

FIRST DRAFT

*La representación léxica en la Gramática del Papel y la Referencia: Hacia un enfoque ontológico*¹

Ricardo Mairal Usón
UNED
rmairal@flog.uned.es

Carlos Perrián Pascual
Universidad Católica San Antonio
Murcia, España
jcperinan@pdi.ucam.edu

María Beatriz Pérez Cabello de Alba
UNED
bperezcabello@flog.uned.es

1. Introducción

El objetivo de este capítulo es ofrecer una propuesta alternativa al sistema de representación léxica formulado en el marco de la Gramática del Papel y la Referencia (de aquí en adelante, GPR) (cf. Van Valin, 2005; Van Valin y LaPolla, 1997). Como tesis central de este trabajo, mantenemos que la adopción de un enfoque de representación léxica inspirado en una semántica ontológica constituye una apuesta más compacta y coherente con las premisas de una teoría funcional del lenguaje que aspira a interpretar y analizar aquellos hechos que traspasan las demarcaciones del propio código lingüístico y así disponer de las herramientas metodológicas necesarias para explicar el propio código lingüístico en el marco de la comunicación y la cognición. Este giro metodológico implica la sustitución de las estructuras lógicas tal como se proponen en la GPR por lo que hemos denominado *estructuras lógico-conceptuales* (de aquí en adelante, ELC), un nuevo sistema de representación léxica que nos permite solventar muchos de los problemas que las estructuras lógicas plantean.

¹ Este trabajo forma parte del trabajo realizado en el marco del grupo de investigación LEXICOM (www.lexicom.es), que está financiado por el proyecto FFI2008-05035-C02-01/FILO.

Hemos organizado este trabajo como sigue: la sección 2 nos ofrece una revisión crítica del sistema de representación léxica propuesto en la GPR, a saber, las estructuras lógicas. Termina esta sección con una serie de desafíos que conformaron el denominador común de un conjunto de propuestas formuladas con el fin de proporcionar un sistema de descomposición semántica más refinado de los primitivos de estado y de actividad. La sección 3 recoge las líneas maestras de estas propuestas, lo que nos sirve como preámbulo para contextualizar una nueva orientación conceptual y presentar las ELCs como alternativa a las estructuras lógicas. Finalmente, cerramos este capítulo con unas conclusiones sobre las implicaciones y la proyección de las ELC en el desarrollo de aplicaciones en el ámbito del procesamiento del lenguaje natural (cf. Perrián Pascual y Mairal Usón, 2009), cuestión que retomamos con más detalle en el capítulo XXX de este volumen.

2. Las estructuras lógicas en la GPR

Recordemos que, a partir del trabajo pionero de Vendler (1967) sobre las distinciones en el modo de acción (*Aktionsart*), la GPR organiza los predicados de una lengua en clases y establece las siguientes: *estados*, *actividades*, *logros*, *semelfactivos*, *realizaciones*, *realizaciones activas* y sus correspondientes causativos (remitimos al trabajo de Cortés, González y Jiménez (este volumen) para una explicación detallada de los tests que se utilizan para distinguir una clase de otra). Cada una de estas clases está representada por una estructura lógica, tal como reproducimos en la siguiente tabla

CLASE VERBAL	ESTRUCTURA LÓGICA (EL)	EJEMPLO	REPRESENTACIÓN <i>EL</i>
Estado	predicado' (x) o (x,y)	querer	want' (x,y)
Actividad	do' (x, [predicado' (x) o (x,y)])	conducir beber	do' (x,[drive' (x)]) do' (x,[drink' (x,y)])
Logro	INGR predicado' (x) o (x,y), o INGR do' (x, [predicado' (x) o (x,y)])	explosionar estallar	INGR exploded' (x) INGR popped' (x)
Semelfactivo	SEML predicado' (x) o (x,y) SEML do' (x, [predicado' (x) o (x,y)])	vislumbrar estornudar	SEML see' (x,y) SEML do' (x,[sneeze' (x)])
Realización	BECOME predicado' (x) o (x,y), o BECOME do' (x, [predicado' (x) o (x,y)])	recibir aprender	BECOME have' (x,y) BECOME know' (x,y)
Realización activa	do' (x, [predicado ₁ ' (x, (y))] & BECOME predicado ₂ ' (z,x) o (y)	pintar, correr beber, escribir	do' (x,[drink' (x,y)]) & INGR consumed' (y)

Causativos	α CAUSE β , donde α y β son estructuras lógicas de cualquier tipo	matar, dar, tomar,	[do' (x,y)] CAUSE [BECOME [dead' (y)]]
-------------------	---	--------------------	--

Figura 1: *Inventario de la estructuras lógicas en la GPR* (adaptado de Van Valin, 2005:45)

Recordemos que una estructura lógica consta de los siguientes elementos:

- las constantes o primitivos, representados en negrita y seguidos por una comilla (') (por ej. **know'**, **want'**, **consumed'**, **drink'** etc.), forman parte de un metalenguaje semántico universal;
- los operadores, representados en mayúsculas (por ej. BECOME, INGR, CAUSE, SEML etc.);
- y las variables, marcadas como *x*, *y* o *z*, que se saturan con las entidades referidas en una expresión lingüística y son elementos propios de cada lengua².

Los estados y las actividades son primitivos y, de hecho, entran como *definiens* en la representación de los logros, semelfactivos, realizaciones y realizaciones activas. Por consiguiente, un predicado de estado como *sentir* vendrá representado por una estructura del tipo **feel'** (x,y), donde tenemos un primitivo y dos variables argumentales. Igualmente, las actividades se expresan mediante un predicado **do'**³, que expresa la marca de actividad propia de esta clase, y un primitivo, que describe la actividad específica, con sus correspondientes variables, pudiendo ser un predicado monoargumental (p.ej. **do'** (x,[**jog'** (x)])), o un predicado biargumental (p.ej. **do'** (x,[**drink'** (x,y)])). Como se indicaba arriba, el resto de las distinciones están formadas a partir de estados y actividades. Así, un logro consta de un predicado de actividad o estado más el operador INGR,⁴ que indica precisamente el rasgo de puntualidad, propio de esta clase: por ej., el predicado *estallar* designa un logro y vendrá representado por INGR **popped'** (x). También los semelfactivos se representan con el operador SEML,

² Este sistema de representación léxica es un ejemplo de lo que Levin (1995) denominó 'sistemas de descomposición léxica basados en el predicado', que, con algunas diferencias aparte (por ej. el tipo y la naturaleza de los primitivos o constantes), coincide con las estructuras conceptuales de Jackendoff (1990) o las plantillas léxicas de Rappaport y Levin (1998) y Levin y Rappaport (2005).

³ Nótese que la GPR hace una distinción entre el primitivo **do'** y **DO'** con el fin de representar la noción de agentividad prototípica en verbos como *asesinar*, *perpetrar*, *prometer* etc. que requieren que el sujeto sea necesariamente animado y dotado de volición; es decir, el evento verbal solo puede realizarse deliberadamente: **El mosquito asesinó al presidente del consejo*.

⁴ Esta abreviatura responde al inglés 'ingressive'.

que señala las propiedades de puntualidad y ausencia de un estado resultante, y pueden estar basados en actividades (por ej. *estornudar*: SEML **do'**(x, [**sneeze'** (x)])) o estados (por ej. *vislumbrar*: SEML **see'** (x,y)). Las realizaciones se representan por medio del operador BECOME, que señala un cambio de estado que tiene una transición interna: por ej. *enrojecer*: BECOME **reddened'** (x). Las realizaciones activas constan de una estructura compleja, que comprende una estructura de actividad y una de logro, que indica precisamente el carácter télico que adquiere el predicado en combinación con un sintagma nominal o adposicional referencial: por ej. el predicado *beber* en una expresión del tipo *Yasmina escribe una novela do'* (x,[**write'** (x,y)]) & INGR **exist'** (y). Esta representación incorpora el signo & que significa 'y entonces'. Finalmente, cada una de estas clases tiene su correspondiente versión causativa⁵.

En este punto, cumple elevar las siguientes consideraciones:

A) Si bien podemos admitir que el uso de palabras inglesas como primitivos potenciales parece obedecer a un criterio notacional (cf. Van Valin, 2005:45), no está tan claro qué criterios debemos utilizar para elegir los referidos primitivos. En este sentido, obtenemos como respuesta que estos primitivos tienen un correlato conceptual, lo que nos parece realmente viable. Sin embargo, si así es, ¿dónde está la ontología que da soporte al léxico? ¿cómo se ha elaborado esa ontología? ¿cómo funciona la interfaz léxico-conceptual? Además, no tenemos constancia o, al menos, no se da ninguna prueba a tal efecto de la naturaleza universal de los primitivos semánticos: ¿cómo sabemos que los referidos primitivos son tipológicamente válidos?

B) Si consideramos las representaciones léxicas en (1) y (2), podemos preguntarnos por qué algunas representaciones descomponen sus primitivos mientras que otras utilizan el mismo *definiendum* como *definiens*:

(1)

<i>ver</i> :	see' (x, y)
<i>cantar</i> :	do' (x, [sing' (x)])
<i>derretirse</i> :	BECOME melted' (x)
<i>romper</i> :	do' (x,φ)] CAUSE [BECOME broken' (y)]

⁵ Remitimos al trabajo de Mairal-Usón y Cortés-Rodríguez (2008) para una explicación detallada de los aspectos técnicos de estas representaciones léxicas.

En este primer grupo, observamos que en las dos primeras representaciones, que corresponden a un predicado de estado y otro de actividad, utilizamos el mismo predicado como *definiens* (**see'** y **sing'**), afirmación que podemos igualmente hacer extensiva a las representaciones en c) y d), con la excepción de que el uso del operador BECOME introduce el estado resultante. Este tipo de práctica contrasta con las representaciones en (2), donde cada predicado se descompone semánticamente: por ejemplo, *aprender* se define como llegar a saber, *recibir* como llegar a tener y *mostrar*, que es causativo, como alguien hacer algo tal que causa que *y* llegue a ver *z*.

(2)

<i>aprender</i>	BECOME know' (x, y)
<i>recibir</i>	BECOME have' (x, y)
<i>mostrar:</i>	[do' (x, φ)] CAUSE [BECOME [see' (y, z)]]
<i>cocinar:</i>	[do' (x, φ)] CAUSE [BECOME baked' (y)]

Es justo reconocer que esta observación no pasó inadvertida en la GPR y Van Valin (1993) ya advierte de la necesidad de mejorar el sistema de representación léxica y lograr descomposiciones semánticas más depuradas.

C) Advertimos que las estructuras lógicas sólo recogen aquellos aspectos del significado de una palabra que son gramaticalmente relevantes, es decir, que tienen un impacto sintáctico. La GPR sigue fielmente los postulados de aquellos enfoques que mantienen este posicionamiento y que, por consiguiente, excluyen todos aquellos factores que no tienen visibilidad sintáctica: por ejemplo, en relación a los verbos de sonido, Pesetsky (1995) argumenta que la intensidad con la que se emite el sonido no es un rasgo dirimente para la estructura sintáctica, razón por la que el referido parámetro no es determinante para la configuración sintáctica del predicado y, por consiguiente, no forma parte de la estructura argumental⁶. Sin embargo, no estamos tan seguros del alcance de esta afirmación pues hay casos en los que aquellos factores semánticos y pragmáticos que forman parte de la definición de un predicado tienen una incidencia sintáctica. Por ejemplo, Goldberg en González García (2008: 350) afirma que la agramaticalidad de **she whispered at him* en contraste con *she shouted/yelled/howled at him* puede fácilmente explicarse si recurrimos al modo en el que se emite el sonido.

⁶ Para una discusión detallada de cómo identificar aquellos aspectos del significado que forman parte de la estructura argumental, remitimos al lector al trabajo de Levin y Rappaport (2005).

Así las cosas, ¿sería posible diseñar un sistema de representación léxica que nos permita ir más allá de aquellos aspectos del significado de una palabra que son gramaticalmente relevantes, es decir, que tienen una presencia sintáctica? ¿Podemos dar cuenta del significado enciclopédico que un hablante posee sobre un determinado predicado? Por ejemplo, como hablantes nativos de español, somos conscientes de las diferencias entre estos verbos: *mandar*, *ordenar*, *decretar*, *preceptuar*, *preinscribir* o *cautivar*, *arrebatar*, *arrobar*, *embelesar*, *extasiar*, *hechizar*. ¿Cómo podríamos recoger en un sistema de representación estas diferencias? Aun más, en estrecha relación con el trabajo pionero de Fillmore (1982), o el más reciente en el marco de FrameNet, o, incluso, la extensión de esta línea de trabajo en el ámbito de la Gramática de Construcciones (cf. Goldberg, 2010), que llegan a proponer el ‘Conventional Frame Constraint’, mediante el cual una palabra evoca una unidad cultural, compartida por todos aquellos que la utilizan, podemos preguntarnos cómo podemos representar mediante un formalismo este conocimiento cultural y enciclopédico que comparten los hablantes de una comunidad lingüística. Mantenemos que este tipo de representación pasa por vincular el componente léxico a un módulo conceptual, como argumentamos en la Sección 3, y obtener así un formalismo con el suficiente alcance explicativo como para dar cuenta de aquellos aspectos del significado que superan el propio código y forman parte del acervo cultural de una comunidad.⁷

En suma, estos desafíos no pasaron desapercibidos para la GPR, que fue consciente de la necesidad de formular un sistema de representación semántica más depurado:

Because there is as yet no adequate decompositional representation for the primitive state and activity predicates which are the building blocks of the system (...).
(Van Valin y LaPolla, 1997:114)

En lo que sigue, presentamos las líneas maestras de estas primeras propuestas, lo que nos servirá para situar el nuevo formalismo que proponemos.

3. Un primer intento con las plantillas léxicas

⁷ Huelga decir que este tipo de formalismo desempeñará un papel fundamental en el desarrollo de aplicaciones en el ámbito del procesamiento del lenguaje natural, por ej. la recuperación de información (cf. Perrián Pascual y Mairal Usón, este volumen).

Como adelantábamos anteriormente, la GPR, ya desde 1993,⁸ empieza a trabajar en el perfeccionamiento del sistema de representación léxica y se proponen plantillas léxicas con un destacado basamento semántico. Consideremos la representación léxica de los verbos de actos de habla propuesta en Van Valin y LaPolla (1977:3.2.3.1.):

(3)

do' (x , [**express** (α).**to** (β).**in.language.**(γ)' (x,y)])

donde $\alpha = x$; $\beta = y$

Este tipo de representación consta de dos tipos de variables: las variables externas representadas por x , y , z etc. y las variables internas codificadas por las letras griegas, α , β , γ , etc. Las variables externas tienen una visibilidad sintáctica mientras que las variables internas dan cuenta de aquellos factores semánticos que forman parte del significado de un predicado. En principio, a partir de la realización de cada una de estas variables internas podemos explicar las diferentes configuraciones de estos verbos sin tener que recurrir a listar tres entradas, por ejemplo, en el caso de 'hablar', donde expresamos el destinatario ($y = \beta$) *Jaime habló con su compañero de trabajo*, o donde expresamos la lengua en la que hablamos ($y = \gamma$) *Jaime habla alemán con fluidez* o, el contenido proposicional ($\alpha = y$), *Jaime habló unas cuantas palabras*. En suma, las variables internas reflejan todos aquellos parámetros que conforman el contenido semántico de un predicado que, además, podemos asociar con cada una de las variables externas. Veamos el caso del predicado *prometer*, que pertenece a esta clase léxica de verbos e introduce un matiz causativo:

(4)

do' (x , [**express** (α).**to** (β).**in.language.**(γ)' (x,y)]) CAUSE [BECOME **obligated'** (x , w)])

donde $\alpha = w$; $\beta = y$ (Van Valin y LaPolla, 1997)

Estamos ante una representación en la que un ejecutor (x) realiza una actividad tal que x expresa algo (α) a un destinatario (β) en una lengua (γ) y esta actividad desencadena un resultado, a saber, que x pasa a adquirir la obligación de algo (w). Nótese que este argumento w está coindizado con la variable interna (α), lo que

⁸ Por ejemplo, el trabajo de Van Valin y Wilkins (1993) sobre el verbo *remember*.

significa que a lo que se compromete x coincide con lo que x ha expresado a y en la primera parte de la representación. Vemos cómo este sistema de representación es más ambicioso que las estructuras lógicas, pues logra incorporar material semántico que nos permite tener una visión más exacta del significado de un predicado.

En esta misma línea, Mairal (2003, 2004) desarrolla la descomposición semántica de varias clases léxicas. Consideremos la representación para los verbos de impacto en inglés, por ej. *strike, knock, tap, rap, cuff, kick, hammer, box, beat*, etc., donde, si dejamos otros detalles técnicos aparte, observamos que esta representación incluye las variables internas *manner, purpose e instrument*:

(5)

[[**do'** (w , [**use.tool.**(α ..**in.**(β)..**manner.for.**(γ)' (w , x)] CAUSE [**do'** (x , [**move.toward'** (x , y) & INGR **be.in.contact.with'** (y , x)], (α) = x .

La realización concreta de cada una de estas variables internas nos permite dar cuenta de las diferencias entre aquellos predicados que pertenecen a esta misma clase léxica. Así, la realización del parámetro semántico *manner* instancia los siguientes predicados:

(6)

MANNER (β); el golpe se puede medir en términos de:

❖ **Fuerza:**

- lightly: *tap, cuff*;
- hard: *strike, punch, thump, knock, clobber, clout, batter, cane, whip, lash, birch*;
- very hard (*sock, knock, clobber, bash*, etc.);

❖ **Sonido:**

- sharp, punctual, loud: (*crack, smack, whack*, etc.);
- dull, punctual (*thump, bump*);

❖ **Movimiento:**

- quick, (*tap, rap, swat*, etc.);
- swinging (*swat*),
- iterative (*rap, jab, batter, clobber*)

Además, a partir de esta representación, junto con un conjunto de reglas, podíamos predecir cada una de las configuraciones sintácticas de estos predicados. No obstante, si bien este enfoque resultaba realmente sugerente en cuanto a su ámbito de

explicación sintáctico, nos dimos cuenta de que resultaba excesivamente complejo: por un lado, formalizar los parámetros semánticos en primitivos y, por otro lado, expresar y acomodar la sucesión de eventos a los requisitos semánticos de cada una de las construcciones sintácticas. Además, desde el punto de vista de la expresión del formalismo, notamos que estas representaciones cada vez resultaban más abigarradas en su expresión pues llegamos a tener casos que constaban de una sucesión de tres o cuatro cadenas causales, motivadas por la necesidad de hacer constar la presencia de una determinada variable que diera cuenta de una determinada construcción sintáctica. En suma, en trabajos posteriores decidimos que era necesario pulir y simplificar el formalismo y mejorar el metalenguaje. A la vera de esta reflexión, propusimos la noción de plantilla léxica, la cual estaba provista de dos módulos de representación: un módulo semántico y un módulo lógico. Sea (7) como muestra del formato de una plantilla léxica:

(7)

predicado' : [módulo semántico <quale> [variables internas ligadas]i]
 [módulo lógico <estructuras lógica> [variable externas]i]

El tipo de representación léxica que proponíamos bajo el término *plantilla léxica*, consta, por tanto, de dos partes con variables de distinta naturaleza ligadas entre sí: unas son de índole netamente semántico-conceptual y recogen aspectos de nuestro conocimiento sobre un concepto que se pueden ligar a otras variables, de estructura lógico-argumental, que indican el potencial sintáctico de la base léxica. Es importante tener en cuenta que una plantilla léxica es algo más que una estructura lógica enriquecida con información semántico-conceptual, puesto que este tipo de información no sólo se liga a variables con proyección sintáctica, lo que no impone límites a la cantidad de conocimiento que se incluye en la representación. Consideremos la siguiente representación:

(8)

considerar: [LOC_{in}^{temp↔}_{1,2} CONT] **think'** (x, y)

x = 1; y = 2

Considerar es un predicado de estado bivalente cuyo *definiens* es el primitivo **think'**, que junto con **know'**, constituyen los términos básicos definitorios en el

dominio de cognición (cf. Faber y Mairal, 1999). El módulo semántico en (8) incluye dos funciones léxicas, $[LOC_{in}^{temp\leftrightarrow}{}_{1,2}]$ y $[CONT]$, con las que se expresan los parámetros semánticos de ubicación temporal (en el presente) y duración. La plantilla debe interpretarse del siguiente modo: una entidad (primer argumento) piensa sobre un percepto mental (segundo argumento) localizado en el presente ($LOC_{in}^{TEMP\leftrightarrow}$) durante un tiempo considerable (CONT).

Una de las cuestiones que más nos preocupaba era encontrar un sistema de representación que nos permitiera expresar con solvencia el contenido semántico de un predicado a la par que arbitrar un formalismo descrito en un metalenguaje simple en su formulación y comprensivo en su ámbito explicativo. A tal efecto, comprobamos que las funciones léxicas propuestas en la *Teoría Sentido Texto (Meaning Text Theory)* (cf. Mel'cuk 1989, Mel'cuk et al. 1995) resultaban insuficientes para este fin pues, recordemos, que éstas fueron utilizadas en un entorno sintagmático y con un fin lexicográfico: el *Diccionario de Colocaciones del Español* para la lengua española (Alonso Ramos, 2004) o el *Dictionnaire explicatif et combinatoire* para la lengua francesa (Mel'cuk et al. 1984). Igualmente, recurrimos a la teoría del lexicón generativo de Pustejovsky (1995, 2002, 2006)⁹. Por consiguiente, después de unos primeros estudios iniciales en los que utilizamos las funciones léxicas propuestas en Melchuk (por ej. Mairal y Faber, 2007), unificamos todo el sistema de representación léxica en torno a la estructura de los *qualia*¹⁰. Reparemos en la representación de los verbos de consumo (*comer, beber, engullir, etc.*), cuya estructura lógica ya mencionamos anteriormente en la figura 1:

(9)

Módulo lógico:

$[do'(x, e_1)]_{E1} \& [INGR NOT exist'(y)]_{E2}, E1 < E2$

Módulo semántico:

{Q_F: MANNER E1

Q_A: e₁: $do'(x, [CAUSE.BECOME.LOC^{in} \cdot (part_of'x, y)]$

Q_T: E2}

⁹ Remitimos al trabajo de De Miguel (2008) para una excelente exposición de esta teoría.

¹⁰ Para una justificación sobre las ventajas que nos ofrece la estructura de *qualia* sobre las funciones léxicas, remitimos al lector al trabajo de Mairal Usón y Ruiz de Mendonza Ibáñez (2008).

Por lo que se refiere a la estructura lógica (o eventiva), esta plantilla incorpora un evento puntual existencial [INGR NOT **exist'** (y)] precedido temporalmente de una actividad [**do'** (x, e₁)]. La naturaleza específica de esta actividad queda descrita en el *Quale* Agentivo [e₁: **do'** (x, [CAUSE.BECOME.LOCⁱⁿ. (**part_of'**x, y))]. Este *quale* viene a indicar el hecho de que los verbos de consumo conllevan una estructura causal por la que un consumidor (x) coloca la entidad afectada (y) dentro de su cuerpo. El *quale* formal especificará las diversas formas de consumir que quedan lexicalizadas en la estructura semántica de los miembros de esta clase y que permiten establecer distinciones semánticas entre ellos. La conexión entre los eventos ligados causalmente en el *quale* agentivo y el logro resultante final (que se corresponde con el *quale* télico) es una relación infraespecificada E1 < E2. Esta relación tiene dos implicaciones básicas: (1) los eventos están ordenados secuencialmente (de hecho el símbolo & en el módulo lógico debe interpretarse como 'y entonces'), y (2) cualquiera de los dos eventos puede ser 'priorizado' (*foregrounded*) en el proceso de enlace semántica-sintaxis.

En suma, a pesar de disponer de un sistema de representación léxicas más ambicioso en cuanto a su alcance explicativo, todavía tenemos tres grandes retos pendientes, como hemos señalado arriba: determinar el mecanismo de selección de los primitivos, concretar su supuesto alcance universal y, finalmente, demostrar cómo podemos obtener conocimiento enciclopédico a partir de una representación léxica, aspecto que nos conduce a delimitar la frontera, con frecuencia borrosa, entre el conocimiento lingüístico y el extralingüístico.

4. La noción de Estructura Lógica Conceptual

Con el fin de buscar respuestas a los tres grandes retos que señalábamos, decidimos imprimir una nueva orientación metodológica a nuestro enfoque y así explorar el potencial explicativo que podía ofrecernos un enfoque ontológico para el tratamiento del léxico. A tal efecto, vinculamos nuestro componente lexicón a FunGramKB, una base de conocimiento léxico-conceptual multipropósito, desarrollada por Perrián Pascual y Arcas Túnez (2004, 2005, 2007, 2010a, 2010b). El nuevo marco de representación tiene ahora el siguiente formato (remitimos al lector al capítulo de Perrián Pascual y Mairal Usón (este volumen) para una explicación más detallada de la arquitectura de la base de conocimiento:

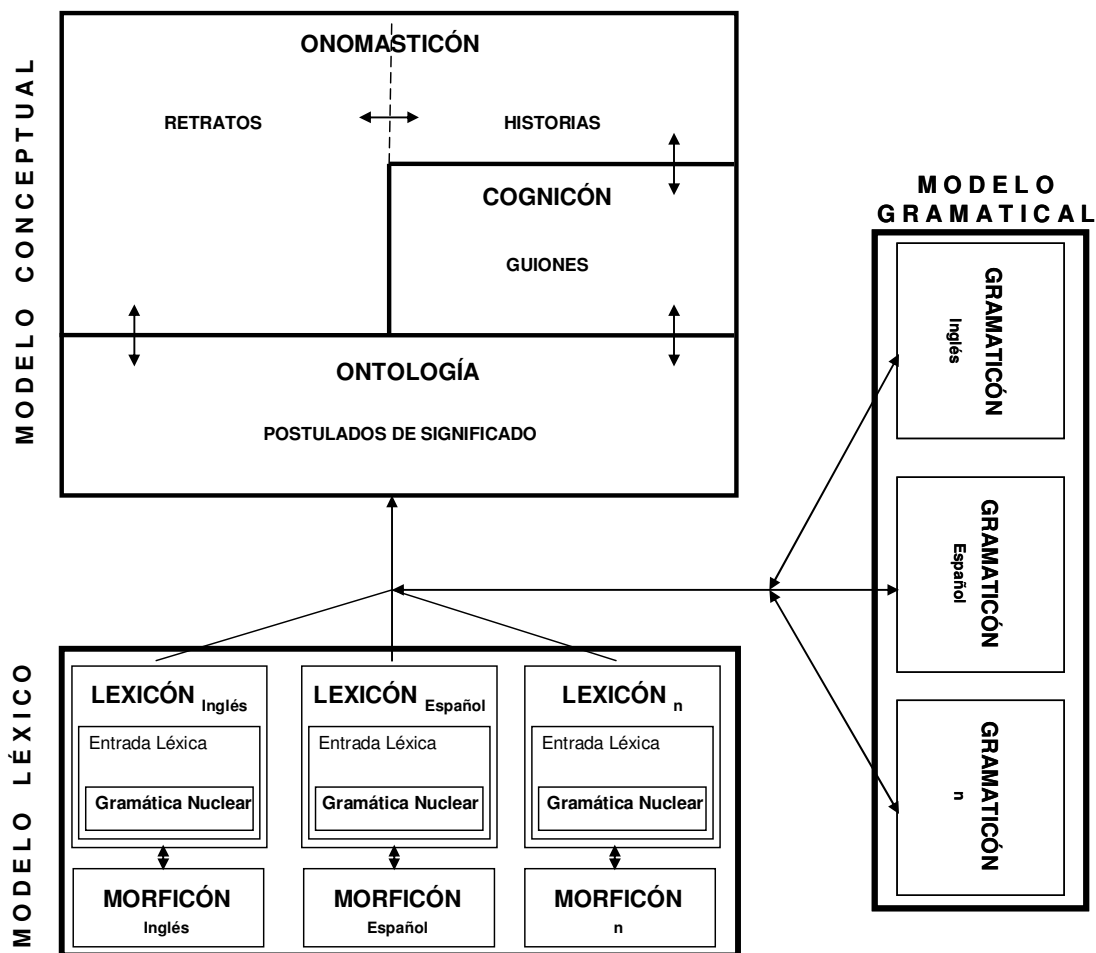


Figura 2: La arquitectura de FunGramKB

Para lo que aquí nos ocupa, nótese que el componente léxico se nutre directamente de la ontología, es decir, postulamos una interfaz léxico conceptual que deberá dar cuenta de cómo construimos las representaciones léxicas resultantes, representaciones que, recordemos, hemos denominado ELCs. La formación de una ELC es un proceso totalmente automatizado a través de una herramienta denominada Constructor de ELC, el cual precisa de la información del *Aktionsart* y de la plantilla léxica (i.e. número de variables, rasgos idiosincrásicos y proyección de los papeles temáticos) en la gramática nuclear de la unidad léxica en cuestión.

Una ELC funciona como una interlingua pivote entre el razonador, que a su vez se nutre de otra interlingua codificada en un lenguaje de representación propio llamado

COREL (*Conceptual Representation Language*),¹¹ y el algoritmo léxico-sintáctico, es decir, entre el componente lingüístico y el conceptual:

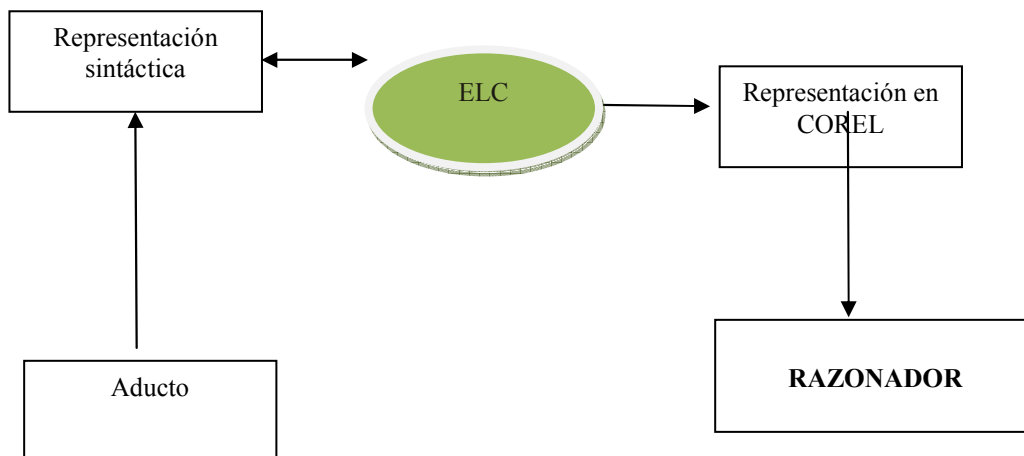


Figura 3: El papel de la ELC como interlingua

Veremos que el hecho de poder traducir una ELC a COREL, i.e. una otra interlingua orientada al razonamiento, nos permitirá enriquecer nuestra ELC con el conocimiento almacenado en toda la base de conocimiento, lo que, en última instancia, constituye un paso firme hacia la consecución de dotar a nuestras representaciones de ese conocimiento cultural y enciclopédico que Goldberg (2010) glosa en torno al *Conventional Frame Constraint*, cuestión que retomaremos en la Sección 4.2.3.

De la Figura 1, podemos colegir la naturaleza ontológica del léxico y el basamento conceptual de las representaciones léxicas. No nos debe sorprender, pues, que cada palabra en los diferentes léxica esté asociada a un concepto en la ontología. Por ejemplo, las entradas léxicas del predicado ‘quejarse’ en español, ‘complain’ en inglés, ‘se plaindre’ en francés, ‘protestare’ en italiano y ‘oplacha se’ en búlgaro están vinculadas al concepto +COMPLAIN_00, como muestra la siguiente ilustración:

¹¹ Para una descripción de la sintaxis y la semántica de COREL, véanse Perrián Pascual y Arcas Túnez (2004) y Perrián Pascual y Mairal Usón (2010).

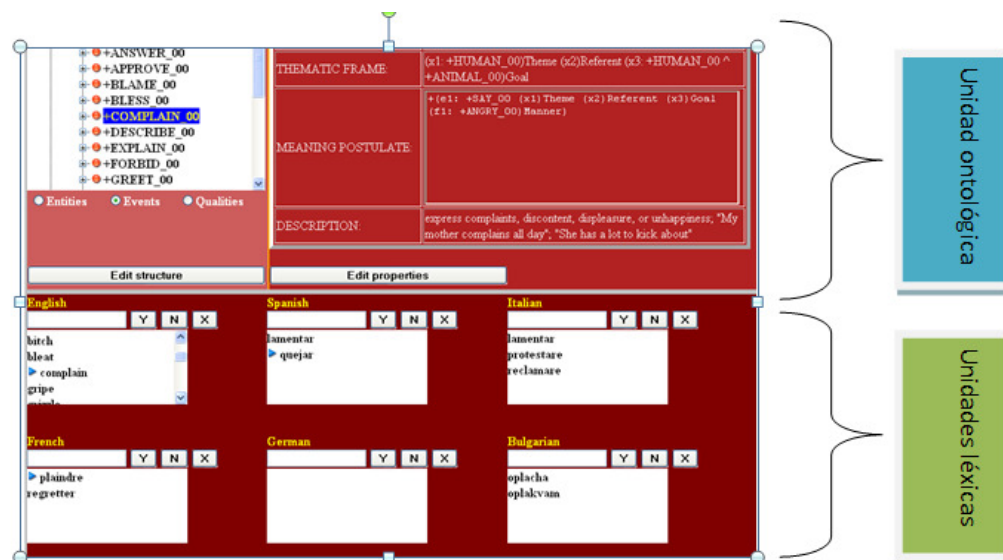


Figura 4: La relación entre los diferentes léxica y la ontología

Partimos de una separación nítida entre el léxico y la ontología, cada uno con cometidos independientes aunque claramente interconectados. Por un lado, la ontología incluye un repositorio de conceptos, mientras que el léxico incluye las propiedades idiosincrásicas de las entradas léxicas para cada una de las lenguas. Por consiguiente, podemos aventurar que la ontología tiene un estatus ‘universal’, en el sentido de que cada concepto que podamos imaginar puede ocupar un lugar adecuado en esta ontología (Periñán Pascual y Arcas Túnez 2007), mientras que el léxico es específico de cada lengua (cf. Mairal Usón y Periñán Pascual, 2009). Este punto es realmente interesante pues, sin entrar en disquisiciones sobre el alcance del adjetivo ‘universal’, nos abre una vía de trabajo para tratar la cuestión de la universalidad de los primitivos. Podemos establecer la siguiente línea de argumentación: si mantenemos que los conceptos tienen un estatus universal y las ELCs están formadas por conceptos, como veremos más adelante, podemos concluir que los primitivos que utilizamos en las ELCs tienen una verdadera justificación sobre su naturaleza universal.

Lo que cumple ahora es argumentar y justificar el por qué de este giro conceptualista. A modo de adelanto, este nuevo enfoque nos permite simplificar mucho el formalismo, pues no es necesario estipular dos módulos, i.e. el semántico y el lógico, como hicimos en las plantillas léxicas. Además, estamos ante un formalismo que no sólo nos permite vincular las representaciones léxicas a una base de conocimiento, lo cual nos permitirá concebir nuestras representaciones en un contexto que trasciende los requisitos lingüísticos, sino también es capaz de incorporar aquel conocimiento cultural

y enciclopédico característico de una comunidad lingüística. Tratemos ahora con mayor detenimiento el alcance explicativo de las ELC.

4.2. Las ELC: implicaciones para la teoría

Esta sección retoma los tres grandes desafíos que señalábamos arriba y nos proporciona las respuestas que podemos obtener en el contexto de un entorno ontológico.

4.2.1. El formato de representación

Si mantenemos que la ontología actúa de pivote central sobre el que descansa y se nutre el léxico, entonces no nos sorprenderá si afirmamos que los primitivos de una ELC no serán predicados sino unidades conceptuales tomadas directamente de la ontología. Sea, entonces, el siguiente formato de representación para cada una de las clases léxicas aspectuales, donde la metavariable $\langle C \rangle$ puede representar a cualquier concepto básico o terminal en la Ontología de FunGramKB:

CLASE VERBAL	ESTRUCTURA LÓGICO CONCEPTUAL
Estado	$\langle C \rangle (x) \text{ o } (x,y)$
Actividad	do $(x, [\langle C \rangle (x) \text{ o } (x,y)]$
Logro	INGR $\langle C \rangle (x) \text{ o } (x,y)$, o INGR do $(x, [\langle C \rangle (x) \text{ o } (x,y)]$
Semelfactivo	SEML $\langle C \rangle (x) \text{ o } (x,y)$ SEML do $(x, [\langle C \rangle (x) \text{ o } (x,y)]$
Realización	BECOME $\langle C \rangle (x) \text{ o } (x,y)$, o BECOME do $(x, [\langle C \rangle (x) \text{ o } (x,y)]$
Realización activa	do $(x, [\langle C \rangle (x, (y))])$ & BECOME $\langle C \rangle (z,x) \text{ o } (y)$
Causativos	α CAUSE β , donde α y β son estructuras lógicas de cualquier tipo

Figura 5: Inventario de ELCs

Si comparamos esta clasificación con la figura 1 vemos que la única diferencia es la sustitución de los primitivos por conceptos. Entonces, ¿qué tipo de ventajas nos reporta este nuevo sistema de representación? Recordemos que una de los retos que apuntábamos concernía la descomposición semántica de los primitivos. Ahora vemos que cada primitivo es un concepto. Cumple preguntarse en qué medida un concepto nos proporciona una descomposición semántica más granular.

En este sentido, es conveniente señalar la orientación semántica de FunGramKB, que, a diferencia de otras bases de conocimiento con un enfoque relacional (por ej. SIMPLE o EuroWordnet, etc.), está inspirada en un enfoque de semántica profunda y, por consiguiente, proporciona definiciones para cada una de las unidades conceptuales de la ontología. Cada concepto consta de una serie de propiedades: los marcos temáticos y los postulados de significado (cf. Perrián Pascual y Arcas Túnez, 2007, 2010a; Perrián Pascual y Mairal Usón, 2010). Un postulado de significado incluye la definición de una unidad conceptual mediante el lenguaje de representación COREL, el cual nos permite definir todas las unidades conceptuales con la ventaja añadida de que las relaciones conceptuales pueden igualmente obtenerse aplicando mecanismos de herencia e inferencia sobre los propios postulados de significado (cf. Perrián Pascual y Arcas Túnez, 2005), lo que nos permitirá utilizar la ELC como aducto para un motor de razonamiento. Con respecto a la descomposición de los primitivos que conforman cada representación, vemos que los postulados de significado nos proporcionan la tan aclamada descomposición semántica de forma natural pues, recordemos, cada pieza léxica está vinculada a un concepto. Consideremos la ELC para el predicado ‘rezar’ en su uso intransitivo:

(10)

rezar: **do** (x-Theme, [+ **PRAY_00** (x-Theme)])

Ante esta representación, deducimos que es un predicado de actividad que toma como primitivo la unidad conceptual +PRAY_00. Pero, ¿no es posible llegar a descomponer el concepto que funciona como primitivo? Si consultamos las propiedades conceptuales de +PRAY_00, obtenemos la siguiente información:

- Marco temático:

(x1: +HUMAN_00)Theme (x2)Referent (x3: +GOD_00 ^ +SAINT_00)Goal

- Postulado de significado:

(e1: +SAY_00 (x1)Theme (x2)Referent (x3)Goal (f1: (e2: +THANK_00 (x1)Theme (x4)Referent (x3)Goal))Purpose | (f2: (e3: +REQUEST_01 (x1)Theme (x4)Referent (x3)Goal))Purpose)

+PRAY_00 forma parte de la dimensión metaconceptual #COMMUNICATION,¹² por lo que su marco temático está definido por tres participantes: un Tema (*Theme*), i.e. entidad que transmite un mensaje, un Mensaje (*Referent*), i.e. el contenido proposicional que se transmite, y una Meta (*Goal*), i.e. la entidad a quien transmitimos el mensaje. Nótese que en este marco temático la tercera entidad está restringida prototípicamente a las unidades +GOD_00 y +SAINT_00. Además, obtenemos una definición en su postulado de significado que nos dice que un emisor dice algo a un receptor con el propósito de que o bien ese emisor agradece algo al receptor, o bien porque el emisor pide algo al receptor, o bien por los dos motivos a la vez. Entonces, vemos cómo la representación para ‘rezar’ incluye una descomposición semántica de este predicado en unidades conceptuales básicas. De hecho, es posible alcanzar una representación semántica en COREL formada por primitivos semánticos cuando aplicamos el mecanismo de “descomposición conceptual gradual” (Periñán Pascual y Arcas Túnez, 2005), donde el *definiens* en un postulado de significado puede convertirse en el *definiendum* de otro postulado de significado. Esto nos permite la construcción de postulados de significado de una manera simple, además de minimizar la redundancia de información.¹³

4.2.2. La elección del primitivo

Una pregunta recurrente que se ha planteado en diversos foros de la GPR y talleres especializados sobre representación léxica está relacionada con la elección de los primitivos que forman parte del formalismo: ¿cómo sabemos qué primitivo debemos elegir? ¿Por qué utilizamos palabras en inglés y no en cualquier otro idioma? De nuevo, entendemos que podemos hallar respuestas a estas preguntas cuando recurrimos a la información que nos proporciona la ontología. Así, podemos estipular el siguiente principio: el significado denotativo de una palabra está ligado a un concepto. Por

¹² De hecho, la ruta conceptual de +PRAY_00 es #EVENT >> #COMMUNICATION >> +SAY_00.

¹³ Este mecanismo está inspirado en el principio de “descomposición léxica gradual” de Dik (1989), excepto por el hecho de que, en lugar de basarnos en predicados de una lengua, nuestros postulados de significado se construyen a partir de unidades ontológicas.

ejemplo, cuando representamos la ELC para el verbo ‘escribir’ en su uso transitivo, proponemos la siguiente representación:

(11)

do (x-Theme, [+WRITE_00 (x-Theme, (y-Referent))] & INGR +WRITE_00 (y-Referent)

Pero, ¿por qué elegimos +WRITE_00, y no +DRAW_00 o +EXIST_00, como se recoge en la estructura lógica estándar de este verbo? La elección del primitivo viene pautada por la propia configuración de la ontología, que nos dice que el verbo ‘escribir’ en español está vinculado al concepto básico +WRITE_00¹⁴. Este concepto, además, está representado por las siguientes propiedades:

(12)

Marco Temático:

(x1: +HUMAN_00 ^ +ANIMAL_00)_{Theme} (x2: +WRITING_00 ^ +IMAGE_00)
Referent

Postulado de Significado:

+(e1: +CREATE_00 (x1)Theme (x2)Referent (f1: +PENCIL_00 ^ +PEN_00 ^ +MACHINE_00)Instrument (f2: +PAPER_00 | +WALL_00)Location (f3: +ON_00)Position)

+WRITE_00 está ligado a la dimensión metaconceptual #CREATE y toma dos argumentos como parte de su marco temático: un Tema, i.e. una entidad que crea otra entidad, y un Referente, i.e. la entidad que es creada por otra entidad. Además, su postulado de significado nos proporciona una descomposición semántica del primitivo. Por otra parte, con respecto a la cuestión de por qué los primitivos suelen presentarse en inglés, nótese que, desde nuestro enfoque ontológico, es realmente irrelevante la lengua utilizada para describir los conceptos, pues no se trata en realidad de palabras específicas de una lengua. En vez de elegir +WRITE_00 como etiqueta, no perderíamos

¹⁴ Evidentemente, una pregunta que surge como reacción a esta afirmación es si disponemos de algún enfoque metodológico que justifique toda la disposición que proponemos en la ontología. La respuesta se encuentra en Perinián Pascual y Mairal Usón (en preparación), donde se describe detalladamente la metodología COHERENT (COncceptualization + HiERarchization + REmodellling + refinemNT), la cual sirvió para realizar el proceso de transducción de las palabras definitorias del *Longman Dictionary of Contemporary English* (Procter 1978) a la taxonomía IS-A de conceptos básicos de FunGramKB.

capacidad explicativa si, en su lugar, decidiéramos usar +667788_00. Ahora bien, en este último caso, la etiqueta del concepto es mucho menos significativa. En esencia, la propia configuración ontológica guía al lingüista en la elección del primitivo que formará parte de la representación.

Además, un enfoque conceptual como el que mantenemos nos ofrece un protocolo pautado en cuando a la elección del primitivo en aquellos predicados de naturaleza polisémica. Como corolario, podemos postular que si el significado denotativo cambia, entonces ese nuevo significado estará vinculado a un concepto diferente. En este sentido, podemos calificar nuestro enfoque como un lexicón enumerativo robusto, el cual proporciona una fina granularidad semántica en sus postulados de significado al tiempo que teje una red conceptual extensa, lo que nos permite esquivar la cuestión controvertida de reducir a una única representación infraespecificada todos los significados de una misma unidad léxica, como es la práctica habitual en la teoría del lexicón generativo. Para FunGramKB, a diferencia de otros enfoques, las diferencias de significado no vienen marcadas por el comportamiento sintáctico de una palabra, sino por su significado denotativo. Por consiguiente, los diferentes matices de significado propios de cada una de las construcciones sintácticas no inciden en la representación conceptual (en la elección del primitivo o unidad conceptual) sino que, recordemos, damos cuenta de esta variación en la propia ELC, la cual variará sus argumentos de acuerdo al tipo de configuración sintáctica. Por ejemplo, codificaremos en el componente léxico los usos transitivos e intransitivos de *romper* mediante una ELC causativa y otra de realización respectivamente. Ilustremos, como coda final a este punto, la elección de los primitivos para el predicado “arreglar” con sus dos sentidos: (i) poner en orden algo (por ej. *Mi madre arregló las flores del jarrón*) y (ii) reparar algo cambiando algunas de las piezas o juntando algo que estaba roto (por ej. *Mi padre arregló el televisor*). Cada uno de estos sentidos está vinculado a un concepto diferente: +TIDY_00 y +REPAIR_00 respectivamente, por lo que su representación tendrá el siguiente formato:

(13)

- a. **do** (x-Theme, [+TIDY_00 (x-Theme, y-Referent)])
- b. **do** (x-Theme, [+REPAIR_00 (x-Theme, y-Referent)])

Estos dos primitivos tienen rutas conceptuales diferentes. En el caso de +TIDY_00, obtenemos la ruta conceptual #MATERIAL >> +DO_00 >> TIDY_00 y el postulado de significado que mostramos en (13): alguien hace algo con el resultado de que ese algo adquiere la propiedad de limpio. En el caso de REPAIR_00, la ruta conceptual es #MATERIAL >> #TRANSFORMATION >> +CHANGE_00 >> +REPAIR_00 y el postulado de significado nos proporciona la siguiente definición: alguien cambia algo de una entidad, porque ésta se dañó, con el resultado de que alguien pueda utilizar dicha entidad.

(14)

- a. +(e1: +DO_00 (x1)Theme (x3)Referent (f1: (e2: +BE_01 (x2)Theme (x4: \$TIDY_D_00)Attribute))Result)
- b. +(e1: +CHANGE_00 (x1)Theme (x2)Referent (f1: (e2: past +DAMAGE_00 (x3)Theme (x2)Referent))Reason (f2: (e3: pos +OPERATE_00 (x4)Theme (x2)Referent))Result)

En suma, si adoptáramos un enfoque lexicista, tendríamos que estipular *ad hoc* aquel primitivo que más se acercase al significado de la palabra, sin ningún tipo de criterio o argumentación que justificara la referida elección. En cambio, en nuestro caso, la propia disposición ontológica determina qué unidad conceptual debe formar parte de la representación.

5. Conclusiones

Este capítulo nos ofrece argumentos a favor de una revisión de las estructuras lógicas de la GPR y nos propone un nuevo sistema de representación léxica denominado ELC. Este giro metodológico hacia una semántica ontológica en detrimento de un enfoque exclusivamente léxico viene motivado por una serie de observaciones que señalamos en la sección 3: por ej. la naturaleza de los primitivos, la representación de aquellos factores del significado que trascienden el ámbito gramatical, etc. Como parte de la nueva orientación ontológica, postulamos una base de conocimiento léxico-conceptual, FunGramKB, en lugar de una mera base de datos léxica. De esta forma, se dota de significación semántica al léxico, es decir, el modelo ya no comienza en el componente léxico sino en el nivel conceptual. Por consiguiente, el componente léxico deja de ser el motor de arranque de la maquinaria lingüística para ser recipiendario de

todo el caudal de información que le aporta el nivel conceptual y, más en particular, la ontología. En la sección 4, explicamos el formato de las ELCs y señalamos las ventajas metodológicas que obtenemos al adoptar un enfoque de representación léxica como éste.

Como coda final a este capítulo, nos gustaría adelantar la línea central del capítulo XX de este volumen, donde, destacamos la potencialidad de este formalismo en el ámbito del procesamiento del lenguaje natural, lo que significa abrir una puerta hasta ahora desconocida para la GPR: su implementación computacional.

6. Referencias

- Alonso Ramos, M. (2004), «Elaboración del Diccionario de colocaciones del español y sus aplicaciones», en P. Battaner y J. DeCesaris (eds.), *De Lexicografía: Actes del I Symposium Internacional de Lexicografía*. Barcelona, IULA, Documenta Universitaria, pp. 149-162.
- De Miguel, E. (2008), «La Teoría del Lexicón Generativo», en E. de Miguel (ed.), *Panorama de la lexicología*, Barcelona, Ariel, pp. 337-368.
- Dik, S.C. (1989), *The Theory of Functional Grammar*, Berlin/New York, Mouton de Gruyter.
- Faber, P. y Mairal Usón, R. (1999), *Constructing a Lexicon of English Verbs*, Berlin/Nueva York, Mouton de Gruyter.
- Fillmore, C.J. (1982), «Frame Semantics», en Linguistic Society of Korea (ed.), *Linguistics in the Morning Calm*, Seoul, Hanshin, pp. 111-138.
- Goldberg, A.E. (2006), *Constructions at Work: The Nature of Generalization in Language*, New York, Oxford University Press.
- Goldberg, A. (2010), «Verbs, constructions and semantic frames», en Rappaport Hovav, M., Doron, E. y Sichel, I. (eds.), *Syntax, Lexical Semantics and Event Structure*, Oxford, Oxford University Press.
- González-García, F. (2008), «Construction Grammar Works: An interview with Adele E. Goldberg», *Annual Review of Cognitive Linguistics*, 6, pp. 345-360.
- Jackendoff, R. (1990), *Semantic structures*, Cambridge, MA, MIT Press.
- Levin, B. (1985), «Lexical Semantics in Review: An Introduction», en Levin, B. (ed.), *Lexical Semantics in Review*, Lexicon Project Working Papers 1, Center for Cognitive Science, Cambridge, MA, MIT, pp. 1-62.
- Levin, B. y Rappaport M. (2005), *Argument Realization*, Oxford, Oxford University Press.
- Mairal, R. (2003) «Why the notion of lexical template?», *Anglogermánica online* 2002–1.
- , (2004), «La reglas léxicas en la Gramática del Papel y la Referencia», en G. Wotjak, G. y J. Cuartero Otal (eds.), *Algunos problemas específicos de la descripción sintáctico-semántica*, Berlín, Frank and Timme Verlag, pp. 175–196.
- Mairal Usón, R. y Cortés, F. (2008), «Modelos teóricos del estudio del léxico: los modelos funcionales», en E. de Miguel (ed.) *Panorama de la Lexicología*, Ariel, Barcelona, pp. 247-280.

- Mairal Usón, R. y Faber, P. (2007), «Lexical templates within a functional cognitive theory of meaning», *Annual Review of Cognitive Linguistics* (5), Amsterdam, John Benjamins, pp. 137-172.
- Mairal Usón, R. y Perrián Pascual, C. (2009), «The anatomy of the lexicon component within the framework of a conceptual knowledge base» *Revista Española de Lingüística Aplicada* 22, pp. 217-244.
- Mairal Usón, R. y Ruiz de Mendoza Ibáñez, F. J. (2008), «New challenges for lexical representation within the Lexical-Constructional Model (LCM)», *Revista Canaria de Estudios Ingleses*. Vol. 57 (2008), pp. 137-158.
- Mel'cuk, I. (1989), «Semantic primitives from the viewpoint of the Meaning-Text Linguistic Theory». *Quaderni di Semantica* 10 (1), pp. 65-102.
- Mel'cuk, I., Arbathefsky-Jumarie, N., Elnitsky, L., Iordanskaja, L. y Lessard, A. (1984), *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques*, Montreal, University of Montreal Press.
- Mel'cuk, I., Clas, A., y Polguère, A. (1995), *Introduction à la lexicologie explicative et combinatoire*, Louvain-la-Neuve (Belgium): Duculot/ Aupelf - UREF.
- Perrián Pascual, C. y Arcas Túnez, F. (2004), «Meaning postulates in a lexico-conceptual knowledge base», *Proceedings of the 15th International Workshop on Databases and Expert Systems Applications*, Los Alamitos (California), IEEE, pp. 38-42.
- , (2005), «Microconceptual-Knowledge Spreading in FunGramKB», *Proceedings on the Ninth IASTED International Conference on Artificial Intelligence and Soft Computing*, Anaheim-Calgary-Zurich, ACTA Press, pp. 239-244.
- , (2007), «Cognitive modules of an NLP knowledge base for language understanding», *Procesamiento del Lenguaje Natural* 39, pp. 197-204.
- , (2010a), «Ontological commitments in FunGramKB», *Procesamiento del Lenguaje Natural* 44, pp. 27-34.
- , (2010b), «The Architecture of FunGramKB», *Proceedings of the Seventh International Conference on Language Resources and Evaluation*, Malta, ELRA, pp. 2667-2674.
- Perrián Pascual, C. y Mairal Usón, R. (2009), «Bringing Role and Reference Grammar to natural language understanding». *Procesamiento del Lenguaje Natural*, 43 , pp. 265-273.
- , (2010), «La gramática de COREL: un lenguaje de representación conceptual». *Onomázein* 21, pp. 11-45.
- , (este volumen) «La dimensión computacional de la GPR: la estructura lógica conceptual y su aplicación en el procesamiento del lenguaje natural».
- , (en preparación) Constructing the FunGramKB basic conceptual level: the COHERENT methodology.
- Pesetsky, D. (1995), *Zero Syntax*, Cambridge, MA, MIT Press.
- Procter, P. (ed.), (1978), *Longman Dictionary of Contemporary English*, Longman, Harlow.
- Pustejovsky, J. (1995), *The Generative Lexicon*, Cambridge, MA, The MIT Press.
- , (2002), «Polysemy and Underspecification», en Behrens, L. y Zaefferer, D. (eds.), *The Lexicon in Focus. Competition and Convergence in Current Lexicology*, Frankfurt am Main, Peter Lang, pp. 187-208.
- , (2006), «Type Theory and Lexical Decomposition», *Journal of Cognitive Science*, 6, pp. 39-76.

- Rappaport, M. y Levin, B. (1998), «Building Verb Meanings», en Butt, M. y Geuder, W. (eds.), *The Projection of Arguments: Lexical and Compositional Factors*, Stanford, CSLI, pp. 97-134.
- Ruiz de Mendoza Ibáñez, F. y Mairal Usón, R. (2008), «Levels of description and constraining factors in meaning construction: an introduction to the Lexical Constructional Model», *Folia Linguistica* 42/2, pp. 355–400.
- Van Valin, R.D. Jr. (ed.), (1993), *Advances in Role and Reference Grammar*, Amsterdam y Philadelphia, John Benjamins.
- , (2004) «Lexical representation, co-composition, and linking syntax and semantics.», [<http://wings.buffalo.edu/soc-sci/linguistics/people/faculty/vanvalin/rrg.html>]
- , (2005), *The Syntax-Semantics-Pragmatics Interface: An Introduction to Role and Reference Grammar*. Cambridge, Cambridge University Press.
- , (2006), «Some universal of verb semantics», en Mairal, R. y Gil, J. (eds.), *Linguistic Universals*, Cambridge, Cambridge University Press, pp. 155-178.
- Van Valin, R.D. Jr. y Wilkins, D. (1993), «Predicting syntax from semantics», en Van Valin, R.D. Jr. (ed.), *Advances in Role and Reference Grammar*, Amsterdam/Philadelphia, John Benjamins, pp. 499–534.
- Van Valin, R.D. Jr. y LaPolla, R. (1997), *Syntax, Structure, Meaning and Function*, Cambridge, Cambridge University Press.
- Vendler, Z. (1967), *Linguistics in Philosophy*, Ithaca, NY, Cornell University Press.