Reducción del Riesgo de Desastres*⁸ de la Oficina de las Naciones Unidas para la Reducción del Riesgo de Desastres y el portal web *ScienceDaily*.⁹

Así pues, se ha podido determinar que los conceptos específicos de este (sub) evento incluyen los siguientes: VOLCÁN, DESLIZAMIENTO DE TIERRA, INUNDACIÓN, AVALANCHA, CICLÓN TROPICAL (incluidos TIFÓN y HURACÁN), TORNADO, SEQUÍA, TERREMOTO, FUEGO y TSUNAMI. Todos ellos forman parte de la categoría conceptual DESASTRE NATURAL. No obstante, otras categorías conceptuales salientes en el evento son ACCIDENTE GEOGRÁFICO, AGENTE ACUÁTICO, AGENTE ATMOSFÉRICO, ÁREA, CONDICIÓN ATMOSFÉRICA, CONSTRUCCIÓN, CURSO DE AGUA, DAÑO, ENERGÍA, EXPLOSIVO, MATERIAL, MUERTE, PÉRDIDA DE VIDA, PÉRDIDA DE PROPIEDAD, PLANTA y SER HUMANO. La Figura 2 muestra la estructuración resultante del EVENTO EXTREMO en *EcoLexicon*, a través de la red conceptual que subyace entre conceptos pertenecientes a las categorías anteriormente mencionadas,¹⁰ así como la definición del propio concepto de EVENTO EXTREMO.

4. Disponible en <http://www.webcorp.org.uk/live>

5. Disponible en <http://www.sketchengine.co.uk/>

6. Disponible en <http://www.eionet.europa.eu/gemet/>

7. Disponible en <http://eurovoc.europa.eu/drupal/>

8. Disponible en http://www.unisdr.org/files/7817_UNISDRTerminologySpanish.pdf

9. Disponible en <http://www.sciencedaily.com/>

10. Para una mayor legibilidad en la imagen, solo se muestran algunas relaciones de tipo hipónimo y meronímico. Se remite a la base de conocimiento *EcoLexicon* para una visualización completa de las relaciones conceptuales del EVENTO EXTREMO.

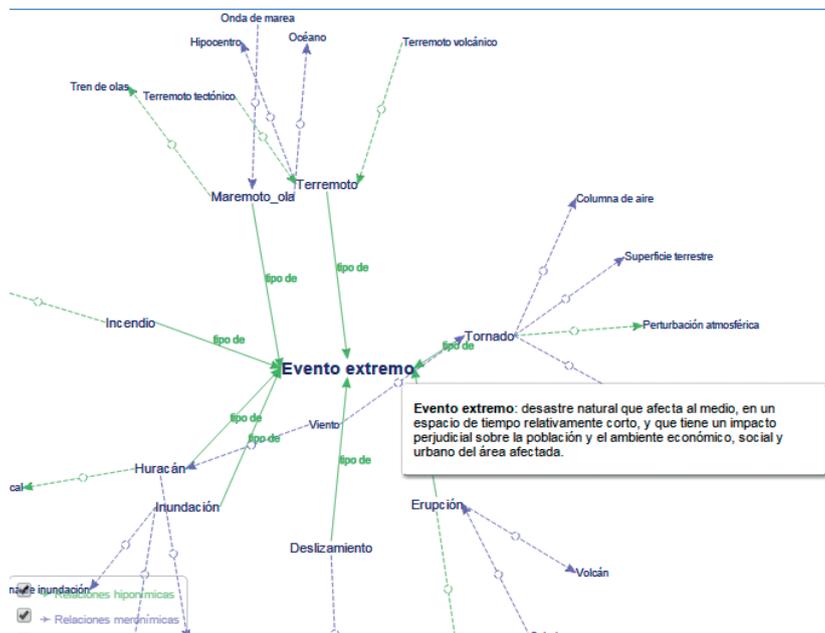


Figura 2. El EVENTO EXTREMO en EcoLexicon

3.3 Extracción de candidatos a verbo

Tras la delimitación conceptual del EVENTO EXTREMO y la categoría DESASTRES NATURALES, se han extraído los verbos más recurrentes del corpus con el extractor terminológico *TermoStat*, una herramienta desarrollada por Drouin (2003) en la Universidad de Montreal.

TermoStat identifica los términos de forma estadística, mediante la comparación de las frecuencias obtenidas en un corpus específico y en uno general. El análisis automático generó un total de 323 lemas considerados *estadísticamente relevantes* en este contexto. En la Figura 3, se muestran algunos de estos *candidatos* a verbos a través de una lista lematizada (primera columna de la izquierda), seguida de su frecuencia absoluta en el texto, la puntuación que se le asigna en función de su especificidad en el ámbito de especialidad y, por último, sus variantes ortográficas. Del total de candidatos extraídos automáticamente, se descartaron aquellos casos que, a pesar de haber sido etiquetados como verbos, se correspondían con otras categorías sintácticas (ej. 'pronosticadores' o 'embudo'). Los verbos restantes son unidades específicas del ámbito de los DESASTRES NATURALES (ej. 'quemar', 'deslavar', 'arrasar', 'propagar' o 'desbordar') o unidades semánticamente relevantes para la descripción de dicho ámbito (ej. 'acrecentar', 'analizar', 'causar', 'ocurrir', 'oscilar', 'variar').

Candidat de regroupement	Fréquence	Score (Spécificité) ‡	Variantes orthographiques	Matrice
quemar	42	63.81	<i>quemar</i> <i>quemar</i>	Verbe
deslavar	6	43.71	<i>deslaves</i>	Verbe
causar	72	37.51	<i>causar</i> <i>causan</i> <i>causen</i>	Verbe
arrasar	11	37.15	<i>arrasar</i> <i>arrasa</i> <i>arrasaron</i>	Verbe
rachar	7	36.46	<i>racha</i>	Verbe
propagar	20	36.1	<i>propagar</i> <i>propagan</i> <i>propagó</i>	Verbe
picar	15	35.55	<i>pico</i>	Verbe
ensanchar	9	35.14	<i>ensanche</i>	Verbe
pronosticadores	4	34.06	<i>pronosticadores</i>	Verbe
embudo	4	34.06	<i>embudo</i>	Verbe
romperán	4	34.06	<i>romperán</i>	Verbe
tormentar	5	33.07	<i>tormentas</i>	Verbe
ocurrir	116	30.12	<i>ocurrir</i> <i>ocurrirá</i> <i>ocurrirán</i> <i>ocumeron</i>	Verbe
desbordar	13	29.86	<i>desborda</i> <i>desbordarse</i>	Verbe

Figura 3. Extracción de candidatos a verbo del corpus (*TermoStat*)

3.4 Clasificación de predicados verbales en dominios y subdominios léxicos

Siguiendo el Modelo de la Gramática Léxica (Faber & Mairal Usón, 1999), los verbos seleccionados se clasificaron en *dominios* y *subdominios léxicos*, de acuerdo con el *genus* y *differentiae* de su definición. En otras palabras, se han asignado a unas estructuras de significado que lexicalizan un dominio conceptual (en todo o en parte) y cuyos miembros se caracterizan por compartir el mismo significado nuclear y ciertos patrones sintácticos (Faber & Mairal Usón, 1999, p. 59).

Desde esta perspectiva, el lexicon verbal de las lenguas europeas se puede dividir en doce dominios léxicos. Cada dominio tiene uno o dos *verbos genéricos* o *superordinados*, a partir de los cuales se definen de forma directa o indirecta todos sus miembros. Así pues, la información semántica del *differentiae* es la que distingue a un verbo de otro dentro del mismo dominio. La Tabla 2 muestra los dominios léxicos (entre corchetes) y los verbos superordinados asociados a cada uno:

Tabla 2. Dominios léxicos del Modelo de la Gramática Léxica (traducción de Faber & Mairal Usón, 1999, p. 88)

-
- (i) *ser* [EXISTENCIA]
 - (ii) *convertirse en algo diferente* [CAMBIO]
 - (iii) *tener/dar* [POSESIÓN]
 - (iv) *decir* [HABLA]
 - (v) *sentir* [EMOCIÓN]
 - (vi) *hacer/fabricar* [ACCIÓN]
 - (vii) *usar* [MANIPULACIÓN]
 - (viii) *saber/pensar* [COGNICIÓN/PERCEPCIÓN MENTAL]
 - (ix) *mover (ir/venir)* [MOVIMIENTO]
 - (x) *ser consciente (notar/percibir)* [PERCEPCIÓN GENERAL]
 - (xi) *ver/oír/gustar/oler/tocar* [PERCEPCIÓN SENSORIAL]
 - (xii) *estar/permanecer/poner* [UBICACIÓN]
-

En consecuencia, el análisis de los verbos del corpus y su clasificación a través de dominios léxicos determina que, por lo general, los dominios más prototípicos o que se activan con mayor intensidad en el EVENTO EXTREMO SON: EXISTENCIA, ACCIÓN, MOVIMIENTO, UBICACIÓN y CAMBIO (Buendía Castro, 2012, p. 157). Esto se debe a que los conceptos activados por el EVENTO EXTREMO (ej. HURACÁN, TSUNAMI, TERREMOTO, etc.) son PROCESOS que empiezan a existir, existen durante un periodo de tiempo y, finalmente, dejan de existir. Durante su existencia (EXISTENCIA), actúan en un lugar determinado (UBICACIÓN) o se desplazan (MOVIMIENTO), lo que genera un cambio negativo (CAMBIO) que afecta a una entidad animada o inanimada (ACCIÓN).

Además, para perfilar el significado de los predicados, se ha estudiado su comportamiento en los textos, gracias a *WordSmith Tools*, a través del análisis de concordancias en el corpus *offline*. Cuando se necesitaron más instancias, se acudió a la *web* como se ha mencionado.¹¹ Esta mayor precisión semántica permite dividir cada dominio léxico en subdominios que se centran en un área de significado más concreta; es decir, los subdominios reflejan las diferentes especificaciones del contenido semántico del dominio. Por ejemplo, dentro del ámbito temático del EVENTO EXTREMO, el dominio de EXISTENCIA (Tabla 3) se subdivide en trece subdominios tales como ‘empezar a existir’, ‘empezar a existir a partir de otra cosa’, ‘empezar a existir convirtiéndose en otra cosa’, etc. (Buendía Castro, 2013, pp. 387–442).

11. Este análisis sirvió, además, para descartar por completo aquellos candidatos a verbo que no estaban directamente asociados con el EVENTO EXTREMO. En otras palabras, se desecharon todos aquellos casos cuyos argumentos no presentaban realizaciones lingüísticas que apuntaran a un desastre natural o un fenómeno relacionado.

Tabla 3. Subdominios léxicos de EXISTENCIA dentro del EVENTO EXTREMO**EVENTO EXTREMO (EXISTENCIA)**

-
- (i) empezar a existir
 - (ii) empezar a existir a partir de otra cosa
 - (iii) empezar a existir convirtiéndose en otra cosa
 - (iv) causar que exista/ocurra
 - (v) evitar que exista/ocurra
 - (vi) ser probable que ocurra (especialmente algo malo)
 - (vii) existir en el tiempo
 - (viii) continuar la existencia en el tiempo (de un desastre natural)
 - (ix) continuar la existencia en el tiempo (de las personas)
 - (x) cesar de existir
 - (xi) cesar de existir en la percepción de otros
 - (xii) causar que cese de existir (una persona)
 - (xiii) causar que cese de existir (un desastre incendiario)
-

A modo de ilustración, dentro del subdominio ‘causar que cese de existir (un desastre incendiario)’ (Tabla 4), se encuentran en español verbos como ‘apagar’, ‘extinguir’ y ‘sofocar’ cuyo *genus*, de forma directa o indirecta, se corresponde con el superordinado de la jerarquía.

Tabla 4. Realización léxica en español del subdominio ‘causar que cese de existir (un desastre incendiario)’

Causar que cese de existir (un desastre incendiario)

apagar: (*referido al fuego*), *hacer que cese o termine*

extinguir: (*referido al fuego*), *apagarlo o hacer que cese.*

sofocar: (*referido al fuego*), *extinguir un fuego.*

4. Los argumentos en el ámbito de los desastres naturales

En este contexto, las colocaciones verbales están formadas por la combinación frecuente de dos o más unidades léxicas, que siguen el patrón sintáctico sustantivo + verbo/ verbo + sustantivo y cuyo significado viene determinado por un patrón conceptual propio del ámbito de los desastres naturales. El significado del verbo (el colocado) viene impuesto por el significado del sustantivo, pero el verbo también restringe los posibles argumentos que se combinan con él. El proceso de análisis y descripción de las colocaciones verbales del corpus se ilustra a partir de los predicados pertenecientes al subdominio léxico ‘causar que cese de existir (un desastre incendiario)’. En este análisis *bottom up*, además de clasificar los predicados en dominios y subdominios léxicos, se analizan los argumentos de las colocaciones

verbales a partir de las líneas de concordancia ofrecidas por *WordSmith Tools* para cada verbo seleccionado.

Concretamente, el proceso seguido para el análisis de los argumentos incluye el estudio de los siguientes aspectos: (i) identificación de las diversas realizaciones lingüísticas de cada argumento y asignación a una categoría conceptual; (ii) asignación de un rol semántico y un macro-rol a cada argumento; (iii) clasificación de cada argumento según su estructura morfosintáctica. Finalmente, todos los datos recabados permiten establecer generalizaciones en el comportamiento semántico y sintáctico de las construcciones fraseológicas pertenecientes al subdominio léxico analizado. Estas se materializan a través de plantillas fraseológicas del subdominio.

4.1 Realizaciones lingüísticas y categorías conceptuales

Los conceptos pertenecientes a una misma categoría conceptual o semántica presentan rasgos comunes en mayor o menor grado (Montero Martínez, Faber, & Buendía Castro, 2011, pp. 103). Así pues, sus proyecciones léxicas en el discurso presentan comportamientos sintácticos y semánticos parecidos (Faber, 2009, p. 123).

Una vez establecidas las categorías conceptuales relevantes en el dominio conceptual del EVENTO EXTREMO, que incluyen ACCIDENTE GEOGRÁFICO, AGENTE ACUÁTICO, AGENTE ATMOSFÉRICO, ÁREA, CONDICIÓN ATMOSFÉRICA, CONSTRUCCIÓN, CURSO DE AGUA, DAÑO, ENERGÍA, EXPLOSIVO, MATERIAL, MUERTE, PÉRDIDA DE VIDA O PROPIEDAD, PLANTA y SER HUMANO, sus realizaciones lingüísticas en el corpus deben materializar las relaciones conceptuales propias de cada categoría. Por ejemplo, la categoría DESASTRE NATURAL mantiene las siguientes relaciones conceptuales con otros conceptos (Tabla 5):

Tabla 5. Principales relaciones conceptuales activadas por la categoría DESASTRE NATURAL

Un DESASTRE NATURAL *causa* PÉRDIDAS HUMANAS, ECONÓMICAS O MATERIALES

Un DESASTRE NATURAL *afecta* al MEDIO AMBIENTE

Un DESASTRE NATURAL *ocurre* en un PERIODO DE TIEMPO CORTO

Así pues, siempre que una realización lingüística cumpla con esas tres condiciones será susceptible de pertenecer a la categoría de DESASTRE NATURAL. Este es el caso de conceptos como TERREMOTO, HURACÁN y FUEGO, que causan pérdidas humanas, económicas y materiales, afectan al medio ambiente y ocurren en un espacio corto de tiempo. En consonancia, los términos y argumentos identificados en el corpus como proyecciones de la categoría de DESASTRE NATURAL en español

son los siguientes: ‘alud’, ‘avalancha’, ‘ciclón’, ‘ciclón tropical’, ‘corrimiento de tierra’, ‘depresión tropical’, ‘desastre natural’, ‘deslizamiento de tierra’, ‘erupción’, ‘fuego’, ‘incendio’, ‘inundación’, ‘huracán’, ‘lluvia’, ‘maremoto’, ‘seísmo’, ‘sequía’, ‘sismo’, ‘temblor’, ‘terremoto’, ‘tifón’, ‘tormenta tropical’, ‘tsunami’ y ‘volcán’.

Finalizada la identificación, cada conjunto de realizaciones lingüísticas que designan el mismo tipo de argumento en el corpus recibió una etiqueta semántica, visualmente diferenciada por colores, que se corresponde con la categoría conceptual pertinente.¹² Por ejemplo, la Tabla 6 muestra la etiquetación, en negrita, de argumentos pertenecientes a la categoría de DESASTRE NATURAL (‘incendio’, ‘fuego’ y ‘lluvias’) que coaparecen con el predicado SOFOCAR. Además, se han etiquetado realizaciones lingüísticas que se corresponden con otras categorías conceptuales o semánticas que se relacionan con los desastres naturales o sirven para describirlos. Este es el caso de ‘efectivos de los bomberos’ o ‘bomberos’, que se corresponden con la categoría de SER HUMANO; ‘12,15h’ que hace referencia al TIEMPO; ‘refinería’ que activa la categoría de CONSTRUCCIÓN; ‘rápidamente’, que se refiere a la MANERA, ‘Navarra’, que se refiere al LUGAR, y ‘grandes dimensiones’ que hace referencia a la MEDIDA de la superficie afectada.

Tabla 6. Etiquetación semántica de ejemplos de uso relativos a la categoría DESASTRE NATURAL

Etiquetación de argumentos	
	1. Al lugar de los hechos se desplazaron efectivos de los Bomberos , que sofocaron el incendio sobre las 12,15 horas .
	2. Sofocan el incendio en la refinería más importante de Venezuela.
	3. Cuatro unidades de bomberos sofocaron el fuego rápidamente en Navarra .
	4. Las lluvias consiguieron sofocar el fuego de grandes dimensiones .

4.2 Roles semánticos y macro-roles

Los argumentos identificados en el corpus se marcaron posteriormente con un *rol semántico*. Esta noción coincide fundamentalmente con los *roles temáticos* más generales de la Gramática del Papel y la Referencia (Van Valin & LaPolla, 1997; Van Valin, 2005) y con los *roles argumentales* propuestos por Goldberg en la Gramática de Construcciones (Goldberg, 1995, 2006). Además, los roles se completan con otros procedentes del inventario de recursos como VerbNet¹³ (*A class-based verb*

12. Este sistema de anotación semántica es similar al utilizado por FrameNet (Ruppenhofer, Ellsworth, Petruck, Johnson, & Scheffczyk, 2010) para los “*frame elements*”, no obstante en este caso solo se anota el núcleo de los sintagmas.

13. Disponible en <http://verbs.colorado.edu/verb-index/>

Tabla 7. Inventario de roles semánticos del EVENTO EXTREMO

Rol semántico	Definición y ejemplo
AGENTE HUMANO	Agente o entidad animada que lleva a cabo una acción de manera intencionada (ej. 'El incendio fue provocado por un pirómano ')
AGENTE NATURAL	Agente inanimado o proceso, generalmente una fuerza natural, que provoca una acción de forma involuntaria (ej. 'El fuego destruyó la ciudad')
DESTINO	Lugar hacia el cual se produce un desplazamiento (ej. 'El incendio se propagó hacia el exterior ')
EXPERIMENTADOR	Entidad animada (generalmente un ser humano) o una entidad inanimada personificada (construcción, planta, zona geográfica) que continua su existencia o cesa su existencia a consecuencia de un desastre natural. (ej. 'Poca gente sobrevivió al terremoto') (ej. 'El Miami Beach Resort ha sobrevivido a huracanes importantes en 1979 y 1992')
FRECUENCIA	Número de veces que algo ocurre durante un período o espacio determinado (ej. 'Este volcán ha entrado en erupción dos veces en los últimos seis meses')
LUGAR	Espacio en el que se desarrolla un evento, acción o estado (ej. 'Cientos de muertos, barcos y casas destrozadas por el tsunami en Japón ')
MEDIDA	Cuantificación del área afectada por el proceso (o entidad) (ej. 'Las lluvias sofocaron el fuego de grandes dimensiones ')
MANERA	Forma o características que distinguen cada realización de un evento, acción o estado (ej. 'El huracán destruyó completamente unas 1980 viviendas')
TRAYECTO	La dirección que sigue el movimiento (ej. 'El volcán arrojó rocas incandescentes alrededor del cráter ')
PACIENTE INANIMADO	Entidad inanimada no personificada que padece el efecto de un agente natural, generalmente una fuerza natural. (ej. 'El huracán devastó la costa ')
ORIGEN	Punto de partida desde el cual se mueve algo, proceso (o entidad) a partir del cual se forma otro (ej. 'El huracán <i>Earl</i> se originó a partir de una onda tropical ')
TEMA	Proceso (o entidad) que conlleva un cambio de estado (ej. 'La tormenta tropical se convirtió en huracán'), que se desarrolla a partir de otro (ej. 'El huracán se desarrolló a partir de una onda tropical') o que simplemente existe sin llevar aparejada ninguna acción (ej. 'El huracán duró dos días')
TIEMPO	Periodo o intervalo temporal en el que tiene lugar la acción (ej. 'El incendio de Cedar comenzó el 25 de octubre de 2003 ')
RESULTADO	Proceso (o entidad) causado o producido por otro, consecuencia o resultado de otro (ej. 'La tormenta tropical se transformó en huracán ')

lexicon) (Kipper, 2005), ADESSE¹⁴ (*Base de datos de verbos, alternancias de diátesis y esquemas sintáctico-semánticos del español*) (Vaamonde, González, & García, 2010) y Sensem¹⁵ (*Sentence semantics*) (Fernández & Vázquez, 2012). Por último, se han creado y adaptado ciertos roles necesarios para el ámbito del medio ambiente.

En concreto, la Tabla 7 enumera el inventario de roles semánticos utilizados para marcar los argumentos del dominio EVENTO EXTREMO (Buendía Castro, 2013, p. 380) y, en consecuencia, los de la categoría DESASTRE NATURAL. Para cada rol, se ofrece una definición y un ejemplo del corpus en el que se ha marcado, en negrita, la realización lingüística del argumento que se corresponde con el rol.

No obstante, una misma realización lingüística de una categoría no siempre activa el mismo rol semántico, ya que el significado contextual puede hacer que este último varíe. Así pues, para reforzar esta distinción, se proporciona también una categorización más general de cada argumento de acuerdo con su macrofunción. En consonancia con la Gramática del Papel y la Referencia (Van Valin & LaPolla, 1997; Van Valin, 2005), existen dos *macro-roles*: el de ACTOR que corresponde a la función generalizada de tipo AGENTE y el de PADECEDOR (“UNDERGOER” en inglés) que se asigna a la función generalizada de tipo PACIENTE (Cortés-Rodríguez & Mairal Usón, 2004, p. 162). Por ejemplo, la asignación de roles y macro-roles de los argumentos y realizaciones lingüísticas de los ejemplos de la Tabla 6 se recogen en la Tabla 8. Como primer argumento del predicado SOFOCAR aparece bien un AGENTE ACUÁTICO (ej. ‘lluvia’) que activa el rol semántico de AGENTE NATURAL, bien un SER HUMANO (ej. ‘bombero’) que activa el rol semántico de AGENTE HUMANO. En ambos casos, activan el macro-rol de ACTOR. El segundo argumento del predicado es un DESASTRE INCENDIARIO (ej. ‘fuego’) con el rol de PACIENTE INANIMADO y el macro-rol de PADECEDOR. Además, de acuerdo con los ejemplos de la Tabla 6, el segundo argumento de SOFOCAR aparece siempre de forma explícita, mientras que el primer argumento algunas veces aparece implícito (ej. ‘Sofocan el incendio en la refinería más importante de Venezuela’). Por último, los roles semánticos de LUGAR, TIEMPO, MANERA y MEDIDA pueden incluirse ocasionalmente (ej. ‘Cuatro unidades de bomberos sofocaron el fuego rápidamente en Navarra’).

Como ejemplo de una misma realización lingüística con distintos roles semánticos, se puede observar el caso del argumento ‘huracán’ que, dependiendo del colocado verbal, activa el rol semántico de TEMA (proceso que meramente existe) en ‘El huracán duró dos días’ o el rol de AGENTE NATURAL (proceso que produce un daño) en ‘El huracán azotó Nueva York’. Así pues, los macro-roles de

14. Disponible en <http://adesse.webs.uvigo.es/datos.html>

15. Disponible en <http://grial.uab.es/sensem/corpus?idioma=es>

Tabla 8. Roles semánticos y macro-roles de argumentos de SOFOCAR

Rol semántico	AGENTE NATURAL/ AGENTE HUMANO	PACIENTE INANIMADO	MEDIDA	LUGAR	TIEMPO	MANERA
Macro-rol	ACTOR	PADECEDOR	_____	_____	_____	_____
Categorías semánticas	AGENTE ACUÁTICO, SER HUMANO	DESASTRE INCENDIARIO	MEDIDA	CONSTRUCCIÓN, ÁREA	TIEMPO	MANERA
Realizaciones lingüísticas	lluvias, efectivos, unidades de bomberos	incendio, fuego	grandes dimensiones	refinería Navarra Venezuela	12,15h	rápidamente

esta proyección varían. En el primer caso, funciona como PADECEDOR puesto que experimenta un proceso y en el segundo como ACTOR.

4.3 Estructura morfosintáctica

Las distintas realizaciones lingüísticas de los argumentos de la categoría DESASTRE NATURAL se analizaron también en términos de su estructura morfosintáctica, al igual que las proyecciones léxicas de otras categorías con las que se relacionan. Se distingue entre sintagma nominal, sintagma adjetival, sintagma adverbial y sintagma preposicional, como se puede observar en la Tabla 9.

Tabla 9. Estructuras morfosintácticas presentes en *EcoLexicon*

Categorías semánticas	AGENTE ACUÁTICO, SER HUMANO	DESASTRE INCENDIARIO	MEDIDA	CONSTRUCCIÓN, ÁREA	TIEMPO	MANERA
Realizaciones lingüísticas	lluvias, efectivos, unidades de bomberos	incendio, fuego	grandes dimensiones	refinería Navarra, Venezuela	12,15h	rápidamente
Sintagma	SN	SN	SAdj (de)	SP (en)	SP (a)	SAdv

Para el caso de los sintagmas preposicionales y adjetivales asignados a las categorías de MEDIDA, LUGAR, TIEMPO y MANERA, el objetivo es identificar el conjunto de preposiciones que generalmente aparecen de forma conjunta con el núcleo. Por ejemplo, para la categoría de LUGAR, en español suele utilizarse la preposición ‘en’ (ej. ‘en la refinería’) y, para la categoría TIEMPO, la preposición ‘a’ (ej. ‘a las 12.15h’).

5. Las plantillas fraseológicas para la codificación y adquisición de conocimiento sobre los desastres naturales

El enfoque teórico y la metodología aplicada por la Terminología basada en Marcos (Faber, 2009, 2011, 2012) para la elaboración de recursos terminográficos como *EcoLexicon* se basan en la premisa de que la comunicación no es posible sin cognición (León Araúz, Faber, & Pérez Hernández, 2008). Así pues, las situaciones comunicativas en las que es necesario codificar conocimiento (general o experto) para poder transmitirlo con la intención deseada presuponen unos conocimientos previos sobre la materia tratada y sobre los usos lingüísticos más adecuados para su transmisión. Desde esta perspectiva, el trabajo del lexicógrafo y del terminógrafo ha de tener en cuenta el perfil de los usuarios de los recursos y adecuar la información que se les ofrece a sus necesidades y conocimientos previos (Tarp, 2008; García Aragón, Buendía Castro, & López Rodríguez, 2014).

5.1 La adquisición y codificación de conocimiento experto

En *EcoLexicon* se asume que los usuarios están familiarizados, en mayor o menor grado, con el lenguaje científico y con su uso, al menos en inglés o en español, ya que son las lenguas de la interfaz (López Rodríguez, Buendía Castro, & García Aragón, 2012, p. 62).

Por tanto, los principales usuarios de este recurso son: (i) estudiantes de ciencias que desean adquirir conocimiento experto; (ii) traductores que buscan la adquisición de conocimiento especializado que les permita comprender o descodificar el texto original o que buscan información relativa a la correspondencia entre lenguas, uso de términos y su combinatoria para ayudarles a codificar el conocimiento en la lengua de llegada y producir una traducción; (iii) expertos interesados en la producción de textos especializados o semiespecializados en una lengua extranjera. A su vez, estos tres grupos se corresponden, por lo general, con perfiles distintos desde el punto de vista del conocimiento experto que poseen *a priori*: usuarios legos (con unos conocimientos básicos sobre el ámbito del medio ambiente), usuarios semiexpertos (con unos conocimientos previos que tienden a especializarse a medida que adquieren conocimientos con mayor grado de especificidad y precisión para el desempeño de sus tareas profesionales) y usuarios expertos (con unos conocimientos especializados o muy especializados en el ámbito de las ciencias del medio ambiente) (Montero Martínez et al., 2011, pp. 24–28).

De acuerdo con la Teoría de las Funciones Lexicográficas (TFL) (Bergenholtz & Tarp, 2003, 2010; Tarp, 2008), las necesidades que presentan estos tres grupos de usuarios se derivan de las llamadas situaciones cognitivas, comunicativas

y operacionales. En el caso de *EcoLexicon*, se atiende a las dos primeras,¹⁶ ya que conocimiento y comunicación son las dos caras de una misma moneda. Concretamente, las *situaciones cognitivas* vienen motivadas por la necesidad de aumentar el conocimiento enciclopédico, cultural o de adquisición de lenguaje especializado (Tarp, 2007; L'Homme & Leroyer, 2009, p. 269). Respecto a las *situaciones comunicativas*, que muchas veces se dan conjuntamente con las situaciones cognitivas, estas incluyen todas aquellas necesidades relacionadas con un problema lingüístico, tanto en la lengua materna como en una lengua extranjera, que puede subsanarse mediante el uso de un recurso lexicográfico o terminográfico (Bergenholtz & Tarp, 2010, p. 31; L'Homme & Leroyer, 2009, p. 270). Así pues, en *EcoLexicon* y por ende en este estudio a través del modelo de clasificación y representación de las colocaciones verbales, se pretende dar respuesta a las siguientes necesidades: adquisición de conocimiento léxico y conceptual sobre temas relacionados con el medio ambiente que permita codificar y descodificar mensajes en el proceso de recepción, traducción, revisión, corrección y producción de textos, bien en lengua materna o en una lengua extranjera del usuario.

5.2 La plantilla del subdominio léxico 'causar que cese de existir (un desastre incendiario)' y su patrón fraseológico

Para llegar a las descripciones fraseológicas que se ofrecerán al usuario final de *EcoLexicon*, en este caso de las colocaciones verbales, la TBM aplica nociones basadas en la estructuración del conocimiento conceptual y léxico (eventos, categorías conceptuales, (sub)dominios léxicos, roles y macro-roles semánticos), tal y como se ha ilustrado. Las clasificaciones semánticas resultantes se complementan y enriquecen con información lingüística propia de cada una de las lenguas presentes en la base de conocimiento. De esta forma, en el análisis de las construcciones fraseológicas se asegura la necesaria descripción del conocimiento conceptual para llegar a la descripción del conocimiento léxico y facilitar el proceso de codificación y descodificación por parte del usuario final. En realidad, se trata de integrar toda la información recopilada en las estructuras de conocimiento ontológico que subyacen en *EcoLexicon* en forma de proposiciones conceptuales o *triples* ontológicos de distinto tipo (León Araúz, Gómez-Romero, & Bobillo, 2012; Montero Martínez, 2003, 2008).

16. Las *situaciones operacionales* se refieren a aquellas en las que “los futuros usuarios precisan de instrucciones para realizar algún tipo de operación física o intelectual” (Fuertes Olivera & Tarp, 2008, p. 79). Hasta la fecha, no parece existir ningún diccionario enfocado a asistir al usuario en dichas situaciones, no obstante se está trabajando en esta línea que podría resultar muy prometedora (Tarp, 2010, p. 50).

Así pues, una vez descritos, caso por caso, todos los verbos del subdominio ‘causar que cese de existir (un desastre incendiario)’, como el predicado SOFOCAR,¹⁷ se establece una plantilla fraseológica del subdominio en cuestión (Tabla 10). Se trata de abstraer toda la información verbal y argumental representándola en forma de requisitos y restricciones, de acuerdo con las categorías semánticas, roles, macro-roles y estructura morfosintáctica. De esta forma, es posible establecer generalizaciones fraseológicas para cada subdominio léxico que se materializan a través de meta-construcciones fraseológicas denominadas *patrones fraseológicos* (Montero Martínez, 2008, p. 1019). Se trata, por tanto, de triples ontológicos en donde se refleja una combinatoria de elementos conceptuales determinada que lleva asociada una semántica concreta. Así pues, estos esquemas constituyen el prototipo para todas aquellas meta-construcciones subordinadas que heredan sus características, si bien pueden añadir rasgos diferenciales como puede ser la aparición de otros elementos conceptuales más periféricos.

Por ejemplo, según la plantilla mostrada en la Tabla 10, cuando un predicado tenga como arg¹ un AGENTE ACUÁTICO o SER HUMANO (que active el rol semántico de AGENTE NATURAL o AGENTE HUMANO y el macro-rol de ACTOR) y como arg² un DESASTRE INCENDIARIO (que funcione como PACIENTE INANIMADO y PADECEDOR) se activará previsiblemente el significado subyacente ‘causar que cese de existir (un desastre incendiario)’. Concretamente, los predicados verbales en español ‘APAGAR’, ‘EXTINGUIR’ y ‘SOFOCAR’, cumplen con estas condiciones. Por tanto, todas estas combinaciones responden a un patrón conceptual superordinado, el patrón fraseológico [AGENTE NATURAL] o [AGENTE HUMANO] *causa que un* [DESASTRE INCENDIARIO] *cese de existir*. A su vez, este patrón combinatorio tiene lugar en el dominio léxico de EXISTENCIA y en el subdominio ‘causar que cese de existir (un desastre incendiario)’. Además, existen ciertas características no tan centrales que pueden activarse en determinados casos como puede ser la aparición de categorías semánticas como LUGAR, TIEMPO, MANERA y MEDIDA. Esta información aparece bajo el epígrafe *Nota*, al igual que información que sirve para especificar características como que normalmente el AGENTE NATURAL es un AGENTE ACUÁTICO (ej. ‘lluvia’).

No obstante, es preciso mencionar que, a nivel de uso lingüístico, no todos los verbos con el mismo número de argumentos y con el mismo tipo de categorías semánticas presentan el mismo comportamiento. Algunas veces hay variaciones en el registro y otras los predicados imponen ciertas restricciones de uso a sus argumentos. Por tanto, en el caso de que haya diferencias, estas se incluyen en el apartado de notas de la plantilla fraseológica de cada verbo en particular.

17. En el anexo 2 se puede observar la plantilla fraseológica completa del predicado SOFOCAR.

Tabla 10. Plantilla fraseológica del subdominio léxico ‘causar que cese de existir (un desastre incendiario)’

Dominio léxico: EXISTENCIA			
Subdominio léxico: causar que cese de existir (un desastre incendiario)			
Patrón fraseológico: [AGENTE NATURAL] o [AGENTE HUMANO] causa que un [DESASTRE INCENDIARIO] cese de existir.			
Nota: El AGENTE NATURAL es normalmente un AGENTE ACUÁTICO (ej. ‘lluvia’). Adicionalmente pueden añadirse las categorías semánticas LUGAR, TIEMPO, MANERA Y MEDIDA.			
Rol semántico	AGENTE NATURAL / AGENTE HUMANO	<i>apagar extinguir</i> <i>sofocar</i>	PACIENTE INANIMADO
Macro-rol	ACTOR		PADECEDOR
Categoría semántica	AGENTE ACUÁTICO/ SER HUMANO		DESASTRE INCENDIARIO
Sintagma	Sintagma nominal		Sintagma nominal

En definitiva, se trata de hacer converger aspectos cognitivos con aspectos de uso lingüístico en la descripción de las construcciones fraseológicas, de tal manera que la representación de esta información en *EcoLexicon* pueda integrarse en las estructuras de conocimiento propias del ámbito del medio ambiente; es decir, en el EVENTO MEDIOAMBIENTAL que subyace, para así garantizar la adquisición de conocimiento por parte del usuario y facilitar sus necesidades de codificación y descodificación lingüística.

6. La representación de colocaciones verbales en *EcoLexicon*: la entrada de INCENDIO

Para facilitar a los usuarios *target* de *EcoLexicon* la tarea de adquisición de conocimiento experto y su codificación y descodificación en las distintas lenguas ofrecidas, las búsquedas pueden llevarse a cabo bien a partir de conceptos (consultas onomasiológicas), bien a partir de términos o colocados verbales (consultas semasiológicas). Así pues, colocaciones como ‘sofocar un incendio’ están accesibles a través de la entrada del término ‘incendio’ o a través de la búsqueda del verbo ‘sofocar’.

Asimismo, es posible personalizar las búsquedas restringiéndolas a ámbitos concretos dentro del área del medio ambiente (ej. geología, ingeniería de costas, derecho medioambiental, etc.) (León Araúz & Faber, 2010), una funcionalidad que permite al usuario optimizar su tiempo de consulta. Por último, los contenidos que conforman la macroestructura y la microestructura de las entradas está

destinado también a facilitar los procesos lingüísticos y cognitivos aludidos que, en muchas ocasiones, van de la mano.

6.1 Información conceptual sobre INCENDIO

Las entradas de *EcoLexicon* ofrecen gran cantidad de información conceptual. A modo de ejemplo, la Figura 4 muestra la entrada correspondiente al concepto INCENDIO. En el panel central, se muestra una red dinámica con el concepto objeto de búsqueda. Así pues, en este caso, se observa cómo el concepto INCENDIO aparece vinculado a otros conceptos con los que guarda alguna relación conceptual próxima. En concreto, en el inventario de relaciones conceptuales (17 en total) se incluyen tres tipos, según consta en la leyenda del cuadro de texto del panel central: (i) relaciones hiponímicas (genérico-específicas), por ejemplo INCENDIO_TIPO-DE_EVENTO EXTREMO; (ii) relaciones meronímicas (parte-todo), que no son pertinentes en este concepto en cuestión; y (iii) relaciones no jerárquicas, como INCENDIO_RESULTADO-DE_COMBUSTIÓN.

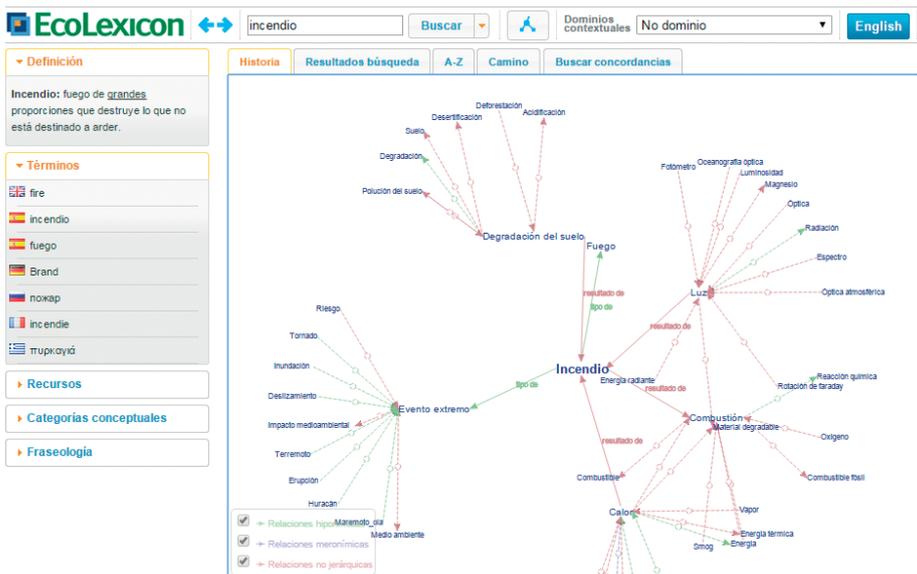


Figura 4. Entrada de INCENDIO en *EcoLexicon*

Esta información conceptual en forma de tesoro visual, que el usuario puede limitar o expandir a través de filtros, agiliza el procesamiento de la información a nivel cognitivo y se completa con una definición textual del concepto. Bajo la pestaña *Definición*, en el menú lateral izquierdo, aparece una definición de INCENDIO que codifica las relaciones conceptuales más salientes: “fuego [*tipo de*] de grandes

proporciones [*atributo_de*], que destruye lo que no está destinado a arder [*causa*]”. Asimismo, bajo la pestaña *Recursos* se incluyen imágenes, documentos, URL y material audiovisual adicional sobre el concepto. Finalmente, bajo *Categorías conceptuales*, se incluye el grupo de conceptos que integran dicha área de significado, por ejemplo, INCENDIO se conceptualiza como un PROCESO NATURAL y, finalmente, como un DESASTRE NATURAL.

De esta forma, el usuario tiene acceso a información conceptual en distintas modalidades (imágenes, redes conceptuales dinámicas, vídeos, textos, definiciones, etc.), lo cual le permite navegar hasta llegar al nivel de especificidad que precise según su grado de conocimiento inicial y la finalidad de su consulta.

6.2 Información terminológica sobre INCENDIO

La información sobre los conceptos representados en *EcoLexicon* incluye también los términos que designan cada concepto en seis lenguas; a saber, inglés, francés, español, alemán, griego moderno y ruso. En la Figura 4, en el menú de la izquierda, aparecen las distintas denominaciones del concepto INCENDIO bajo la pestaña *Términos* ('fire', 'incendio', 'fuego', 'Brand', 'пожар', 'incendie', 'πυρκαγιά'). Al pinchar en cada uno de ellos, se ofrece una descripción lingüística detallada, como muestra la Figura 5 para el término español 'incendio'.

Term information	
Term:	incendio
Language:	Spanish
Term type:	main term
Context:	ev-incendio1-jose.txt
Part of speech:	common noun
Gender:	masculine

[View concordances](#)

Phraseological section	
ACTION	
to produce fire	arder2
CHANGE	
to cause sth/sb to change for the worse	afectar arder1 arrasar asolar calcinar castigar dañar demoler derribar derruir derrumbar destrozar destruir devastar quemar

[Phraseological entry](#)

Figura 5. Microestructura de la entrada terminológica 'incendio'

Además del idioma, la descripción terminológica incluye los siguientes campos: (i) el tipo de término (término principal, sinónimo, variante geográfica y acrónimo); (ii) un archivo con una selección de contextos de uso; (iii) información morfosintáctica acerca de su categoría gramatical (sustantivo, verbo, adjetivo y adverbio); (iv) el género gramatical para las lenguas que lo requieran (masculino, femenino y neutro); (v) acceso al módulo de búsqueda de concordancias en el corpus integrado; (vi) una sección fraseológica breve y (vii) acceso a una entrada fraseológica más completa relativa al término.

6.3 Información fraseológica sobre INCENDIO

Bajo la *Sección fraseológica*, aparece información relativa a los predicados verbales que coaparecen con el término dando lugar a colocaciones verbales. En la Figura 6, los predicados reflejan la clasificación semántica en dominios léxicos (significado nuclear) y en subdominios (dimensión de significado concreta). Así pues, cuando el término español *incendio* activa el significado de existencia (EXISTENCIA) y, concretamente, la dimensión ‘causar que cese de existir (un desastre incendiario)’, los verbos que se activan en español son ‘apagar’, ‘extinguir’ y ‘sofocar’.

EXISTENCE	
to begin to exist	surgir
to cause somebody to cease to exist	matar
to cause to cease to exist (of fire disaster)	apagar extinguir sofocar
to cause to exist/happen	causar generar ocasionar originar producir provocar
to continue to exist (of natural disaster)	durar extenderse perdurar persistir
to continue to exist (of people)	recuperarse de sobrevivir a
to exist in time	ocurrir pasar producirse sobreenir suceder tener lugar
to begin to exist	comenzar desarrollarse empezar formarse originarse

Figura 6. Información semántica sobre los verbos de las colocaciones

Además, a través de una ventana emergente, cada uno de los verbos ofrece ejemplos de uso y una sección de notas con información relativa a las restricciones de uso o significado. Al pinchar en ‘sofocar’ (Figura 7), el usuario tiene acceso a cinco ejemplos de uso, así como a un apartado de notas que establece que el AGENTE NATURAL es normalmente un AGENTE ACUÁTICO (ej. ‘lluvia’) y que el AGENTE es normalmente un SER HUMANO. Asimismo, se especifica que ‘sofocar’, también puede combinarse con complementos de lugar (LUGAR), tiempo (TIEMPO) o modo (MANERA).

Verb details	
Verb	sofocar
Usage examples	<ol style="list-style-type: none"> 1. Al lugar de los hechos se desplazaron efectivos de los Bomberos, que sofocaron el incendio sobre las 12,15 horas. 2. Sofocan el incendio en la refinería más importante de Venezuela. 3. Cuatro unidades de bomberos sofocaron el fuego rápidamente en Navarra. 4. Las lluvias consiguieron sofocar el fuego de grandes dimensiones.
Note	The NATURAL AGENT is usually a WATER AGENT (e.g. 'lluvia'). LOCATION, TIME, MANNER and MEASURE can also be specified.

Figura 7. Ejemplos de uso y restricciones semánticas del verbo *sofocar*

Además de la sección fraseológica, el usuario tiene una *Entrada fraseológica* más completa. La Figura 8 muestra un extracto de la entrada del término 'incendio', en la que se incluyen los verbos 'apagar', 'extinguir' y 'sofocar' junto con su patrón fraseológico, su dimensión semántica y su significado nuclear. Por último, al igual que ocurría con la sección fraseológica abreviada, el usuario puede pinchar en cada verbo para obtener ejemplos y notas de uso.

Phraseology	
Nuclear meaning	EXISTENCE
Meaning dimension	to continue to exist (of people)
Phraseological pattern	EXPERIENCER continues to exist in its former state after a NATURAL DISASTER
Verbs	sobrevivir a , recuperarse de
<hr/>	
Nuclear meaning	EXISTENCE
Meaning dimension	to cause somebody to cease to exist
Phraseological pattern	NATURAL DISASTER causes HUMAN BEING to cease to exist
Verbs	matar
<hr/>	
Nuclear meaning	EXISTENCE
Meaning dimension	to cause to cease to exist (of fire disaster)
Phraseological pattern	WATER AGENT or HUMAN BEING causes a NATURAL FIRE DISASTER to cease to exist
Verbs	apagar , extinguir , sofocar

Figura 8. Extracto de la entrada fraseológica del término 'incendio'

7. Conclusiones

Las necesidades de codificación y descodificación de conocimiento sobre el medio ambiente que se atienden en la aplicación *EcoLexicon* obligan a formular una definición del fenómeno de las colocaciones verbales acorde con la estructura ontológica, basada en el conocimiento, de dicho recurso. Así pues, se ilustra una metodología de tipo *top-down* y *botton-up* que permite dar cuenta de las características más sobresalientes de este fenómeno en un dominio de especialidad: i) que son unidades cuya semántica responde a un patrón conceptual propio del ámbito, ii) que son combinaciones frecuentes, iii) que están formadas por dos o más unidades, iv) que presentan una estructura de tipo sustantivo + verbo o verbo + sustantivo.

Para ello, es necesario analizar tanto el comportamiento de los argumentos nominales como de los predicados verbales de las colocaciones. A partir de la clasificación de los predicados en dominios y subdominios léxicos presentes en el ámbito del medio ambiente, sus argumentos se clasifican y estructuran en función de las categorías semántico-conceptuales propias del ámbito, junto con los roles semánticos y macro-roles que activan. De esta forma, es posible establecer generalizaciones en forma de patrones fraseológicos que permiten codificar la información semántico-conceptual pertinente para la descripción de un conjunto de unidades relacionadas con una misma combinación conceptual. Así pues, es posible predecir el abanico de verbos generalmente asociados a un evento especializado concreto.

Este método de codificación de la información, enriquecida también con anotaciones sobre restricciones y requisitos de uso lingüístico en cada lengua particular, resulta de gran utilidad para la adquisición de conocimiento sobre el medio ambiente y su lenguaje, así como para la producción y recepción de textos especializados en este dominio. El usuario de *EcoLexicon* tiene acceso a las colocaciones de cada término de la base de datos de forma amigable y se permiten búsquedas de tipo semasiológico y onomasiológico para atender a sus necesidades de consulta.

Para completar la información fraseológica contenida en *EcoLexicon*, actualmente se están codificando construcciones fraseológicas de tipo nominal, adjetival y adverbial en torno a patrones fraseológicos, ya que estos son la estructura paraguas que posibilita la vinculación de la información léxica y conceptual con los triples ontológicos que conforman la base de conocimiento *EcoLexicon*. El objetivo final es alcanzar la idea de diccionario descrita por Sinclair (2004, p. xxiv): "A dictionary containing all the lexical items of a language, each one in its canonical form with a list of possible variations, would be the ultimate dictionary". Por último, este proceso de estructuración facilitará la integración de la base de

conocimiento de *EcoLexicon* en la web de datos, consiguiendo enlazarse y estar enlazada con otros recursos terminológicos abiertos (*Linked Linguistic Data*).

Agradecimientos

Esta investigación se ha llevado a cabo en el marco del proyecto CONTENT: *Bases cognitivas y neurológicas de la terminología asistida para la traducción/ Cognitive and Neurological Bases for Terminology-enhanced Translation* (FF2014-52740-P), financiado por el Ministerio de Economía y Competitividad.

Referencias

- Baroni, M., & Bernardini, S. (Eds.). (2006). *Wacky!: Working papers on the web as corpus*. Bolonia: GEDIT.
- Bergenholtz, H., & Tarp, S. (2003). Two opposing theories: On H. E. Wiegand's recent discovery of lexicographic functions. *Hermes: Journal of Linguistics*, 31, 171–196.
- Bergenholtz, H., & Tarp, S. (2010). LSP lexicography or terminography?: The lexicographer's point of view. En P. A. Fuertes-Olivera (Ed.), *Specialized dictionaries for learners* (pp. 27–37). Berlín/Nueva York: Mouton de Gruyter.
- Benson, M., Benson, E., & Ilson, R. (1986). *Lexicographic description of English*. Amsterdam/Filadelfia: John Benjamins. doi: 10.1075/slcs.14
- Benson, M., Benson, E., & Ilson, R. (2009). *The BBI combinatory dictionary of English* (3ª edición). Amsterdam/Filadelfia: John Benjamins.
- Buendía Castro, M. (2012). Verb dynamics. *Terminology*, 18(2), 149–166. doi: 10.1075/term.18.2.01bue
- Buendía Castro, M. (2013). *Phraseology in specialized language and its representation in environmental knowledge resources*. Tesis doctoral, Universidad de Granada
- Buendía Castro, M., & Ureña Gómez-Moreno, J. M. (2009). Parameters of evaluation for corpus design. *International Journal of Translation*, 21(1–2), 73–88.
- Corpas Pastor, G. (2003). *Diez años de investigación en fraseología: Análisis sintáctico semánticos, contrastivos y traductológicos*. Fráncfort/Madrid: Vervuert/Ibero-Americana.
- Corpas Pastor, G. (2008). *Investigar con corpus en traducción: Los retos de un nuevo paradigma*. Fráncfort: Peter Lang.
- Cortés-Rodríguez, F., & Mairal Usón, R. (2004). La formación de palabras, de nuevo, como encrucijada: La interfaz semántica-morfología-sintaxis-fonología en el Modelo de Gramáticas Léxicas. En G. Wotjak & J. Cuartero Otal (Eds.), *Algunos problemas específicos de la descripción sintáctico-semántica* (pp. 159–174). Berlín: Frank & Timme Verlag.
- Cowie, A. P. (2012). Dictionaries, language learning and phraseology. *International Journal of Lexicography*, 25(4), 386–392. doi: 10.1093/ijl/ecso23
- De Schryver, G. M. (2002). Web for/as corpus: A perspective for the African languages. *Nordic Journal of African Studies*, 11, 266–282. Recuperado de <http://tshwanedje.com/publications/webtocorpus.pdf>

- Drouin, P. (2003). Term extraction using non-technical corpora as a point of leverage. *Terminology*, 9(1), 99–117. doi: 10.1075/term.9.1.o6dro
- Faber, P. (2009). The cognitive shift in terminology and specialized translation. *MonTI: Monografías de Traducción e Interpretación*, 1(1), 107–134. doi: 10.6035/MonTI.2009.1.5
- Faber, P. (2011). The dynamics of specialized knowledge representation: Simulational reconstruction or the perception-action interface. *Terminology*, 17(1), 9–29. doi: 10.1075/term.17.1.o2fab
- Faber, P. (Ed.). (2012). *A cognitive linguistics view of terminology and specialized language*. Berlin/Boston: Mouton de Gruyter. doi: 10.1515/9783110277203
- Faber, P., León Araúz, P., & Prieto Velasco, J. A. (2009). Semantic relations, dynamicity, and terminological knowledge bases. *Current Issues in Language Studies*, 1, 1–23.
- Faber, P., & Mairal Usón, R. (1999). *Constructing a lexicon of English verbs*. Berlín/Nueva York: Mouton de Gruyter. doi: 10.1515/9783110800623
- Fernández, A., & Vázquez, G. (2012). Análisis cuantitativo del corpus SenSem. En I. Elorza, O. Carbonell i Cortés, R. Albarrán, B. García Rianza & M. Pérez-Veneros (Eds.), *Empiricism and analytical tools for 21st century Applied Linguistics* (pp. 157–170). Salamanca: Ediciones de la Universidad de Salamanca.
- Fillmore, C. J., & Atkins, B. T. (1992). Toward a frame-based lexicon: The semantics of RISK and its neighbors. En A. Lehrer & E. F. Kittay (Eds.), *Frames, fields and contrasts: New essays in semantic and lexical organization* (pp. 75–102). Hillsdale: Erlbaum.
- Firth, J. R. (1968). *Selected papers of J. R. Firth 1952–1959* (editado por F. R. Palmer). Londres: Longman.
- Fletcher, W. H. (2007). Implementing a BNC-compare-able web corpus. En C. Fairon, H. Naets, A. Kilgarriff, & G. M. De Schryver (Eds.), *Building and exploring web corpora* (pp. 43–56). Lovaina la Nueva: Cahiers du Central.
- Fuertes Olivera, P. A., & Tarp, S. (2008). La teoría funcional de la lexicografía y sus consecuencias para los diccionarios de economía del español. *Revista de Lexicografía*, 14, 75–95.
- García Aragón, A., Buendía Castro, M., & López Rodríguez, C. I. (2014). Evaluación de una base de conocimiento terminológica sobre el medio ambiente en el aula de la traducción especializada. En C. Vargas Sierra (Ed.), *TIC, trabajo colaborativo e interacción en terminología y traducción* (pp. 477–487). Granada: Comares.
- García-Page, M. (2008). *Introducción a la fraseología española: Estudio de las locuciones*. Barcelona: Anthropos.
- Gläser, R. (1994/1995). Relations between phraseology and terminology with special reference to English. *ALFA*, 7(8), 41–60.
- Gläser, R. (1998). The stylistic potential of phraseological units. En A. P. Cowie (Ed.), *Phraseology: Theory, analysis, and applications* (pp. 125–143). Oxford: Clarendon Press.
- Goldberg, A. (1995). *Constructions: A Construction Grammar approach to argument structure*. Chicago: University of Chicago Press.
- Goldberg, A. (2006). *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.
- Grinev, S., & Klepalchenko, I. (1999). Terminological approach to knowledge representation. En *TKE '99: Proceedings of the 5th International Congress on Terminology and Knowledge Engineering* (pp. 147–151). Viena: TermNet.
- Gross, G. (1994). Classes d'objets et description des verbes. *Langages*, 115, 15–30. doi: 10.3406/lgge.1994.1684

- Hausmann, F. J. (1989). Le dictionnaire de collocations. En F. J. Hausmann, O. Reichmann, H. E. Wiegand, & L. Zgusta (Eds.), *Wörterbücher/Dictionaries/Dictionnaires: Ein internationales Handbuch zur Lexikographie/An international encyclopedia of lexicography / Encyclopédie internationale de lexicographie* (pp. 1010–1019). Berlin/Nueva York: Walter de Gruyter.
- Heid, U. (1992). Décrire les collocations. *Terminologie et Traduction*, 2(3), 523–548.
- Heid, U. (1994). On ways words work together: Topics in lexical combinatorics. En W. Martin, W. Meys, M. Moerland, E. ten Pas, P. van Sterkenburg, & P. Vossen (Eds.), *Euralex 1994: Proceedings* (pp. 226–257). Amsterdam: International Congress on Lexicography.
- Heid, U. (2001). Collocations in sublanguage texts: Extraction from corpora. En S. E. Wright (Ed.), *Handbook of terminology management: Applications-oriented terminology management*, Vol. II. (pp. 788–808). Amsterdam/Filadelfia: John Benjamins. doi: 10.1075/z.htm2.31hei
- Keller, E. A., & Blodget, R. H. (2007). *Riesgos naturales: Procesos de la tierra como riesgos, desastres y catástrofes*. Madrid: Pearson Educación.
- Kipper, K. (2005). *VerbNet: A broad-coverage, comprehensive verb lexicon*. Tesis doctoral, University of Pennsylvania. Recuperado de <http://verbs.colorado.edu/~kipper/Papers/dissertation.pdf>
- León Araúz, P., Faber, P., & Pérez Hernández, C. (2008). LSP dictionaries and their genuine purpose: A frame-based example from MARCOCOSTA. En E. Bernal & J. DeCesaris (Eds.), *Proceedings of the XIII EURALEX International Congress*. Barcelona: Institut Universitari de Lingüística Aplicada (Universitat Pompeu Fabra). Recuperado de <http://lexicon.ugr.es/pdf/leonfaberperez2008.pdf>
- León Araúz, P., & Faber, P. (2010). Natural and contextual constraints for domain-specific relations. En V. Barbu Mititelu, V. Pekar, & E. Barbu (Eds.), *Proceedings of the Workshop Semantic Relations. Theory and Applications*, (pp. 12–17). La Valeta.
- León Araúz, P., Gómez-Romero, J., & Bobillo, F. (2012). A fuzzy ontology extension of WordNet and EuroWordnet for specialized knowledge. En G. Aguado de Cea, M. C. Suárez-Figueroa, R. García-Castro, & E. Montiel-Ponsoda (Eds.), *Proceedings of the 10th Terminology and Knowledge Engineering Conference (TKE-2012): New frontiers in the constructive symbiosis of terminology and knowledge engineering* (pp. 139–154). Madrid.
- López Rodríguez, C. I., Buendía Castro, M., & García Aragón, A. (2012). User needs to the test: Evaluating a terminological knowledge base on the environment by trainee translators. *Jostrans: The Journal of Specialized Translation*, 18, 57–76. Recuperado de http://www.jostrans.org/issue18/art_lopez.pdf.
- L'Homme, M. C. (1998). Le statut du verbe en langue de spécialité et sa description lexicographique. *Cahiers de Lexicographie*, 73(2), 61–84. Recuperado de <http://www.ling.umontreal.ca/lhomme/docs/cahiers-lexico-98.PDF>
- L'Homme, M. C., & Leroyer, P. (2009). Combining the semantics of collocations with situation-driven search paths in specialized dictionaries. *Terminology*, 15(2), 258–283. doi: 10.1075/term.15.2.05lho
- Mel'čuk, I. et al. (1984–1999). *Dictionnaire explicatif et combinatoire du français contemporain: Recherche lexico-sémantiques I, II, III, IV*. Montreal: Les Presses de l'Université de Montréal.
- Montero Martínez, S. (2003). *Estructuración conceptual y formalización terminográfica de frasesmas en el subdominio de la oncología*. Madrid: CSIC/Elies. Recuperado de <http://elies.rediris.es/elies19>
- Montero Martínez, S. (2008). A constructional approach to terminological phrasemes. En E. Bernal & J. DeCesaris (Eds.), *Proceedings of the XIII International EURALEX Congress* (pp. 1015–1022). Barcelona: IULA.

- Montero Martínez, S., & Buendía Castro, M. (2010). Las construcciones fraseológicas en la Terminología basada en Marcos. Plenary session presented at the *Seminario de Lingüística Española: Los retos informáticos del español en Lexicología, Terminología y Fraseología*, Universidad de Amberes, 26–27 de febrero.
- Montero Martínez, S., & Buendía Castro, M. (2012). La sistematización en el tratamiento de las construcciones fraseológicas: el caso del medio ambiente. En I. Elorza, O. Carbonell i Cortés, R. Albarrán, B. García Riaza, & M. Pérez-Veneros (Eds.), *Empiricism and analytical tools for 21st century Applied Linguistics* (pp. 711–724). Salamanca: Ediciones de la Universidad de Salamanca.
- Montero Martínez, S., Faber, P., & Buendía Castro, M. (2011). *Terminología para traductores e intérpretes*. Granada: Ediciones Tragacanto.
- Montero Martínez, S., Fuertes Olivera, P. A., & García de Quesada, M. (2001). The translator as “language planner”: Syntactic calquing in an English-Spanish translation of Chemical Engineering. *META*, 46, 687–98. doi: 10.7202/003591ar
- Moon, R. (2008). Sinclair, phraseology, and lexicography. *International Journal of Lexicography*, 21(3), 243–254. doi: 10.1093/ijl/ecn027
- Pawley, A. (2001). Phraseology: Linguistics and the dictionary. *International Journal of Lexicography*, 14(2), 122–134. doi: 10.1093/ijl/14.2.122
- Roberts, R. P. (1998). Phraseology and translation. En P. Fernández Nistal & J. M. Bravo Gozalo (Eds.), *La traducción: Orientaciones lingüísticas y culturales* (pp. 61–78). Universidad de Valladolid: SAE.
- Rossenbeck, K. (1989). Lexikologische und lexicographische Probleme fachsprachlicher Phraseologie aus kontrastiver Sicht. En M. Snell-Hornby & E. Pöhl (Eds.), *Translation and lexicography* (pp. 197–210). Amsterdam/Filadelfia: John Benjamins. doi: 10.1075/z.40.23ros
- Ruppenhofer, J., Ellsworth, M., Petruck, M., Johnson, C., & Scheffczyk, J. (2010). FrameNet II: Extended theory and practice. ICSI Technical Report. Recuperado de <http://framenet2.icsi.berkeley.edu/docs/r1.5/book.pdf>
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Sinclair, J. (2004). In praise of the dictionary. En G. Williams & S. Vessier (Eds.), *Proceedings of the Eleventh EURALEX International Congress, EURALEX 2004*. (pp. 1–11). Lorient: Université de Bretagne-Sud.
- Sinclair, J. (2007). Preface. *International Journal of Corpus Linguistics*, 12(2), 155–157. doi: 10.1075/ijcl.12.2.03sin
- Tarp, S. (2007). ¿Qué requisitos debe cumplir un diccionario de traducción del siglo XXI? En P. A. Fuertes Olivera (Ed.), *Problemas lingüísticos en la traducción especializada* (pp. 227–256). Valladolid: Universidad de Valladolid.
- Tarp, S. (2008). *Lexicography in the borderland between knowledge and non-knowledge: General lexicographical theory with particular focus on learner’s lexicography*. Tübinga: Max Niemeyer Verlag. doi: 10.1515/9783484970434
- Tarp, S. (2010). Functions of specialized learners’ dictionaries. En P. A. Fuertes-Olivera (Ed.), *Specialized dictionaries for learners* (pp. 39–54). Berlín/Nueva York: De Gruyter.
- Vaamonde, G., González, F., & García, J. M. (2010). ADESSE: A database with syntactic and semantic annotation of a corpus of Spanish. En N. Calzolari, K. Choukri, B. Maegaard, J. Mariani, J. Odiijk, S. Piperidis, M. Rosner, & D. Tapias (Eds.), *Proceedings of the Seventh Conference on International Language Resources and Evaluation (LREC’10)* (pp. 1903–1910). La Valeta: European Language Resources Association (ELRA).

- Van Valin, R. D. Jr. (2005). *The syntax-semantics-pragmatics interface: An introduction to Role and Reference Grammar*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511610578
- Van Valin, R. D. Jr., & LaPolla, R. (1997). *Syntax: Structure, meaning and function*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9781139166799
- Williams, G. (2013). Review of Seretan, V. (2011). *Syntax-based collocation extraction*. *International Journal of Lexicography*, 26(1), 90–94. doi: 10.1093/ijl/ecs015
- Zuluaga, A. (1997). Sobre las funciones de unidades fraseológicas en textos literarios. *Paremia*, 6, 631–640.

Anexo 1. Relación de textos del corpus

TEXTOS DEL CORPUS

Actas de congresos

- Chacón, J., & Irigaray, C. (1996). *Riesgos naturales, ordenación del territorio y medio ambiente: Sexto Congreso Nacional y Conferencia Internacional de Geología Ambiental y Ordenación del Territorio = Natural hazards, land-use planning and environment: 6th Spanish Congress and International Conference on Environmental Geology and Land-Use Planning*. Madrid: Sociedad Española de Geología Ambiental y Ordenación del Territorio.
- Benito, G., & Díez Herrero, A. (2004). *Riesgos naturales y antrópicos en geomorfología: Libro de Actas de los simposios desarrollados durante la VIII Reunión Nacional de Geomorfología*, Toledo, 22–25 de septiembre de 2004. Madrid: Sociedad Española de Geomorfología.

Monografías

- Olcina Cantos, J. (2006a). *¿Riesgos naturales?* Vol.1. Barcelona: Davinci.
- Olcina Cantos, J. (2006b). *¿Riesgos naturales?* Vol.2. Barcelona: Davinci.
- Keller, E. A., & Blodgett, R. H. (2007). *Riesgos naturales: Procesos de la Tierra como riesgos, desastres y catástrofes*. Madrid: Pearson Educación. [Traducción de Pilar Gil Ruiz].
- Naciones Unidas. (2009). *UNISDR Terminología sobre reducción del riesgo de desastres*. Ginebra: Estrategia Internacional para la Reducción de Desastres de las Naciones Unidas. <http://www.unisdr.org/files/7817_UNISDRTerminologySpanish.pdf>.

Revistas divulgativas

National Geographic: <http://www.nationalgeographic.com.es/>

Muy Interesante: <http://www.muyinteresante.es/>

Anexo 2. Plantilla fraseológica de SOFOCAR

Dominio léxico: EXISTENCIA

Subdominio léxico: causar que cese de existir (un desastre incendiario)

Patrón fraseológico: [AGENTE ACUÁTICO] o [SER HUMANO] causa que un [DESASTRE INCENDIARIO] cese de existir.

Nota: El AGENTE NATURAL es normalmente un AGENTE ACUÁTICO (ej. 'lluvia'). Adicionalmente pueden añadirse las categorías semánticas LUGAR, TIEMPO y MANERA y MANERA.

Rol semántico	SOFOCAR	AGENTE NATURAL/ AGENTE HUMANO	PACIENTE INANIMADO	MEDIDA	LUGAR	TIEMPO	MANERA
Macro-rol		ACTOR	PADECE-DOR	_____	_____	_____	_____
Categoría semántica		AGENTE ACUÁTICO, SER HUMANO	DESASTRE INCENDIARIO	MEDIDA	CONSTRUCCIÓN, ÁREA	TIEMPO	MANERA
Realizaciones lingüísticas		lluvias, efectivos, unidades de bomberos	incendio, fuego	grandes dimensiones	refinería Navarra, Venezuela	12,15h	rápidamente
Sintagma		SN	SN	S. (de)	Adj SP (en)	SP (a)	S. Adv
Ejemplos de uso	<ol style="list-style-type: none"> 1. Al lugar de los hechos se desplazaron efectivos de los Bomberos, que sofocaron el incendio sobre las 12,15 horas. 2. Sofocan el incendio en la refinería más importante de Venezuela. 3. Cuatro unidades de bomberos sofocaron el fuego rápidamente en Navarra. 4. Las lluvias consiguieron sofocar el fuego de grandes dimensiones. 						

Abstract

Lexicographic and terminographic resources have generally focused on the description and analysis of noun + noun or noun+ adjective collocations to the exclusion of combinations of other categories of language, such as verbs, despite the fact that verbs are considered to be the most important category of language (L'Homme, 1998). This article proposes a methodology for classifying and describing verb collocations based on their meaning, with a view to enhancing the acquisition and codification of specialized knowledge. The basic premise is that when verbs are classified in lexical domains and subdomains according to the Lexical Grammar Model (Faber & Mairal Usón, 1999), and the structure and semantic characteristics of their arguments

are specified, it is possible to predict and make generalizations in regards to phraseological patterns (Montero Martínez, 2008). For this purpose, we classified the arguments of a set of verbs in a set of conceptual-semantic categories typical of a given specialized domain, along with the semantic roles activated (Van Valin & LaPolla, 1997; Van Valin, 2005), and their morphosyntactic structure. These descriptions are represented in the environmental knowledge base, *EcoLexicon*.

Keywords: collocations, verbs, phraseology, Lexical Grammar Model, Frame-based Terminology, *EcoLexicon*

Authors' addresses

Silvia Montero Martínez
Facultad de Traducción e Interpretación,
Departamento de Traducción e
Interpretación
Universidad de Granada
C/Buenuceso N°11
18071 Granada
smontero@ugr.es

Míriam Buendía Castro
Facultad de Letras, Departamento de
Filología Moderna
Universidad de Castilla-La Mancha
Avda. Camino José Cela S/N
13071 Ciudad Real
miriam.buendia@uclm.es

Biographical notes

Silvia Montero Martínez es doctora en Lingüística Española por la Universidad de Valladolid y tiene un máster en traducción especializada inglés/español por la misma universidad. En la actualidad es profesora Titular de Universidad en el Departamento de Traducción e Interpretación de la Universidad de Granada. Es autora de más de 40 artículos, libros y capítulos de libro en revistas y editoriales de prestigio como son *Terminology*, *Perspectives*, *META*, *Babel*, *Journal of Pragmatics*. Sus principales áreas de investigación son la terminología, la traducción especializada y la ingeniería del conocimiento.

Míriam Buendía Castro es doctora en Traducción e Interpretación por la Universidad de Granada. Obtuvo el premio extraordinario de doctorado de la Universidad de Granada por su tesis *Phraseology in specialized language and its representation in environmental knowledge resources*. En la actualidad es profesora contratada doctor interina en el Departamento de Filología Moderna en la Universidad de Castilla-La Mancha. Es autora de más de 30 artículos, libros y capítulos de libro en revistas y editoriales de prestigio como son *Terminology*, *Fachsprache* o *Yearbook of Phraseology*. Sus principales áreas de investigación son la terminología, la fraseología y la lingüística de corpus.