

## CAPÍTULO 5

### LA TEORÍA DEL LÉXICO EN LOS NUEVOS DICCIONARIOS<sup>1</sup>

OLGA BATIUKOVA

*Brandeis University / Universidad Autónoma de Madrid*

#### 1. Introducción

La lexicografía y la lexicología, a pesar de compartir el objeto del estudio, no han interactuado de manera productiva hasta hace relativamente poco. Una de las razones de ello radica seguramente en la diferencia de objetivos de ambas disciplinas: la lexicología como parte de la lingüística teórica ha de dar respuestas sobre la estructura y el funcionamiento del léxico; la meta de la lexicografía es, en cambio, de tipo aplicado: consiste en recoger el vocabulario y describirlo en forma de repertorios. Esta diferencia de objetivos tiene varias consecuencias (cfr. Geeraerts, 1989): 1) la teoría lingüística no tiene por qué dar cuenta de todos los datos —se ignoran los fenómenos que se supone que deben tratar otras (sub)disciplinas—; la lexicografía, a su vez, tiene que integrar todos los aspectos importantes del uso de una palabra; 2) la asimilación lexicográfica de los modelos lingüísticos no siempre es perceptible y coherente desde el punto de vista teórico, y 3) los principios teóricos solo son uno de los parámetros que tiene que considerar la lexicografía.<sup>2</sup>

En cuanto a los beneficios de la colaboración entre lingüistas y lexicógrafos, se puede decir que la teoría del lenguaje proporciona la base lingüística a los diccionarios, dotándolos de coherencia interna. Las obras lexicográficas, a su vez, son una importante fuente de datos para los estudios lingüísticos, y al mismo tiempo su banco de pruebas.

1. No puedo dejar de agradecer a todas las personas en Madrid y Brandeis que me han dado ánimos y han sacrificado parte de su tiempo por ayudarme con comentarios, referencias y redacción de estilo. En primer lugar, a Elena de Miguel por su confianza y apoyo; su papel aquí supera con mucho el de una editora. Mil gracias asimismo a Carlos Piera, Santiago Urbano Sánchez, Roser Saurí, José Antonio Pascual, Pedro Álvarez de Miranda, Joaquín Garrido, Anna Rumshisky, James Pustejovsky y Polonca Kocjančič.

2. Entre otros factores específicos se encuentra, por ejemplo, el tipo de usuario al que va dirigida la obra.

La dificultad de distinguir entre los aspectos lexicográficos prácticos y teóricos hace que a veces se identifiquen erróneamente la lexicología y la lexicografía teórica (la *metalexicografía*). Véase al respecto Wiegand (1998), Zgusta (2006: 121-122), Porto Dapena (2002: § 1.1), y Azorín Fernández (2003, § 2).

El desarrollo de la relación entre la lexicología y los diccionarios se vio obstaculizado por ciertas premisas teóricas, como, p. ej., que el léxico es un conjunto no ordenado de propiedades irregulares e impredecibles de las palabras (para el estructuralismo americano de principios del siglo xx y la primera etapa de la gramática generativa, p. ej.).<sup>3</sup> Esta concepción (que se supera dentro del generativismo con el «giro lexicista», cfr. el capítulo 4 de la tercera parte de este volumen), aparte de separar drásticamente el léxico y la gramática, se corresponde con la organización de las entradas léxicas en forma de compartimientos aislados de rasgos (como en los diccionarios de lengua convencionales, que no suelen aludir a las relaciones entre palabras, limitándose a indicar casos de sinonimia/antonimia y familia de palabras). En un modelo así, solo se dan dos tipos de situaciones: las piezas léxicas relacionadas se incluyen en la misma entrada, y las no relacionadas, en entradas distintas. Esta premisa es tanto más perniciosa en cuanto que las relaciones paradigmáticas y sintagmáticas de una palabra son las que determinan en gran medida qué es dicha palabra.

La superación de estos supuestos teóricos ha ido acompañada de ciertos avances en el seno de la lexicografía que han cambiado su naturaleza de puramente práctica o técnica a una materia «de base lingüística sobre el método utilizado en la redacción de aquellas obras y sobre los principios que las inspiran» (Rafel, 2005: 13).<sup>4</sup> Uno de los avances es el cambio de la perspectiva normativa a la descriptiva: los diccionarios normativos reflejan aquellos usos lingüísticos que en la época dada se consideran como un buen uso; los diccionarios descriptivos, al menos en principio, se limitan a registrar los fenómenos lingüísticos.<sup>5</sup> Otro cambio atañe a la aplicación del criterio diacrónico en la elaboración de diccionarios, que repercute, p. ej., en la ordenación de las acepciones (se menciona en primer lugar la más antigua o la más cercana al étimo) y en la selección de material léxico (se incluyen voces anticuadas). Hoy en día se considera que el uso del criterio diacrónico más allá de los diccionarios históricos y etimológicos no está justificado (Shcherba, 1940). Aparte de estos y otros avances, la introducción del ordenador ha tenido que ver con muchos cambios en el trabajo lexicográfico (§ 2.2).

Lo que determinará el curso de la exposición en lo que sigue serán los aspectos del estudio teórico del léxico que revisten mayor relevancia para la lexicografía:

1. Las diferentes maneras de definir los elementos del significado léxico.
2. La relación entre los rasgos léxico-semánticos de una palabra y su combinación con otras palabras.
3. La búsqueda de grupos de palabras (*tipos lexicográficos* de Apresjan, 2000,

3. En Atkins *et al.* (1994) y Zgusta (2006: § 2.2) se trata el desarrollo histórico del estudio del léxico en relación con la lexicografía y la lexicología computacional.

4. El inicio de la transición a la «nueva lexicografía» se suele hacer coincidir con los trabajos clásicos de 1971 de L. Zgusta, J. Rey-Debove, y J. y C. Dubois. Cfr. Pascual (2006) para una reflexión sobre los cambios que ha experimentado la disciplina lexicográfica en el siglo xx.

5. Todo diccionario posee determinado efecto regulador por ser una obra de referencia. Por un lado, incluso en los diccionarios descriptivos los datos léxicos pasan por un proceso de selección; por otro, lo favorecen las marcas que aluden al estatus de las palabras (coloquial, dialectal, anticuado, etc.) (Zgusta, 2006: 195).

o *clases léxicas* de Bosque, 2004a) con al menos una propiedad compartida, que sean sensibles hacia las mismas reglas lingüísticas y tengan que estar descritos de manera uniforme en el diccionario.

## 2. «Teorización» de los diccionarios

### 2.1. LOS POSTULADOS DE LA LINGÜÍSTICA TEÓRICA EN LOS DICCIONARIOS

Uno de los resultados de la reflexión lexicográfica y lexicológica de nuestra época ha sido la conclusión de que, para poder reflejar adecuadamente el sistema léxico, un diccionario no puede basarse solo en la intuición lingüística de quien lo elabora; esta intuición se debe apoyar en un modelo de organización léxica. A día de hoy no existe un modelo léxico comúnmente aceptado; más bien las distintas teorías desarrollan su propia propuesta (cfr. la tercera parte de este libro). A pesar de ello, como vamos a comprobar, los análisis de la estructura y el funcionamiento del léxico no suelen ser incompatibles: a menudo se recurre al mismo tipo de rasgos léxicos (para explicar el comportamiento sintáctico de distintos grupos de palabras, por ejemplo). Como apuntan Zaenen y Engdahl (1994), un problema mucho mayor es que para la codificación de gran parte de la información léxica que tienen que manejar los diccionarios no existe ningún tipo de propuesta teórica, o bien existen propuestas incompletas.

La aplicación lexicográfica de una teoría está sujeta a parámetros de varios tipos, como su verificabilidad y la solidez de sus planteamientos (el lexicógrafo tiene que seleccionar los resultados de la investigación más aceptados y que más perduren en el tiempo), y la clase de diccionario que se quiere elaborar en función de las necesidades del potencial usuario. En este sentido, Geeraerts (1985) destaca tres tipos básicos de diccionarios: *diccionario enciclopédico* o *terminológico*, que proporciona información científica; *diccionario de lengua*, que contiene datos lingüísticos y enciclopédicos en proporción variable, y cuya función es facilitar la comunicación y difundir la norma lingüística; y *diccionario lingüístico* o *teórico*, que sirve para el propósito científico de la descripción del lenguaje y su evolución. Los diccionarios lingüísticos son los que más nos van a interesar aquí (aunque son también los menos numerosos porque van destinados principalmente a investigadores): estos sí parten de un modelo teórico que condiciona de una forma determinada la presentación de los datos, y poseen como función complementaria la de servir de fuente de enriquecimiento para otros repertorios lexicográficos (cfr. Lorente Casafont, 1996).

A la hora de hablar del contenido lingüístico de los diccionarios, cabe mencionar la diferencia entre los diccionarios «para humanos» y «para máquinas». Los diccionarios usuales cuentan con la competencia del hablante de alguna de las lenguas y se centran en los aspectos que pueden resultar problemáticos para su uso. Como el funcionamiento de los ordenadores se basa en el conocimiento almacenado, los diccionarios propiamente electrónicos tienen que representar las piezas léxicas de forma necesariamente más detallada, incluyendo la información lingüística que los hablantes dominan de forma inconsciente (cfr. Jiménez Ríos, 2000) y también ma-

yores cantidades de conocimiento enciclopédico que los diccionarios convencionales (dependiendo del propósito de la aplicación informática concreta).

El grado de repercusión de las teorías léxicas en los diccionarios se ve limitado por su naturaleza de repertorios que «no pueden registrar más que lo lingüísticamente realizado y, por lo general, con algún retraso en el tiempo y de forma incompleta» (Haensch *et al.*, 1982: 359). Al no ser obras de referencia atemporal, no pueden predecir la evolución de los rasgos semánticos y sintácticos de las palabras ni la aparición de unidades léxicas nuevas. Lo que sí pueden llegar a hacer con ayuda de la lingüística teórica es esclarecer las relaciones entre diferentes unidades léxicas y describir el cauce que su evolución ha seguido dentro del alcance temporal del diccionario.

## 2.2. INTRODUCCIÓN DEL ORDENADOR EN EL MÉTODO DE TRABAJO

El encuentro entre el diccionario y el ordenador es inevitable: los diccionarios almacenan grandes cantidades de datos lingüísticos y el ordenador es el procesador universal de información (Knowles, 1989). Los diccionarios se usan en aplicaciones informáticas relacionadas con la investigación y el tratamiento del lenguaje (traducción asistida por ordenador, investigación en inteligencia artificial, etc.) así como con fines puramente utilitarios (en aplicaciones ofimáticas, por ejemplo).

La aplicación del ordenador a la labor lexicográfica ofrece *mejoras cuantitativas* (una mayor eficacia, rapidez y comodidad de consulta, inclusión de un mayor número de datos, etc.) y *cualitativas*, que suponen un cambio más profundo, como la repercusión en la base teórico-lingüística del diccionario y la modificación del método de trabajo lexicográfico (estas son las que más interesan a nivel teórico).<sup>6</sup>

La utilización de procedimientos informáticos exige que la descripción lexicográfica sea coherente y uniforme (Pascual, 1996): solo puede y debe incluir lo que está inicialmente previsto por el programa (en términos de atributos de las bases de datos, p. ej.). El procesamiento informático ayuda a mantener la coherencia interna de un diccionario señalando sobre todo cuando esta se infringe: cuando aparece una fórmula definitoria no establecida previamente o se sobrepasa el vocabulario definidor, cuando no se rellenan todos los campos en la microestructura, etc.

En cuanto al método de trabajo, quizá el avance más significativo ha sido el paso de la redacción secuencial de artículos a la estructura de una base de datos

6. Ambos tipos de mejoras se corresponden con la existencia de dos clases de diccionarios en soporte electrónico (cfr. Rafel, 2005, Gelpí Arroyo, 2003 y Fernández-Pampillón Cesteros y Matesanz del Barrio, 2006):

- *Diccionarios digitalizados o automatizados*: versiones de diccionarios convencionales en soporte digital o diccionarios elaborados en formato electrónico, pero con características análogas a las de los diccionarios comunes, que suelen incorporar lematizador automático (véase el capítulo 4 de la cuarta parte de este volumen) y permiten búsquedas básicas. Estos diccionarios presentan los mismos problemas que los diccionarios en papel.
- *Diccionarios electrónicos*: son aplicaciones informáticas que integran una base de datos léxica rica en formatos y contenidos (así como herramientas para su consulta, visualización, impresión, etc.). Estos se suelen utilizar como fuentes de información en aplicaciones de procesamiento del lenguaje.

léxica (BDL). Los campos de una BDL, diferenciados según el tipo de información recogida,<sup>7</sup> multiplican las posibilidades de acceso a los datos, de ahí su carácter *multifuncional*: cada usuario, en cada consulta específica, puede extraer la porción de conocimiento léxico que necesite y que satisfaga sus exigencias de contenido y de formato.

Para que la representación computerizada del sistema léxico sea posible, aparte de los recursos computacionales necesitamos disponer de un modelo de organización léxica: los elementos de las entradas y las relaciones entre ellos serán las que el lexicólogo haya establecido y formalizado. Insistamos en esta puntualización: los procedimientos informáticos no pueden sustituir la reflexión sobre la lengua en cualquier área de lingüística aplicada incluida la lexicografía, solo la facilitan.<sup>8</sup>

En Calzolari (1994) se expone precisamente una visión de una BDL relacional flexible y dinámica. La autora incluye en una representación sistemática del léxico los rasgos de las piezas léxicas (categoría sintáctica, paradigmas flexivos, etc.), y las relaciones entre palabras: la hiperonimia y otras relaciones semánticas ('especie de', 'tipo de', 'parte-todo', 'conjunto de'), relaciones de modificación ('usado para', 'provisión de', 'en forma de'), relaciones de derivación morfológica; relaciones argumentales (papeles temáticos y otras); sinonimia y antonimia; restricciones de selección (del tipo 'dícese de', que se establecen entre los nombres y los adjetivos); campos semánticos, etc.<sup>9</sup> Una BDL así permite acceder a los datos a través de las palabras, sus rasgos o el tipo de relación que estas mantienen posibilitando también desta-

7. La información codificada en BDL avanzadas incluye desde datos estadísticos de uso y claves para resolver crucigramas hasta fragmentos de videos y características tomadas de las teorías de descripción lingüística, a las que volveremos en los siguientes apartados, como las funciones léxicas de Mel'čuk, los primitivos semánticos y los patrones actanciales (cfr. Knowles, 1989: § 5; Zampolli, 1991 y Lenders, 2002). Un ejemplo de organización de entradas léxicas en forma de BDL puede ser el *Diccionari descriptiu de la llengua catalana* (proyecto del Institut d'Estudis Catalans, dirigido por J. Rafel). Consúltese su estructura, muy rica en cuanto al tipo de datos lingüísticos proporcionados, en [dcc.iecat.net/ddlc/](http://dcc.iecat.net/ddlc/).

8. Hablando de la «distribución de trabajo» entre los métodos computacionales en general y los modelos cognitivos o mentalistas del lenguaje, se puede decir que ambos se complementan. Por un lado, aunque los métodos estadísticos son capaces de producir cierto tipo de generalizaciones sobre la base de los datos observados, no pueden sustituir la elaboración de modelos cognitivos, porque se acercan a la competencia humana mediante el análisis de su producto final, los datos lingüísticos, y no dan respuestas sobre su naturaleza y funcionamiento. En este sentido, Bosque (2001) y Haensch *et al.* (1982) critican el tratamiento de las colocaciones solo en términos de frecuencia de coaparición, el único criterio seguido en algunos diccionarios: *The BBI Dictionary of English Word Combinations*, *Collins Cobuild English Collocations* y *LTP Dictionary of Selected Collocations*. Según Haensch *et al.* (1982: 252), «Si se renuncia a los conocimientos de la lingüística, se obtienen únicamente colecciones de colocaciones según su frecuencia, que, en el mejor de los casos, representan cierto progreso (abarcan más colocaciones que los diccionarios de estilo basados en procedimientos de selección más subjetivos) frente a los trabajos de hace 60 años, pero que, aparte de esto, no responden al estado actual de la investigación».

Por otro lado, los métodos probabilísticos permiten emular con altas cuotas de éxito la capacidad lingüística humana en tareas de procesamiento del lenguaje (Lavid, 2005: § 2.4), lo que demuestra que los datos estadísticos han de ser tenidos en cuenta por los modelos cognitivos. Los lingüistas no deberían ignorar el hecho de que unas palabras tienen mayor peso en la lengua en términos de distribución y frecuencia que otras (cfr. las conclusiones del estudio estadístico del corpus CUMBRE en Almela *et al.*, 2005).

9. Muchos de estos datos están contenidos en los diccionarios convencionales; de hecho, se aprovechan en los lexicones computacionales una vez extraídos y formalizados.

car sublexicones más específicos según el parámetro de acceso elegido (en términos lexicológicos, estos serían grupos dinámicos de palabras con una o varias propiedades compartidas; la misma palabra sería miembro de varios grupos según el criterio dado).<sup>10</sup> Una entrada léxica de este tipo ya no sería una unidad cerrada en sí misma, sino un punto de convergencia de todas las relaciones de una pieza léxica.

Hoy en día resulta impensable la recogida y preparación de datos para un diccionario sin la ayuda de corpus informatizados: estos contienen datos reales de la lengua y constituyen por tanto una base objetiva para su estudio, a diferencia de métodos de análisis más subjetivos (como los basados en la intuición del lingüista).<sup>11</sup> Un corpus lingüístico es:

una colección de textos orales o escritos de una lengua, que han sido seleccionados a partir de unos criterios lingüísticos explícitos, pasados a soporte electrónico y mínimamente procesados,<sup>12</sup> y que se utilizan como muestra representativa de esta para el abastecimiento de datos destinados a su estudio sistemático (Saurí, 2003).

Los corpus combinados con técnicas estadísticas se utilizan para determinar la distribución de términos específicos (neologismos, tecnicismos, etc.) en diferentes niveles de lenguaje; para estudiar las restricciones de coaparición entre palabras; para desambiguarlas gramatical y semánticamente, etc. (Köhler, 2002; Lavid, 2005). En la lexicografía, los datos de los corpus ayudan, por ejemplo, a determinar qué palabras y acepciones deben entrar en el diccionario,<sup>13</sup> a establecer el orden de las acepciones (a menudo se da prioridad a las más usadas), a detectar patrones de comportamiento sintáctico y combinaciones estables de palabras y a seleccionar ejemplos de uso.

### 3. Reflejo de las teorías léxicas en los nuevos diccionarios

#### 3.1. PROBLEMAS GENERALES DEL TRATAMIENTO LEXICOGRÁFICO DE LAS PIEZAS LÉXICAS

Antes de hablar sobre como se ven reflejadas diversas teorías léxicas en los diccionarios, nos centraremos en algunos problemas de presentación de la información

10. Así, podemos destacar grupos semánticos de verbos que comparten la estructura argumental y la red temática cotejando las fórmulas hiperonímicas de las definiciones: *rodar* y *volar* se definen como un tipo de movimiento ('moverse' es su hiperónimo o género próximo) y poseen la misma estructura argumental.

11. COBUILD (1987) es el primer diccionario basado en un corpus electrónico (*Bank of English*).

La ley de Zipf, que relaciona la frecuencia de aparición de una palabra con el tamaño del corpus, proporciona una justificación estadística para el uso de corpus: cuanto más reducido sea, mayor será la frecuencia de las palabras en los primeros rangos. Por tanto, cuanto más amplio es el corpus más posibilidad hay de dar con los vocablos que menor frecuencia de uso poseen.

12. Es decir, anotados o provistos de información metalingüística extratextual (fonética, morfológica, sintáctica, semántica y pragmática) que facilita su consulta. Para una revisión de las técnicas de anotación, véase Lavid (2005) y los trabajos allí citados. Para más información sobre los corpus, cfr. la web del grupo EAGLES: [www.ilc.cnr.it/EAGLES/home.html](http://www.ilc.cnr.it/EAGLES/home.html).

13. Saber qué palabras conforman el vocabulario básico de un idioma es importante para elaborar diccionarios escolares y diccionarios para estudiantes extranjeros de un idioma. Además, este *modus operandi* permite evitar la aparición de voces anticuadas, desusadas o inexistentes.

semántica y sintáctica en un artículo lexicográfico. Los lexicógrafos siempre han dado importancia a la parte semántico-conceptual de la definición, pero hay un progresivo aumento de interés por la información sintáctica, relativamente reciente, que tiene que ver con que la palabra ya no es el único foco de atención, sino que se suele considerar como parte de constituyentes mayores: el sintagma, la oración y el texto.

### 3.1.1. *Componente léxico-semántico de la definición*

Parece haber un acuerdo más o menos general sobre el hecho de que una definición tiene que representar un análisis léxico-semántico del lema (Porto Dapena, 2002: 275; Mederos, 1994; Martín Mingorance, 1994). Si se asume que el léxico es un sistema de estructuras interrelacionadas, la mejor manera de definir una palabra es aludiendo al tipo de estructura en la que se integra, que serviría como guía o patrón natural de la definición (Mederos, 1994). Sin embargo, a la hora de hablar de las estructuras léxicas nos encontramos con algunas dificultades.

De entrada, no parece que el léxico en su totalidad esté estructurado en función de un único criterio, más bien evoca una red multidimensional. La conveniencia de dividir los criterios de estructuración entre los propiamente lingüísticos y los enciclopédicos (el conocimiento del mundo) es ciertamente una cuestión de principio sobre la que no existe acuerdo general.<sup>14</sup> Aquí asumiremos que esta distinción sí es relevante y que las nociones del lenguaje solo representan una parte del dominio conceptual en el que almacenamos la información. El léxico es el portador del conocimiento extralingüístico en tanto que relaciona el contenido conceptual con una forma fónica.

Las diferentes teorías y subdisciplinas lingüísticas suelen adoptar una visión propia acerca del grado de «enciclopedismo» admisible. La lexicografía tiene que aludir a lo extralingüístico: por ejemplo, para explicar qué significa un sustantivo concreto, a menudo hay que describir el objeto denotado.<sup>15</sup> El conocimiento del mundo es asimismo un elemento habitual en las representaciones de inteligencia artificial y lingüística computacional (Lavid, 2005: § 2.3.2). Las teorías léxicas, como iremos viendo, tienden a recurrir a informaciones cada vez más específicas, pero esta «expansión» tiene que estar justificada por razones independientes, como la imposibilidad de deducir ciertas oposiciones léxicas (p. ej., continuo-discontinuo, individual-colectivo) del conocimiento del mundo, o su repercusión en la sintaxis u otros niveles del lenguaje (Bosque, 1982).

Otra fuente de controversia concierne al metalenguaje utilizado en las definiciones: en la mayoría de los diccionarios este proviene de la lengua analizada, con algunas consecuencias problemáticas como la circularidad (cuyo caso extremo se da cuando dos palabras se definen una en términos de la otra). La otra opción es utilizar

14. Uno de los modelos que más ha contribuido a que esta distinción se cuestione es la lingüística cognitiva. Peeters (2000) resume de manera clara y abarcadora las diferentes posturas adoptadas sobre este tema, incluidos los modelos de la sección § 3.2 de este capítulo.

15. La distinción entre lo lingüístico y lo enciclopédico en los diccionarios es más una cuestión de grado que una división estricta (cfr. Rafel, 2005), e influye, por ejemplo, en el carácter de las definiciones (son más específicas las de un diccionario enciclopédico que las de uno de lengua) y en la nomenclatura (los nombres propios no se incluyen en los diccionarios de lengua por no tener contenido lingüístico; a su vez, las categorías gramaticales, de escaso contenido conceptual, no aparecen en las enciclopedias).

un lenguaje formal (que coincida o no con la lengua objeto), con el riesgo de que las definiciones no sean entendidas por el lector.

Para examinar más de cerca estos y otros problemas, tomemos como punto de partida la clase de definición más extendida, la *hiperonímica* (llamada también *aristotélica* o *definición por género próximo y diferencia específica*). En esta definición, un vocablo se identifica con su género o hiperónimo (*coche* y *autobús* se caracterizan como hipónimos de ‘vehículo automóvil’ en el DRAE), y luego se distingue de otros miembros del mismo género (co-hipónimos) mediante la enumeración de todas las diferencias sistemáticas («de tamaño pequeño o mediano» para *coche* o «[destinado al] transporte público» para *autobús*). Este tipo de definición plantea algunas dificultades:

- Parte de una clasificación establecida para el estudio de los conceptos y fenómenos del mundo real (como en la clasificación de organismos), no para el de las palabras, por eso funciona bien con expresiones nominales de carácter terminológico o denotativo, pero no con el vocabulario más abstracto expresado por otras categorías léxicas.
- Como todas las taxonomías,<sup>16</sup> está basada en un solo criterio, la inclusión hiperonímica. Este puede ser un principio adecuado para estructurar ciertos campos semánticos, pero obviamente no agota todas las posibles relaciones entre palabras.
- Si se lleva a cabo una serie de definiciones en las que cada palabra es definida mediante el género, un vocablo menos específico y más frecuente (Weinreich, 1975), al final de la cadena tendremos términos muy abstractos que solo se pueden definir de forma circular, mediante otro término general (*cosa*, *objeto*, etc.) o de forma icónica (una ilustración).<sup>17</sup>

Los aspectos problemáticos que tiene que tratar una teoría de semántica léxica aplicada a la definición lexicográfica son, como vamos viendo, tres:

1. El tratamiento de la información enciclopédica frente a la lingüística: (no) distinción entre ambas, criterios para su distinción, (no) inclusión de información enciclopédica.
2. La elección de qué estructuras léxicas se destacan y de acuerdo con qué criterios.
3. El estatus del vocabulario definidor: qué tipo de elementos básicos y no analizables se utilizan para definir la totalidad del vocabulario.

Las dos últimas cuestiones tienen que ver con como se da cuenta de la composicionalidad del significado léxico. Un sistema de representación es composicional

16. Un término relacionado con el de *taxonomía* es *ontología*, entendida como jerarquía de dominios semánticos. Entre los tipos ontológicos de grado de generalidad más alto se incluyen: cosa, cualidad, cantidad, lugar, tiempo, estado, proceso, evento, acción, relación, manera, etc. (cfr. Cruse, 2000: 47).

17. Cruse (1986: § 6) señala que las taxonomías léxicas no suelen superar cinco niveles: *planta* (reino) - *arbusto* (forma de vida) - *rosa* (género) - *híbrido de té* (especie) - *cover girl* (variedad).



cuando 1) contiene tanto símbolos sintácticos y semánticamente primitivos como símbolos complejos; y 2) cuando estos heredan sus propiedades de aquellos. Muchas de las teorías de § 3.2 llevan a cabo una descomposición del significado para detectar los elementos primitivos y recurrentes que producen y permiten interpretar los elementos complejos.<sup>18</sup>

### 3.1.2. *Componente sintáctico de la definición*

Aun bajo el supuesto tradicional de que el objeto de los diccionarios son las palabras y no la gramática, los diccionarios siempre han aludido a los rasgos sintácticos básicos de las piezas léxicas (así, la categoría léxica se utiliza para separar y ordenar acepciones). Cuando surgió la necesidad de explicar, además, las propiedades combinatorias de las palabras (p. ej., para reflejar en los diccionarios didácticos las restricciones que intervienen en los patrones oracionales), se tuvo que recurrir a especificaciones más sutiles, como, en términos modernos, las de *valencia*, *subcategorización* y *selección*, según veremos más adelante en distintos ejemplos.

Muchos diccionarios de lengua incluyen indicaciones sintácticas<sup>19</sup> en el *contorno definicional*, que completa y especifica el contenido del término definido, y por tanto lo diferencia del de otras palabras o de otras acepciones de la misma palabra. Como la inclusión de esta información hace a menudo que el enunciado definitorio y el vocablo definido dejen de ser mutuamente sustituibles en el contexto de una oración (las oraciones de (2) pueden servir de ejemplo), es necesario delimitar la parte propiamente semántica y el contorno. Por ejemplo, el *Diccionario Salamanca* marca el sujeto con paréntesis angulares y los complementos con corchetes, como en (2a). En la planta del *Diccionario Coruña* (proyecto de la Universidade da Coruña, dirigido por J. A. Porto Dapena) se especifica que el contorno se registrará entre corchetes, con los elementos no obligatorios incluidos entre paréntesis angulares (se aludirá, además, a la función sintáctica de cada elemento, como en [2b]).

- (2) a. Enfermar v.tr. Causar <una cosa> una enfermedad a [una persona, un animal o una planta]  
 b. Satisfacer 1. tr. [~ alguien (suj.) un deseo, necesidad, condición u obligación (od.) <a otra persona o a sí mismo> (oi.) <con algo> (cp.)]. Hacerlo desaparecer total o parcialmente [...].

Un buen ejemplo de tratamiento de la información sintáctica en los diccionarios son los diccionarios de valencias. La noción de valencia parte de la idea de que el verbo, como elemento central de la oración, «atrae» determinado tipo y número de argumentos, sus actantes. Algunos ejemplos del género son: el *Wörterbuch zur Va-*

18. Algunos de los problemas que tienen que afrontar estas teorías son: justificar el estatus y el número admisible de los primitivos y decidir cuál es el límite de la descomposición, así como dar cuenta de casos de lexicalización en los que la composicionalidad no se cumple o está restringida (locuciones, colocaciones, etc.). Para un análisis detallado de la descomposición léxica, cfr. Cruse (2000: § 13).

19. En el ámbito hispánico, el primer referente del estudio sistemático de las propiedades sintagmáticas de las palabras es el *Diccionario de construcción y régimen* (Rufino José Cuervo / Instituto Caro y Cuervo).

*lenz und Distribution Deutscher Verben* de Helbig y Schenkel, los diccionarios de valencia y distribución de adjetivos, sustantivos y verbos alemanes de Sommerfeldt y Schreiber y el *Valency Dictionary of English* de Herbst.<sup>20</sup>

Nos centraremos en los proyectos que contemplan los datos sintácticos en relación con los léxico-semánticos. La distribución del peso de la interpretación contextual de las palabras será diferente en diferentes modelos: en algunos la sintaxis se deriva de la semántica de manera determinista (como en la clasificación verbal de Levin, y en menor medida en el Modelo Sentido-Texto, § 3.2.4 y § 3.2.2), en otros el significado léxico es completado por el de la estructura sintáctica (como en el Lexicón Generativo, §3.2.3 y en la Gramática de Construcciones). También serán distintos los mecanismos de interacción entre ambos tipos de rasgos (a través de las funciones léxicas, la estructura argumental o actancial, los roles semánticos, etc.). Destacaremos aparte los proyectos que van más allá de la información sintáctica, dando un salto al dominio pragmático para explicar la interpretación de las piezas léxicas en el contexto (como FrameNet, § 3.2.5).

### 3.2. LAS TEORÍAS LÉXICAS EN LOS NUEVOS DICCIONARIOS

#### 3.2.1. *Teorías de rasgos: la escuela funcional-estructuralista y la teoría de prototipos*

La aplicación al ámbito semántico de la noción de *rasgo distintivo* es una de las principales contribuciones de la escuela funcional-estructuralista al estudio sistematizado del léxico. En el plano paradigmático, los *semas* (rasgos semánticos distintivos mínimos) permiten por un lado caracterizar el denominador común de un grupo de elementos y por otro diferenciarlos entre ellos. Así, los verbos *ejecutar*, *asesinar* y *masacrar* tienen como núcleo común ‘quitar la vida, matar’<sup>21</sup> y se diferencian entre ellos por los siguientes elementos de significado: ‘a una persona condenada en cumplimiento de la sentencia de muerte’, ‘a una persona con premeditación’ y ‘a personas de forma colectiva’, respectivamente. Un conjunto de lexemas que, como los que hemos mencionado, «resulta de la repartición de un contenido léxico continuo entre diferentes unidades dadas en la lengua como palabras y que se oponen de manera inmediata unas a otras por medio de rasgos distintivos mínimos» constituye, en términos de Coseriu (1977: 146), un *campo léxico*.<sup>22</sup>

20. Entre los proyectos incipientes cabe señalar el *Diccionario contrastivo de valencias español-alemán*, de la Universidad de Santiago de Compostela, que parte de la Base de Datos Sintácticos del español actual (BDS), desarrollada en la misma universidad bajo la dirección de Guillermo Rojo. La BDS ha dado lugar a varios proyectos de interés lexicológico (como ADESSE, § 3.2.5). Esta base de datos contiene el análisis manual de 160.000 cláusulas de la parte contemporánea del archivo ARTHUS, e incluye información sobre la cláusula como conjunto y sobre las funciones sintácticas que se dan dentro de ella (cfr. [www.bds.usc.es](http://www.bds.usc.es)).

21. *Matar* sería el *archilexema* del campo léxico formado por *ejecutar*, *asesinar*, *masacrar*, etc. El *archilexema* integra los rasgos semánticos compartidos por los miembros del mismo campo léxico.

22. Para la revisión de la teoría de campos léxicos, cfr. Geckeler (1976, 2002) y el capítulo 1 de la tercera parte de este volumen.

El paralelismo entre el género próximo de la definición aristotélica y el archilexema por un lado, y entre la diferencia específica y los rasgos distintivos por otro, parece obvio. En este sentido, el método estructural proporciona al menos en principio una base teórico-lingüística para la estructuración de la definición (Escobedo Rodríguez, 2000), y los campos léxicos podrían ser una forma de organizar las entradas léxicas. Sin embargo, un diccionario así nunca se ha llegado a crear, y de existir tendría que afrontar algunas de las dificultades que suscita el método estructural. Por una parte, no está demostrado que todo el léxico se pueda estructurar en campos. Por otra, el método estructural excluye una parte importante del vocabulario, en concreto la terminología y las nomenclaturas, considerando que solo se pueden definir y estructurar desde la realidad que representan, no desde el lenguaje: los términos científicos y los tecnicismos significan lo que designan, no forman parte de las estructuras léxicas del mismo modo que el resto de palabras. Los diccionarios onomasiológicos (analógicos e ideológicos) se acercan bastante a la idea de la clasificación semántica del vocabulario en campos, pero se construyen a partir de conceptos o ideas, asignando después estos a una serie de palabras que a su vez se relacionan entre sí a través de algún tipo de afinidad asociativa no estrictamente lingüística.<sup>23</sup>

Como muchos modelos simbólicos altamente restringidos, la semántica de rasgos distintivos posee un poder explicativo superior a su poder descriptivo.<sup>24</sup> No es capaz de manejar casos marginales: *pingüino* no posee el rasgo esencial ‘volar’, pero se define como pájaro. Tampoco da cuenta de la gradualidad de propiedades: si *león* se describe en el *Diccionario Salamanca* como «buen depredador», será que hay depredadores menos «buenos» o menos representativos de la categoría ‘depredador’.<sup>25</sup> Nótese que no hablamos del conocimiento científico (depredador como

23. Según Shcherba (1940), la elaboración de los diccionarios onomasiológicos presenta dos tipos de dificultades: 1) «hace falta tener un listado completo y muy exacto de palabras como conceptos para una lengua dada, y para compilar un listado así hay que describir con precisión todos los significados de las palabras en diccionarios convencionales»; 2) habría que dar respuesta a la pregunta de «cómo clasificar las palabras en tanto que conceptos». Los diccionarios onomasiológicos del español más citados son el *Diccionario ideológico* de Julio Casares, y el más reciente *Diccionario ideológico de la lengua española Vox*. Algunos ejemplos del ámbito anglosajón son el *Longman Language Activator*, *Longman Lexicon of Contemporary English* y los diccionarios de la serie *Cambridge Word Selector*.

24. Las ventajas y los inconvenientes de la teoría en cuestión se explican en parte por su adscripción al modelo de categorización aristotélico (modelo de condiciones necesarias y suficientes). En este modelo:

- Las unidades de análisis (categorías, unidades, rasgos) se consideran discretas (claramente delimitadas).
- La pertenencia a una categoría se define mediante condiciones necesarias y suficientes: una unidad que cumple todas esas y sólo esas condiciones se considera miembro de la categoría, y en caso contrario se queda fuera. Estas condiciones en el lenguaje serían los rasgos (contrastivos o definitorios), que forman parte de la definición de la palabra (de manera enumerativa, no estructurada) y se consideran lingüísticamente relevantes; el resto de los rasgos pertenecen al saber extralingüístico del hablante.
- Como todos los miembros de una categoría dada cumplen las mismas condiciones de pertenencia, su estatus es el mismo (no hay miembros más idóneos que otros).

25. Otras de las palabras utilizadas en las definiciones lexicográficas para aludir a entidades y propiedades más o menos representativas son: *especialmente*, *típicamente*, *normalmente*, *a menudo*, etc.

animal que da caza y muerte a su presa), sino de lo que los enfoques cognitivos llaman «el cuadro *naif* del mundo», la realidad tal y como la percibe el hablante, al margen de la realidad «objetiva».

En este sentido, la lingüística cognitiva constituye una contrapartida al modelo clásico estructural. La semántica cognitiva, que estudia el reflejo lingüístico del dominio cognitivo general, se basa en la noción de *prototipo*, el punto clave de la teoría de la categorización de Eleanor Rosch (Rosch y Mervis, 1975). El prototipo se interpreta como el representante más idóneo de una categoría (*manzana* para la categoría ‘fruta’), como un conjunto de propiedades ‘típicas’ (no exactamente ‘necesarias’) de una categoría (p. ej., ‘ser capaz de volar’, ‘ser ovíparo’, ‘tener plumas’, ‘tener alas’ y ‘tener pico’ para la categoría *pájaro*), o bien como el sentido básico del que se derivan los demás.<sup>26</sup>

La semántica de prototipos posee un aparato descriptor muy potente dado que:

- Permite que los conceptos tengan límites borrosos: si ningún rasgo es necesario sino tan solo típico, obtenemos definiciones muy flexibles que llevan a que el estatus de los elementos que abarca un concepto sea desigual, y se defina bien como más o menos central o bien en términos de *parecido de familia*.<sup>27</sup> Este tipo de estructuración permite tratar eficazmente la polisemia: el prototipo en un diccionario es el significado principal y las demás acepciones se derivan de él mediante un proceso de desplazamiento.<sup>28</sup>
- Amplía la noción de lo ‘lingüísticamente pertinente’: como ninguna propiedad es necesaria ni suficiente, hace falta recurrir a descripciones semánticas muy ricas. No basta con indicar lo que diferencia las palabras, hay que definir las positivamente, enumerando todos sus rasgos. Esto, unido a que la única restricción para clasificar un rasgo como lingüístico (en las corrientes que mantienen la dicotomía lingüístico-conceptual) es que este forme parte del conocimiento compartido por los hablantes (Wierzbicka, 1985), hace las definiciones muy extensas y detalladas.
- Recurre al código visual (ejemplificación ostensiva en los diccionarios) porque un prototipo puede ser una imagen o un esquema (cfr. *Longman Dictionary of Contemporary English* o *Oxford Advanced Learner’s Dictionary*).

Otro rasgo importante de la semántica cognitiva es que resalta la dimensión antropomórfica de la significación en el lenguaje; mencionaremos dos de sus aspectos. El primero es el *estereotipo* (Putnam, 1975), un prototipo «atado» a una lengua concreta y en especial a su dimensión sociocultural. Luis Fernando Lara, uno de los lexicógrafos que propugna su aplicación en los diccionarios, lo define como «un núcleo de significación de origen cognoscitivo, pero estabilizado por el consenso social mediante la enseñanza y la educación, que ofrece una organización primaria

26. Cfr. Kleiber (1995), Cruse (2000), y el capítulo 3 de la segunda parte y el capítulo 3 de la tercera parte de este volumen.

27. La relación de parecido de familia implica que un rasgo ha de ser compartido por al menos dos miembros de una categoría: AB-BC-CD-DE, etc.

28. Véase Geeraerts (2001) para un análisis de la polisemia en términos prototípicos, que representa una alternativa multidimensional a la estructuración lineal de las acepciones en los diccionarios.

del sentido en su comunidad» (Lara, 1992: 9; 1989). Sirvan de ejemplo dos entradas abreviadas del *Diccionario del español usual en México* (obsérvense especialmente los elementos que aluden al entorno cultural y los rasgos funcionales):

- (3) *a.* **mestizaje** s m **1** Mezcla de razas distintas, y en América especialmente el cruzamiento entre indígenas y españoles: *el mestizaje en la Nueva España* [...] *b.* **palma**<sup>1</sup> s f **1** Planta angiosperma monocotiledónea de diversas variedades y especies, que se caracteriza por tener el tallo leñoso, cilíndrico y recto, liso, áspero o espinoso, con un penacho de hojas muy grandes [...]. De algunas especies se aprovecha el tallo para construcciones rústicas y de otras, el aceite que producen; es propia de climas tropicales, aunque algunas variedades se pueden cultivar en climas templados **2** pl Familia que forman estas plantas **3** [...] **4** Hoja de estas plantas que se emplea en el tejido de sombreros, canastas, etc. o en la construcción de techos y muros.

El segundo aspecto antropomórfico es que el hombre representa en sí mismo un criterio estructurador de conceptos y palabras relacionadas. La metáfora del cuerpo humano es un ejemplo muy citado:  *cuello de la botella, cabeza/boca de un río, puñado de manifestantes*. Wierzbicka (1985: § 5.7) estudia casos más complejos, p. ej., el papel de la función (de nuevo, noción antropocéntrica) en la estructuración del significado. Por su importancia, la información sobre la función se incluye en primer lugar: es la que dota de estabilidad al concepto y también ayuda a entender, en el caso de los términos que representan objetos creados por el hombre, el resto de su descripción. Véase en (4) el principio de la definición de *cup* ‘taza’ (Wierzbicka, 1985: 33):

- (4) CUPS  
A KIND OF THING THAT PEOPLE MAKE  
IMAGINING THINGS OF THIS KIND PEOPLE WOULD SAY THESE  
THINGS ABOUT THEM:  
they are made for people to use repeatedly for drinking hot liquids from  
such as tea or coffee  
one person from one thing of this kind  
being able to put them down on something else

Tazas

Una cosa que hacen las personas

Imaginando cosas de este tipo las personas dirían estas cosas sobre ellas:  
están hechas para que las personas las usen reiteradamente para beber líquidos calientes como té o café

una persona de una cosa de este tipo

siendo capaz de ponerlas sobre algo más

[La traducción es mía, O. B.]

En toda la entrada (cfr. el trabajo original citado *supra*), la información sobre la función determina el tipo y la ordenación de los otros datos: primero el material (sugiriendo que es duro y fácil de lavar, resistente al contacto con el líquido), luego la apariencia (forma específica que garantiza la estabilidad y fácil manejo, proporciones

adecuadas para mantener el líquido caliente) y el tamaño (justo para contener cierta cantidad de líquido y para poder ser levantado con una mano).

El estudio de Wierzbicka es interesante también porque desarrolla una propuesta de unidades de análisis del significado. Su *metalenguaje semántico natural* está compuesto por primitivos semánticos, conceptos básicos indefinibles que se utilizan para describir la totalidad del vocabulario y que son el producto final del análisis reduccionista. Los primitivos son universales en cuanto entidades conceptuales, y a la vez específicos para la lengua que los plasma.<sup>29</sup> La lista de primitivos y cuasi-primitivos para el inglés incluye, entre otros: *I, you, something, somebody, thus, everything, word, two, to say, (not) to want, to feel, to think, to know, to be able, to do, to happen, to move, to see, good, bad, similar, far, near, the same, where, when, after, because of*.<sup>30</sup>

Véase otro ejemplo de aplicación de los primitivos semánticos del *English Speech Act Verbs. A semantic dictionary* (Wierzbicka, 1987):

(5) REPROVE ('reprender, recriminar')

Meaning

I say: you have done something bad

I assume you understand that you shouldn't do such things

I assume you understand that I am someone who can say this to you

I say this because I want to cause you not to do such things<sup>31</sup>

Significado

Yo digo: tú has hecho algo mal

Yo asumo que tú entiendes que tú no deberías hacer este tipo de cosas

Yo asumo que tú entiendes que yo soy alguien que puede decirte esto

Yo lo digo porque quiero que tú no hagas este tipo de cosas

[La traducción es mía, O. B.]

La ventaja de los primitivos semánticos es que cualquiera puede entenderlos. Sin embargo, el análisis reduccionista combinado con la exhaustividad de las definiciones hace que las entradas se vuelvan larguísimas y por eso no mucho más fáciles de procesar que las definiciones basadas en términos más formales. Además, queda abierta la cuestión de si son realmente primitivos los elementos postulados: el hecho de que constituyan el vocabulario básico y no definible de un idioma (al menos mediante otras palabras del mismo idioma) no quiere decir que las nociones que representen sean conceptual y lingüísticamente sencillas; su complejidad pone en duda su universalidad. Otro defecto achacable a esta teoría es la ausencia de información sintáctica.

29. Peeters (2006) reúne trabajos que aplican la teoría de Wierzbicka a varias lenguas romances.

30. Wierzbicka (1996) recoge una lista completa de primitivos, clasificados según el tipo de concepto codificado. Esta no es la única propuesta de primitivos semánticos. Compárese, por ejemplo, con la lista de A. Zholkovsky, representante de la Escuela Semántica de Moscú y creador, junto con I. Mel'čuk y Y. Apresjan, del Modelo Sentido-Texto: *set, object, property, relation, include, no, and, or, if-then, identical, true, all, there is, more [than], dimension, number, time, space, person, want, think, speaker [=I], norm* (citado por Mel'čuk, 1989).

31. Adviértase que la fórmula *I assume* 'yo asumo' introduce la presuposición, información de carácter pragmático estrechamente relacionada con la semántica léxica.

El potencial descriptivo de la teoría de los prototipos se alcanza a costa de la capacidad explicativa. Es cierto que el supuesto de los límites borrosos facilita el tratamiento de muchos fenómenos lingüísticos, pero también complica el análisis de la composicionalidad: si ningún rasgo es necesario ni suficiente, resulta difícil hablar de elementos compartidos por palabras diferentes (cfr. la argumentación en Fodor y Lepore, 2002). También la identificación entre las unidades del dominio cognitivo y las del lenguaje conlleva el riesgo de reducir lo lingüístico a lo puramente conceptual.

La utilidad de las nociones prototípicas para la lexicografía depende del tipo de diccionario. Como indican Geeraerts (1985) y Zgusta (1994), los diccionarios técnicos y enciclopédicos dan una descripción científico-técnica del significado; los diccionarios generales monolingües (*diccionarios de lengua* en Geeraerts, 1985) se centran en la presentación de estereotipos comunes para una comunidad lingüística, útiles asimismo en los diccionarios históricos; por último, el análisis semántico en los diccionarios lingüísticos está determinado por la noción de prototipo: los conceptos se describen y se ordenan con respecto a su relación con el prototipo (el concepto o la acepción central).

### 3.2.2. *Las Funciones Léxicas de la Teoría Sentido-Texto*

Frente a propuestas como la de Wierzbicka, que parten de un conjunto de primitivos, las teorías que se describen en este apartado y el siguiente hacen hincapié en el proceso mismo de la descomposición o la descomposición parcial, ya que esta es imprescindible y a menudo suficiente para descubrir las relaciones entre palabras. La descomposición es relevante dentro de la Teoría Sentido-Texto (cfr. referencias *infra*) porque la correspondencia entre las dos entidades que la configuran (el significado y la señal o el texto) se establece mediante paráfrasis sinónimas: a un mismo significado pueden corresponder varios textos sinónimos y viceversa.

El Modelo Sentido-Texto (MST) es en cierto modo único porque su componente léxico es, de hecho, un diccionario teórico. El *Diccionario Explicativo y Combinatorio* (DEC) es un diccionario *codificador* (orientado hacia la producción del texto), de *base semántica* (se da prioridad a la descripción del significado léxico, que determina en gran medida las propiedades sintácticas), *exhaustivo* (recoge todo lo que un hablante nativo sabe sobre una palabra) y centrado en casos de coaparición restringida. En este momento existen dos diccionarios de este tipo: el *Diccionario explicativo y combinatorio del Ruso Moderno* (*Tolkovo-kombinatornyj slovar' sovremennogo russkogo jazyka*, Mel'čuk y Zholkovsky, 1984), y el *Diccionario explicativo y combinatorio del francés contemporáneo*, DECFC (*Dictionnaire explicatif et combinatoire du français contemporain*, Mel'čuk et al., 1984-1999), con cuatro tomos publicados a día de hoy.

Para facilitar el análisis de los elementos básicos de las entradas del DECFC, ofrecemos en (6) el ejemplo del sustantivo francés *reproche* ('reproche'), con los términos lexicográficos en español, tal y como aparece en Mel'čuk (2001: 305-306):<sup>32</sup>

32. Las limitaciones de este trabajo nos impiden comentar con detalle la entrada de (6), por lo que remitimos al lector al artículo citado, a otros trabajos de Mel'čuk incluidos en la bibliografía y a la web del OLST (Observatorio de Lingüística Sentido-Texto: [olst.ling.umontreal.ca](http://olst.ling.umontreal.ca)). Si nos detendremos a continuación en su parte más innovadora, y la que más nos interesa: las funciones léxicas.

(6) **REPROCHE 1a.** ‘reproche’, nombre, masculino**Definición:**

*Reproche de X à Z à propos de Y = Fait que X reproche1a Y à Z*  
 [=S<sub>0</sub>(reprocher1a)]

**Régimen**

<i>X=I</i>	<i>Y=II</i>	<i>Z=III</i>
1. <i>de N</i>	1. <i>à propos &lt;au sujet&gt; de N</i>	1. <i>à N</i>
2. <i>de la part de N</i>	2. <i>pour N</i>	2. <i>à l'endroit de N</i>
3. <i>A<sub>pos</sub></i>	3. <i>sur N</i>	3. <i>envers N</i>
	4. <i>de V<sub>inf</sub></i>	
	5. <i>pour V<sub>inf. passé</sub></i>	

- 1)<sup>33</sup> C<sub>II.4</sub> : M<sub>1</sub>(M<sub>2</sub>) = M<sub>3</sub>  
 2) C<sub>I.1</sub> + C<sub>II.4, II.5, III.1</sub>  
 3) C<sub>I.2</sub> + C<sub>II.3</sub>  
 4) C<sub>II.4</sub> + C<sub>III.2, III.3</sub> } : **no deseable** [=uso censurable]

*le reproche de (la part de) Marie, son reproche; des reproches à propos de <=au sujet de, pour, sur> sa mauvaise conduite; le reproche de Marie <son reproche> à l'endroit de <=envers> ses professeurs, son reproche à ses professeurs*

**No deseable:**

*?les reproches de Marie de s'être <=pour s'être> mal conduit, ?les reproches de Marie à Pierre (2) [=les reproches de Marie à propos de la mauvaise conduite de Pierre, les reproches de Marie à l'endroit de Pierre]*

**Funciones léxicas**

Syn<sub>n</sub> : accusation, blâme1, désapprobation, grief, réprobation  
 Anti<sub>n</sub> : approbation; pardon  
 V<sub>0</sub> : reprocher1a  
 S<sub>2</sub> : cause, motif, objet, sujet [de (ART) ~], matière [à ~] | C<sub>i</sub> = Λ  
 PredAble<sub>3</sub> : mériter [ART ~]  
 AntiAble<sub>3,3</sub> : irréprochable1  
 Magn : sérieux | antepuesto, grave | antepuesto o pospuesto, lourd | antepuesto  
 AntiMagn : léger | antepuesto o pospuesto, petit | antepuesto  
 Ver : juste | antepuesto o pospuesto, justifié, mérité

33. Esta parte describe la combinatoria de los actantes de la tabla de régimen: C es columna; los números romanos identifican las columnas y los números árabes las líneas. M es ‘actante sintáctico profundo’, la fórmula dice que el primer actante del infinitivo coincide con el tercer actante del verbo *reprocher*.



AntiVer	: injuste   antepuesto o pospuesto, injustifié, non fondé
Oper <sub>1</sub>	: faire [à N]   C2 = de V <sub>inf</sub>
Oper <sub>3</sub>	: encourir [les ~s]   R. en plural, C2 ≠ Λ
Caus <sub>3</sub> Func <sub>3</sub>	: (se) mériter, s'attirer [les ~s]   R. en plural, C1 ≠ Λ
A <sub>2</sub> Manif	: de [~] [ton <air; expression> de reproche; lettres <paroles> de reproche]
Magn+A <sub>2</sub> Manif	: lourd, plein, rempli 1 [de~] G= air, expression, paroles, regard, ton, voix, yeux
nonPerm <sub>1</sub> Manif	: inavoué

El artículo se divide en varias zonas; las más importantes son: la semántica o la definición (fórmula proposicional con variables que representan actantes semánticos); la sintáctica: el régimen o el marco de subcategorización, que especifica para cada actante semántico el actante sintáctico correspondiente y su forma de realización; y la léxica (o zona de coaparición léxica), donde se registran las funciones léxicas relacionadas con la acepción (*lexía*) que trata el artículo.

La *función léxica* (FL) se define como «una función en el sentido matemático: una dependencia, o correspondencia, **f** que asocia a una lexía L, llamada el *argumento* [o *palabra clave*] de **f**, un conjunto de lexías **f(L)** —el *valor* de **f**— (Mel'čuk, 2001). Una relación entre una lexía L y un conjunto de lexías **f(L)** se considera una FL si: 1) se da en un número considerable de pares de palabras (como la relación de 'totalidad' o 'agregado' en las expresiones *ramo de flores*, *rebaño de ovejas*, *jauría de perros*, y muchas otras); y 2) posee diversos medios de expresión, que revelan una relación fraseológica entre el valor de **f** y sus argumentos: p. ej., el significado de intensidad, la FL **Magn** de (6) se expresa de manera diferente en las combinaciones *chillar como un descosido*, *sordo como una tapia* y *llover a cántaros*.

Hay dos tipos de FLL: paradigmáticas (relacionadas con la selección de un valor adecuado del argumento) y sintagmáticas (combinatorias, que cifran el potencial colocativo del argumento). En (6), las FLL **Syn**, **Anti**, **V<sub>0</sub>**, **S<sub>2</sub>**, **Able** son paradigmáticas y **Pred**, **Magn**, **Ver**, **Oper**, **Caus**, **Func**, **Manif** y **Perm** son sintagmáticas.

Describamos primero las FLL paradigmáticas. **Syn** y **Anti** proporcionan para el lema una lista de sinónimos y antónimos. **V<sub>0</sub>** da como valor el verbo derivado del sustantivo clave sin cambio de significado (*reproche-reprochar*). **S<sub>2</sub>** es una función nominal que da el nombre típico para el segundo actante de la palabra clave (*causa*, *motivo*, *objeto* o *sujeto de reproche*). **Able** codifica una propiedad que el participante potencial de una situación debe cumplir ('tal que puede' o 'tal que se le puede'). El valor de **Able**(*reproche*) en español sería *reprochable*; en (6) **Able** forma parte de la combinación **AntiAble**, que relaciona *reproche* con *irreprochable*.

En el conjunto de FLL sintagmáticas, **Pred** cubre la combinación de un verbo copulativo (*ser*; *estar*) con la palabra clave y aparece combinado con otras FLL (en (6), **PredAble**(*reproche*) = *merecer* [*el reproche*]). Las funciones **Magn** y **Ver** son adjetivales. **Magn** es un intensificador: **Magn**(*reproche*) = *serio*, *grave*. **Ver** codifica la relación 'tal como debe ser' o 'que cumple los requisitos imprescindibles': **Ver**(*reproche*) = *justificado*, *merecido*. Mencionemos tres de las FLL verbales: **Oper**, **Caus** y **Func**. **Oper** y **Func** formalizan la relación de verbo soporte y solo se diferen-

cian por el papel sintáctico que cumple la palabra clave: **Oper** toma la palabra clave como su objeto directo (**Oper**(*reproche*) = *hacer [un reproche]*) y **Func** como su sujeto (**Func**(*silencio*) = [*el silencio*] *reina*). En nuestro caso, **Func**(*reproche*) podría tener el valor *proviene de*, pero su estructura actancial se ve modificada por la FL **Caus** (que introduce el actante de causa): **CausFunc**(*reproche*) = *atraer [reproches]*.<sup>34</sup>

Una ventaja de las FFL en punto de eficacia y aplicabilidad es su número reducido (alrededor de 60) y su universalidad: el carácter abstracto y general de las relaciones que expresan las FFL garantiza su presencia en todas las lenguas, y el metalenguaje formal no está condicionado por la influencia de una lengua específica (a diferencia, p. ej., de los primitivos de Wierzbicka, que se formulan en inglés).

Este breve análisis de algunas de las FFL previstas por el MST permite sacar conclusiones sobre su alcance teórico y su utilidad para la lexicografía. La noción de FL permite describir las relaciones de una palabra a partir de un único criterio básico: los significados de las FFL, que son lo suficientemente amplios como para permitir generalizaciones importantes salvando el obstáculo de la idiosincrasia léxica. En este sentido, las FFL son un elemento clave de la relación léxico-sintaxis: codifican los rasgos aspectuales de los predicados, las propiedades de su estructura argumental, etc. En el plano sintagmático, las FFL proporcionan una base lingüística para el estudio de las colocaciones, que serían un producto de las restricciones que impone la base (la palabra clave) sobre el colocativo (el valor de la FL). Este enfoque contrasta con el análisis puramente estadístico de la coaparición de unidades léxicas (cfr. *supra* nota 8).

Además del DECFC, el MST se aplica en proyectos centrados específicamente en la combinatoria léxica y las colocaciones: el DiCo (*Dictionnaire de combinatoire*, OLSST, Universidad de Montreal), y el DiCE (*Diccionario de colocaciones del español*, Universidade da Coruña, dirigido por Margarita Alonso Ramos). (7) reproduce una entrada abreviada del DiCE:

- (7) **Orgullo** 1a Sentimiento [o = orgullo]  
**orgullo de individuo X de Y**  
**Cuasisinónimos:** *satisfacción*1a **Cuasiantónimos:** *vergüenza*1.1a
- ATRIBUTOS DE LOS PARTICIPANTES  
**que siente mucho** o FL Magn+A<sub>1</sub> *lleno [de ~]*  
**causa del** o FL S<sub>II</sub> *fuelle [de ~]*
- ACEPCIÓN + ADJETIVO ACEPCIÓN + VERBO  
**intenso** o FL Magn ~ *absoluto* **el** o **empieza** FL IncepFunc<sub>1</sub> *invadir [a X]*  
**justificado** o FL Ver ~ *legítimo*
- VERBO + ACEPCIÓN NOMBRE DE LA ACEPCIÓN  
**sentir** ~ FL Oper<sub>1</sub> *tener [ART ~]* **aparición repentina de** ~ FL Sing *acceso [de ~]*

En la parte inicial de la entrada se representa el significado central de palabra ('sentimiento'), una forma proposicional que codifica su estructura actancial ('orgullo

34. La lista completa de las FFL y su descripción detallada se puede consultar en Mel'čuk (1996, 2001).

de individuo X de Y'), y sus sinónimos y antónimos. A partir de ahí son posibles varios tipos de consulta, que llevan a las palabras vinculadas con el lema; dichas relaciones se codifican por medio de FLL (**Magn, Ver**, etc.) y de sus paráfrasis en el lenguaje natural (p. ej., 'intenso', 'justificado'). Se describen los *atributos de los participantes* del predicado que se basa en el lema (como *lleno de orgullo, fuente de orgullo*), las combinaciones 'acepción+adjetivo' (*orgullo absoluto, legítimo*, etc.), 'verbo+acepción' (*tener el orgullo*), 'acepción+verbo' (como *invadir [a X]*) y varias combinaciones reunidas bajo la rúbrica 'nombre de la acepción' (p. ej., *acceso de orgullo*).<sup>35</sup>

El uso de diccionarios como los reseñados tiene un potencial muy amplio (cfr. Mel'čuk, 2001; Alonso Ramos, 2004), desde su aplicación en sistemas de traducción automática (para dar cuenta de la equivalencia semántica de expresiones lexicalizadas en idiomas diferentes) hasta la enseñanza de lenguas extranjeras (el uso de colocaciones es un indicador importante del dominio de segundas lenguas).

### 3.2.3. *El Lexicón Generativo*

Al igual que el MST, la teoría del Lexicón Generativo (LG, cfr. Pustejovsky, 1995, y trabajos posteriores) presenta un lexicón complejo y relacional, integrado por varios tipos de estructuras conectadas entre sí por medio de mecanismos generativos, que son los que explican la creatividad y la flexibilidad del nivel léxico.<sup>36</sup> El MST y el LG son modelos hiperespecíficos, i. e., incluyen informaciones léxico-semánticas que van más allá de lo que se suele entender como lingüísticamente relevante. El LG combina la exhaustividad en lo que a los tipos de estructuras y sus rasgos se refiere con el recurso a la *subespecificación* o falta de especificación de las expresiones léxicas que explica su polisemia y las capacita para intervenir en diferentes estructuras sintácticas y operaciones de composición semántica (cfr. Pustejovsky, 1995. Cfr. asimismo el cap. 5 de la tercera parte de este volumen, donde se define el concepto con el término de *infraespecificación*, usado a menudo en la bibliografía del español). Por eso podemos expresar diferentes facetas del significado léxico y gramatical de una palabra como potenciales o fijadas parcialmente, obteniendo una interpretación plena dependiendo del contexto.<sup>37</sup> Esto evita el tener que describir todos sus rasgos, tarea imposible tanto por la flexibilidad inherente del significado como por los problemas que implica la descomposición (§ 3.1.1). En el ámbito lexicográfico, la subespecificación se traduce

35. Véase el apartado *Consultas* de [www.dicesp.com/](http://www.dicesp.com/) para más información.

36. El LG tiene una ventaja importante con respecto al MST: la información heterogénea comprimida en las FLL se distribuye dentro del LG en varias estructuras (cfr. el capítulo 5 de la tercera parte, de este volumen): p. ej., los rasgos aspectuales codificados por las FLL **Caus, Incept** y **Temp** se incluyen en la estructura eventiva. Este tipo de organización ofrece una mayor coherencia en cuanto al tipo de información que abarca la misma estructura y dota el modelo de una mayor flexibilidad y un potencial explicativo superior.

37. Esta idea está detrás del *Corpus Pattern Analysis* (CPA, Análisis de Patrones de Corpus, cfr. Hanks y Pustejovsky, 2005; Pustejovsky *et al.*, 2004). Si las palabras aportan su potencial semántico al contexto oracional, entonces su significado básico debería ser deducible del contexto. El CPA agrupa los patrones de uso de los verbos y los analiza a nivel sintáctico (en términos de número y categoría sintáctica de los miembros del predicado, incluyendo los opcionales) y semántico (en términos de *tipos semánticos*, como *entidad, persona, objeto, institución, evento, proceso*, etc., y *roles semánticos*, como *agente, paciente, doctor, beneficiario, pagador*). Esta información sirve para elaborar patrones semánticos que identifican las acepciones del predicado (el patrón del verbo *tostar*, p. ej., sería: [Persona] *tostar* [Comida]).

en la creación de entradas léxicas con un grado de especificidad variable, lo que facilita el paso de un lexicón de tipo general a uno técnico (de la medicina, la ingeniería, etc.) sin cambiar la arquitectura general de la descripción lexicográfica.

No es nuestro propósito presentar el LG en detalle aquí (al que se dedica el capítulo 5 de la tercera parte de este volumen). Me centraré en exclusiva en la estructura de *qualia* (EQ) porque constituye la parte más innovadora del LG y la que más repercusión teórica y aplicada ha tenido, y porque es el elemento central del LG, que repercute en los demás niveles.

La EQ es un sistema semántico relacional cuyo funcionamiento se determina por unos parámetros o roles que permiten describir un objeto, un evento o una propiedad:

- *Rol constitutivo*: relaciona una entidad con sus partes constituyentes.
- *Rol formal*: codifica los rasgos distintivos de una entidad (orientación espacial, tamaño, forma, dimensionalidad, color, etc.).
- *Rol télico*: denota el propósito y la función de una entidad.
- *Rol agentivo*: alude a los factores implicados en el origen o la creación de la entidad.

La prominencia de unos u otros roles de la EQ en la descripción de diferentes clases de palabras permite elaborar un sistema de tipos semánticos. Se destacan:

- *Los tipos naturales o simples*, basados en los roles formal y constitutivo (como en *roca, tigre, grande, plano, llover, tener*). Son simples porque se pueden describir en función de un solo criterio y tener subespecificados los demás (del sustantivo *tigre* se destaca su posición en la clasificación biológica y su descripción física —rol formal y constitutivo—, pero su rol télico estará subespecificado, al no haber una función inherente).
- *Los tipos unificados*: aluden a entidades creadas o artefactos (p. ej., *cerveza, maestro, cuchillo, estropear*), y resultan de la aplicación de los valores agentivo y télico a las especificaciones de los tipos naturales. Son tipos multidimensionales, clasificados con respecto a varios criterios, de ahí el operador conjuntivo  $\otimes$  que se utiliza para estos casos, como en Pustejovsky (2006):

- (8) a. *cerveza*: líquido $\otimes$ <sub>T</sub>beber [T= rol télico]  
 b. *cuchillo*: obj.físico $\otimes$ <sub>T</sub>cortar  
 c. *casa*: obj.físico $\otimes$ <sub>T</sub>habitar

- *Los tipos complejos*: representan conceptos que incorporan el significado de dos o más tipos semánticos relacionados de forma inherente (piénsese en *libro*, que es a la vez del tipo ‘información’ y ‘objeto material’ o *concierto*, que es ‘obra musical’ y ‘evento’).

La EQ es el elemento del LG más cuestionado en lo teórico por acercarse mucho más al conocimiento enciclopédico que, por ejemplo, los rasgos de selección tradicionales (como ‘animado’, ‘concreto’, ‘abstracto’, ‘continuo’). Esto no es de extrañar porque

los *qualia* son «herederos» de los *aitiai* de Aristóteles, cuyo objetivo no era establecer un repertorio de rasgos léxicos, sino analizar las posibles configuraciones de los elementos de la realidad y explicar como entendemos el mundo. El criterio teórico para incluir la información semántica en el LG consiste en mostrar que repercute en la sintaxis. Su justificación lexicográfica, por otra parte, viene dada por el hecho de que la EQ alude a unos aspectos básicos del significado (origen, forma, constitución, función) con los que todos estamos familiarizados y sin los que la definición de muchas palabras sería muy difícil (sobre todo, pero no únicamente, la de los sustantivos concretos).<sup>38</sup>

La especificidad de los *qualia* se explica además porque el LG ha sido concebido como un componente básico de sistemas de procesamiento del lenguaje que, como hemos dicho, suelen recurrir a mayores cantidades de conocimiento que los diccionarios usuales para poder simular la competencia de los hablantes.<sup>39</sup>

Quizá la aportación de más peso del LG a la lexicografía computacional haya tenido lugar en el marco del proyecto SIMPLE (*Semantic Information for Multipurpose Plurilingual Lexicons*), cuyo objetivo es desarrollar dentro de un modelo homogéneo lexicones semánticos de amplio alcance para 12 lenguas europeas: alemán, catalán, danés, español, finés, francés, griego, holandés, inglés, italiano, portugués y sueco.

En SIMPLE, la EQ es el fundamento de su núcleo conceptual, la ontología básica (el sistema de tipos compartido por todos los lexicones); de hecho, los roles de la EQ son los nodos más altos de la jerarquía semántica. La EQ permite paliar las deficiencias de las ontologías hiperonímicas (§ 3.1.1), que se limitan a proporcionar la información de tipo ISA<sup>40</sup> (en el LG de esto se encarga el rol formal). Este principio (que se adopta, p. ej., en el proyecto de WordNet),<sup>41</sup> lleva a que sus nodos tengan que asumir varios tipos de información a la vez, haciéndola menos accesible. En el ejemplo de (10), tomado del MultiWordNet español y abreviado, *la pulsera* se describe en el paso 1 como adorno hecho de metales y piedras preciosos, mencionando sus partes

38. Recuérdese la propuesta de Casares (1950), que destaca dentro de la definición real varios subtipos: la *genética*, que «procura explicar el definido como el efecto de la acción que lo engendra», la *teleológica*, que «nos informa de la naturaleza de las cosas en razón del fin a que se destinan», y la *descriptiva*, que contiene entre otras la información sobre la forma (de las plantas y los animales, p. ej.). Es obvio el paralelismo entre la definición genética y el rol agentivo de la EQ, entre la teleológica y el valor del rol télico, y finalmente entre la descriptiva y el rol formal.

39. Piénsese, por ejemplo, en las *clases de objetos* (como *adecuado, defecto físico/moral/intelectual/psicológico, edad colectiva, función, cualidad moral/psíquica/intelectual/física, movimiento político*, etc.) del modelo de G. Gross, que fueron introducidas porque los rasgos de subcategorización y selección no bastan para delimitar el dominio de los argumentos de un predicado (Gross, 1994, y Herrero Ingelmo, 2000).

40. ISA es la relación 'x es (un tipo de) y'. La lexicografía moderna utiliza ISA para reflejar las propiedades compartidas por las piezas léxicas a consecuencia de la herencia. Por ejemplo, *cuchara* es un (ISA) *utensilio*, el primero es el hipónimo del segundo y hereda sus rasgos básicos ('es una cuchara' implica 'es un utensilio'). *Utensilio*, a su vez, es hipónimo de *instrumento*, y, en un nivel aún más general, de *cosa*. Véase también el capítulo 4 de la cuarta parte de este volumen.

41. Además de relaciones hiponímicas, WordNet recurre a la meronimia, la implicación (*comprar* implica *pagar*, p. ej.), la troponimia (que relaciona dos verbos de forma que uno especifica una determinada manera de realizar la acción denotada por el otro, como *abofetar* y *pegar*) y la causalidad (*matar* causa *morir*). No obstante, esto no neutraliza los inconvenientes de las jerarquías en cuestión: el criterio hiperonímico es el único que rige la herencia (cfr. Fellbaum, 1998, y el capítulo 4 de la cuarta parte de este volumen).

En EuroWordNet los *qualia* constituyen los *top concepts* de las entidades de primer orden (entidades concretas), es decir, se usan para clasificarlas según su origen, forma, composición y función (Vossen, 2001).

constitutivas, en los pasos 2 y 3 se remite a su función ('un elemento de decoración que se añade para evitar la monotonía' y 'algo utilizado para embellecer'); y en el 4 se alude a su origen (objeto hecho por el hombre):

## (10) Sentido 1 'brazalete, pulsera'

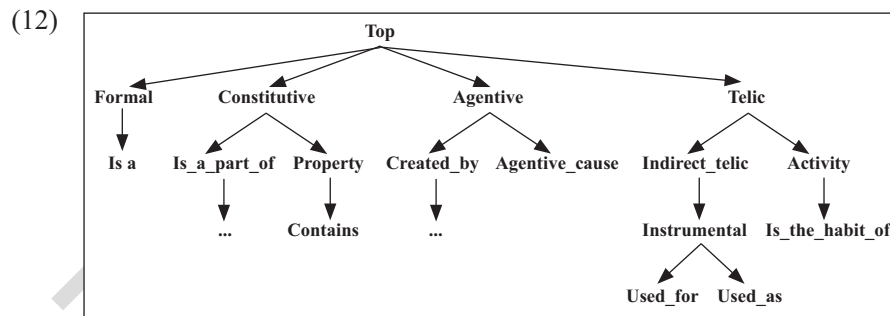
- 1 → joyas, joyería - [*an adornment made of precious metals and set with gems*]
- 2 → adorno, ornamento - [*a decoration of color or interest that is added to relieve plainness*]
- 3 → adorno, decoración, ornamento - [*something used to beautify*]
- 4 → artefacto - [*a man-made object*]
- 5 → cosa, objeto, objeto\_físico, objeto\_inanimado
- 6 → entidad

La utilización de los *qualia* permite evitar estos inconvenientes y dotar el sistema de unos patrones de definición completos y eficaces. Cfr. el patrón para el tipo 'obra de arte' tal y como se presenta en Lenci *et al.* (2000a), con la información para el ejemplo concreto de *pulsera* tomada de la muestra de SIMPLE español disponible en la web:

## (11)

		Traducción [O. B.]
<b>UseM</b>	<pulsera>	
<b>BC number</b>	<Ø>	
<b>Template_Type</b>	[Artwork]	[obra de arte]
<b>Unification_path</b>	[Concrete_entity   Artifact <sub>Agentive</sub> ]	[Entidad_concreta   Artefacto <sub>Agentivo</sub> ]
<b>Domain</b>	Arts	Arte
<b>Semantic Class</b>	Artifact	Artefacto
<b>Gloss</b>	<Ø>	
<b>Pred_Rep.</b>	<Ø>	
<b>Selectional Restr.</b>	<Ø>	
<b>Derivation</b>	<Ø>	
<b>Formal</b>	<i>isa</i> (<joya>, <Artwork>: [Artifact])	<i>isa</i> (<joya>, <obra de arte>: [Artefacto])
<b>Agentive</b>	<i>created_by</i> (<pulsera>, <hacer>: [Creation])	<i>creado_por</i> (<pulsera>, <hacer>: [Creación])
<b>Constitutive</b>	<i>constitutive</i> <muñeca>, <Bodypart>	<i>constitutivo</i> <muñeca>, <parte del cuerpo>
<b>Telic</b>	<i>used_for</i> (<adornar>, <CauseChange State>)	<i>usado_para</i> (<adornar>, <Causar-CambioEstado>)
<b>Synonymy</b>	<Ø>	
<b>Collocates</b>	<i>Collocates</i> (<UseM1>, ..., <UseMn>)	<i>Colocativos</i> (<UseM1>, ..., <UseMn>)
<b>Complex</b>	<Ø>	

Los valores *isa*, *created by* y *used for* corresponden a la *EQ extendida* (para los roles formal, agentivo y télico), que proporciona un conjunto de subtipos para cada rol. Algunos de los valores de la EQ extendida se representan en (12) (Lenci *et al.*, 2000b):



Hasta ahora hemos estado hablando de la aplicación de *qualia* a las características intensionales de las piezas léxicas, pero podemos también interpretarlos como una relación significativa entre dos elementos (al modo de las FFLL), que identifica el conjunto de palabras combinables con una palabra dada y su interpretación. Veamos el caso del verbo *rematar*:

- (13) a. El pistolero remató a su víctima [mató a la víctima].  
 b. Los puños de la camisa estaban rematados por un bordado [el bordado es una parte de los puños de la camisa].  
 c. Tardaron dos días en rematar las provisiones [en comérselas].  
 d. Por fin sacó tiempo para rematar su segunda novela [para acabar de redactarla].

*Rematar* establece una serie de relaciones con sus argumentos (que precisan su sentido general de ‘poner fin a algo’) explotando diferentes aspectos del significado de estos: *víctima* (rol formal, ‘humano o animado’); *puños* (rol constitutivo, ‘parte exterior, margen’); *provisiones* (rol télico, ‘consumir’); *novela* (rol agentivo, ‘crear escribiendo’).

Al menos algunas de las oraciones de (13) se perciben como traslaticias, y sin embargo se pueden interpretar en los mismos términos que los ejemplos que no tienen este matiz. Si es cierto que la EQ, las FFLL, o algún mecanismo o estructura similar pueden servir para determinar lo que podemos hacer con una palabra dentro de un predicado, entonces nos encontramos frente a un fenómeno global, que no solo permite descartar la existencia de la combinatoria libre, sino estudiar dentro del mismo marco (el de la *selección léxica*) fenómenos tan diversos como las colocaciones y otros casos de coaparición frecuente de palabras, la modificación adverbial y adjetival, la caracterización semántica de los argumentos del predicado en términos más sutiles que los papeles temáticos, etc.

El término de selección léxica ha sido utilizado por Ignacio Bosque para de-

finir las condiciones léxicas que los predicados imponen a sus argumentos dentro del *Diccionario Combinatorio del Español Contemporáneo* REDES. REDES parte del estudio extensivo de los argumentos que seleccionan las categorías predicativas o colocativos (adjetivos, verbos y adverbios)<sup>42</sup> y posteriormente los agrupa en *clases léxicas* (se trata, por tanto, de un proceso de inducción empírica). Por ejemplo, la locución *en son de paz* se combina con las siguientes clases léxicas: A) verbos de movimiento (*venir, acudir, regresar* y otros); B) verbos que designan la acción de poner algo de manifiesto (como *llamar* y *señalar*); C) sustantivos que denotan encuentro (p. ej., *visita, reunión*), y D) sustantivos que designan personas o colectivos cuando estos se asocian con los movimientos del grupo A (*soldado, okupa*, etc.). El cruce de los predicados que seleccionan una pieza léxica como su argumento permite descubrir aquellos de sus rasgos léxicos «que pueden no aparecer en su definición, pero que la sintaxis saca a la luz de manera bastante objetiva» (Bosque, 2004a: CXLIV).

#### 3.2.4. *Las clases verbales de Levin: VerbNet*

El abanico cada vez más amplio de tareas de procesamiento del lenguaje que se sustentan en los lexicones computacionales exige, entre otras cosas, que estos incluyan información sobre la estructura argumental de las piezas léxicas (básicamente verbos, por su carácter predicativo) y su realización concreta. En el MST estos datos se distribuyen entre la zona semántica (la estructura actancial) y la sintáctica (los patrones de subcategorización) de las entradas léxicas. El mismo papel es desempeñado por la estructura argumental del LG, cuyas propiedades están restringidas por la estructura eventiva y la EQ.

La propuesta teórica que describimos en este apartado se enmarca dentro de los estudios que derivan el comportamiento sintáctico de las palabras de su significado (hipótesis conocida como *determinismo semántico*, cfr. el cap. 4 de la tercera parte de este volumen). En su libro *English Verb Classes and Alternations* (1993), Beth Levin ofrece una clasificación léxico-semántica de los verbos del inglés en función de las alternancias que presentan en la realización sintáctica de sus argumentos. Una vez establecidos los grupos de verbos según las alternancias detectadas, se determina qué elementos de significado son comunes para todos sus miembros, con lo que se obtiene un repertorio de rasgos lingüísticamente relevantes. Una consecuencia inmediata del determinismo semántico con respecto al contenido de las entradas es que, dados ciertos rasgos semánticos, ya no es necesario especificar las propiedades combinatorias de la palabra (como el marco de subcategorización y de selección de los verbos).

En (14) se resume un ejemplo del análisis de Levin (1993); en la tabla se establece una relación entre las alternancias sintácticas en las que cada grupo de verbos participa y determinados elementos de su significado, y en (14a-c) se dan ejemplos de dichas alternancias en inglés:

42. En esto REDES se diferencia del DEC: este parte de la base y llega a través de las FLL a los colocativos; aquel, en cambio, se centra en el colocativo que se contempla como predicado selector.



(14)

	<i>Ascenso de Poseedor Inalienable 'contacto'</i>	<i>Alternancia de conato 'contacto+movimiento'</i>	<i>Alternancia media 'cambio de estado'</i>
<i>touch (pat, stroke, tickle)</i>	Sí	No	No
<i>hit (bash, pick, pound)</i>	Sí	Sí	No
<i>cut (hack, saw, slash)</i>	Sí	Sí	Sí
<i>break (crack, rip, snap)</i>	No	No	Sí

a. Ascenso de Poseedor Inalienable:

*Margaret cut {Bill's arm / Bill on the arm} versus Janet broke {Bill's finger / \*Bill on the finger}*

Lit.: *Margaret cortó {el brazo de Bill / a Bill en el brazo} versus Janet rompió {el dedo de Bill / a Bill en el dedo}*

b. Alternancia de conato:

*Carla hit {the door / at the door} versus Janet broke {the vase / \*at the vase}*

Lit.: *Carla golpeó {la puerta / en la puerta} versus Janet rompió {el florero / en el florero}*

c. Alternancia media:

*The bread cuts easily versus \*Cats touch easily*

Lit.: *El pan se corta fácilmente versus Los gatos se tocan fácilmente*

El proyecto VerbNet es la implementación más consistente de las clases verbales de Levin. Después de varias extensiones de su propuesta (adición de subclases a las clases originales, reorganización de estas e introducción de clases nuevas), se consiguió una cobertura más completa del léxico verbal inglés: el número de clases de primer nivel pasó de 191 a 274, y el número de acepciones de 3445 a 3769 (Kipper *et al.*, 2006).<sup>43</sup>

Cada clase verbal se describe en términos de papeles temáticos, restricciones de selección y unos marcos (*frames*)<sup>44</sup> que reúnen la descripción sintáctica, semántica y temporal. Reproducimos en (15) la descripción de la clase verbal *patent-101* tal y como se da dentro de VerbNet v. 2.3, con una representación ligeramente modificada:

(15)

<b>Patent-101</b>		Class Hierarchy: no subclasses
<b>Members:</b>	Certify, copyright, credential, license, patent, register, trademark	
<b>Roles and restrictions:</b>	• Agent [+animate, +organization] • Theme	
<b>Frames</b>	<b>Name:</b>	NP
	<b>Example:</b>	«I patented my idea»
	<b>Syntax:</b>	Agent V Theme
	<b>Semantics:</b>	license(result(E), Theme) cause (Agent, E)

42. Como resultado de la extensión ha sido cubierta una laguna notable dentro de la clasificación original: la ausencia de información sobre complementos adjetivales, adverbiales, oracionales y algunos otros.

43. Este uso del término *marco* no debe confundirse con el descrito en § 3.2.5 en relación con la semántica de marcos. La lista de etiquetas para cada tipo de información se puede consultar en la dirección electrónica: [verbs.colorado.edu/verb-index/reference.php](http://verbs.colorado.edu/verb-index/reference.php).

### 3.2.5. *La semántica de marcos: FrameNet*

La *semántica de marcos* (*Frame Semantics*) destaca entre otras teorías por la amplitud de su alcance: establece un vínculo directo entre la semántica léxica y la textual y ofrece a la vez un análisis de los fenómenos sintácticos basado en las propiedades semántico-cognitivas de las palabras. Dicho tratamiento abarca un amplio abanico de fenómenos pragmáticos o discursivos (por ejemplo, la presuposición, los postulados y las implicaturas conversacionales), que intervienen en la interpretación de una unidad léxica en el contexto definido de forma amplia.<sup>45</sup>

La teoría en cuestión se desarrolla a partir de la premisa cognitivista de que la experiencia y el conocimiento del hablante se configuran en torno a unas estructuras conceptuales específicas (creencias sobre el mundo, prácticas convencionales o rutinarias, instituciones socioculturales, etc.), que reciben el nombre de *escenas* (*scenes*) o *scripts* dependiendo del enfoque.<sup>46</sup> Estas estructuras cristalizan en el lenguaje en forma de *marcos semánticos*, «representaciones esquemáticas estructuradas de una situación, objeto o evento que proporcionan un trasfondo y motivación necesarios para la existencia y uso cotidiano de las palabras de una lengua» (Petrucci *et al.*, 2004).<sup>47</sup> La idea de la descripción del léxico en función de situaciones esquematizadas en las que intervienen varios tipos de entidades (participantes y otros elementos accesorios) representados en el lenguaje mediante *roles semánticos* o *casos* (como *agente, paciente, experimentante, instrumento, locativo, tiempo*) pertenece a Charles Fillmore, quien la aplicó en su *Gramática de Casos* (Fillmore, 1968, 1982 y 1985). Los roles semánticos (o papeles temáticos, cfr. Gruber, 1965) jugaron un papel importante dentro de la gramática generativa facilitando la proyección de la información semántica en la sintaxis. Dentro de la semántica de marcos, la noción de rol semántico evolucionó desde un conjunto universalmente válido y reducido de etiquetas a un repertorio en principio infinito de roles semántico-situacionales específicos para cada marco concreto.

La primera aplicación de los marcos a la descripción lexicográfica se encuentra en Fillmore y Atkins (1994), quienes analizan la codificación de la palabra *risk* ('riesgo, arriesgar') en varios diccionarios del inglés y proponen un tratamiento alternativo, que parte de una representación conceptual de la situación de riesgo, y relaciona los elementos de esta situación con su función sintáctica y su realización categorial. Estos elementos son el Protagonista que arriesga algo, las Consecuencias, negativas o positivas, la Decisión del protagonista, la Posesión (algo valorado por el protagonista y que está en peligro), y la Fuente (algo o alguien que puede causar daño).

45. Ciertas palabras, aparte de significado propiamente léxico, tienen asociada información pragmática que las encaja en el contexto. Garrido Medina (1990) las llama *activadores de presuposición* y pone como ejemplo el verbo *olvidarse de*, que implica la existencia de una obligación de hacer algo (*Juan se olvidó de lavar los platos* implica que tenía la obligación de hacerlo).

46. Estos términos se utilizan asimismo en el ámbito de la inteligencia artificial y la psicología cognitiva.

47. Adviértase que frente a la semántica de campos (§ 3.2.1), que define una palabra según las relaciones que esta establece con otras del mismo campo (p. ej., el sustantivo *profesor* como miembro del campo 'profesional', en el que estarían igualmente *albañil, militar*, etc.), la semántica de marcos le atribuye una interpretación en función del marco al que pertenece —*profesor* como *agente* o participante activo dentro del marco 'enseñanza' o 'transmisión de conocimiento', que incluiría además elementos como 'receptor de información' (*alumno*), 'lugar' (*aula, colegio*), etc.—.

La semántica de marcos ha inspirado la creación de la base de datos relacional FrameNet, inicialmente para el inglés (a partir de 1997 en la Universidad de Berkeley), y posteriormente para otras lenguas, entre ellas el español (el FrameNet Español se desarrolla en la Universidad Autónoma de Barcelona, dirigido por Carlos Subirats).

Las principales tareas que se realizan dentro de FrameNet son: la identificación de los marcos semánticos y los elementos conceptuales (*argumentos semánticos* o *frame elements*) que los configuran; la extracción de datos del corpus sobre las construcciones sintácticas en las que se proyectan los elementos de cada marco; la anotación semántica de las oraciones extraídas con el tipo de argumento semántico, la función sintáctica y su realización categorial; y, por último, la verificación de los resultados de anotación vía web y el procesamiento automático de los resultados de la anotación (Subirats, 2004).

Damos en (16) el ejemplo abreviado del marco *Predicting*<sup>48</sup> para que el lector pueda hacerse una idea sobre la organización de FrameNet (cfr. el comentario *infra*):

(16) **PREDICTING** **Definition:**

A <b>Speaker</b> states or makes known a future <b>Eventuality</b> on the basis of some <b>Evidence</b> .
---

**Core FEs:**

<b>Eventuality [Eve]</b>	This FE marks expressions that indicate an event or state that is predicted to occur or hold at a future time. A single <b>Eventuality</b> may be expressed discontinuously in a raising construction. This FE is also used for expressions that denote entities with variable attributes about which the <b>Speaker</b> makes predictions.
Ej.: <i>La muchacha <b>PREDIJO</b> que después del accidente su marido volvería a andar.</i>	
<b>Medium [Med]</b>	The physical or abstract setting in which the <b>Eventuality</b> is conveyed.
Ej.: <i>Los resultados de las elecciones corroboran las <b>PREDICCIONES</b> de las encuestas.</i>	
<b>Speaker [Spe]</b> <b>Semantic Type</b> Sentient	The <b>Speaker</b> makes a prediction.
Ej.: <i>Billy Hunter no quiso hacer <b>PREDICCIONES</b> acerca de cómo resultarían las pláticas.</i>	
<b>Topic [Top]</b>	<b>Topic</b> is the subject matter about which a <b>Speaker</b> states or makes known a future <b>Eventuality</b> . It is normally expressed as a PP-complement headed by <i>sobre</i> (about).
Ej.: <i>Se hicieron <b>PREDICCIONES</b> completamente erróneas sobre este tema.</i>	

48. Tomado de gemini.uab.es:9080/SFNsite. FrameNet Español utiliza la terminología inglesa de FrameNet. Para ampliar la información sobre FrameNet y su versión española, remitimos a las páginas web [framenet.icsi.berkeley.edu](http://framenet.icsi.berkeley.edu) y [gemini.uab.es:9080/SFNsite](http://gemini.uab.es:9080/SFNsite) y las publicaciones allí citadas.

*Non Core FEs:*

Accuracy [Acc]	This FE indicates the Accuracy of the prediction. Ej.: <i>Ningún especialista puede <b>PREDECIR</b> exactamente dónde golpeará y qué daños causará a su compañero.</i>
Concessive [Con]	A modifier expressing the deviation of the actual level of security from the expected value given the Speaker and the state indicated by the target itself. Ej.: <i>Max <b>PREDIJO</b> que su equipo vencería a pesar de tenerlo todo en su contra.</i>
Domain [Dom]	The aspects of the Eventuality (or Topic) which are predicted. Ej.: <i>El vidente hizo ayer sus <b>PREDICCIONES</b> astrológicas para el año 1997.</i>
Focal participant [Foc]	The Focal participant FE refers to an entity that is related to the Eventuality which is stated or made known by the Speaker. Ej.: <i>Las <b>PREDICCIONES</b> más sombrías para los demócratas circulaban el miércoles en las filas de la mayoría republicana.</i>
Manner [Man] Semantic Type Manner	Manner marks any description of the Speaker's prediction making that is not covered by more specific FEs. Ej.: <i>Eva <b>desesperadamente</b> <b>PREDIJO</b> que nada iba a ser tan malo como lo que le sucede ahora.</i>

Inherits From:      Is Inherited By:      Subframe of:      Has Subframes:  
 Uses: Expectation      Is Used By:      Is Inchoative of:      Is Causative of:      See Also:  
**Lexical Units**      *predecir.v, predicción.n*

En la descripción del marco se incluyen la definición y una lista de argumentos semánticos (*frame elements*, FE) principales o nucleares (*core*) y periféricos (*non-core*); se especifica el tipo de relación existente entre los marcos (en términos de herencia, submarco, causación, etc.) y, finalmente, se indican las piezas léxicas predicativas (verbos, sustantivos, adjetivos) que evocan el marco,<sup>49</sup> y que forman una *clase semántica* caracterizada por la identidad en la selección de argumentos conceptuales. Una entrada léxica de FrameNet contiene asimismo una tabla de realización sintáctica (la función y la categoría) de los argumentos semánticos y una tabla con los patrones actanciales en los que aparecen los argumentos semánticos de la unidad léxica dada.

Otro proyecto español que comparte algunas de las premisas teóricas de FrameNet es ADESSE (Alternancias de Diátesis y Esquemas Sintáctico-Semánticos del Español, Universidad de Vigo, coordinado por J. M. García-Miguel). ADESSE parte de la información sintáctica de la BDS (cfr. nota 20 *supra*) y le añade los datos semán-

49. La unidad léxica se define dentro de FrameNet como combinación de una palabra fonológica con un significado particular. Cuando a la misma palabra fonológica le corresponde más de un sentido (polisemia u homonimia), se habla de unidades léxicas diferentes que forman parte de marcos semánticos diferentes.

ticos que también aparecen en FrameNet: los sentidos verbales, los roles semánticos y las clases de verbos (que se corresponden en general con los marcos de FrameNet), con la diferencia fundamental de que dentro de ADESSE el punto de partida no es el marco sino el verbo. Siguiendo los postulados de la Gramática Cognitiva de Langacker y de la Gramática de Construcciones de Goldberg, ADESSE contempla la realización del verbo en el contexto como resultado de la interacción del significado verbal y el significado de la construcción sintáctica que lo acoge. De la construcción depende qué componentes del marco conceptual evocado por el verbo quedarán expresados en el contexto. Véase en (17) la realización de los argumentos semánticos de *prever* (A1-Conocedor, A2-Contenido) en dos construcciones sintácticas (tomado de [adesse.uvigo.es/](http://adesse.uvigo.es/)):

- (17) a. Soloviev [1] preveía una colaboración entre la técnica japonesa y la mano de obra china [2].  
b. Se prevé para el año próximo un «agujero» del 50 % de las horas [2].

En cuanto a su aplicación al ámbito lexicológico, la información que proporcionan FrameNet y ADESSE es relevante, p. ej., para establecer las bases conceptuales de la organización del léxico y para investigar la proyección de las propiedades semánticas de los predicados en la sintaxis.

#### 4. Conclusión

La pequeña muestra de aplicación de la teoría del léxico en la elaboración de diccionarios ofrecida en este capítulo confirma que la investigación lexicológica tiene mucho que aportar a diferentes tipos de repertorios lexicográficos (desde los diccionarios tradicionales hasta las bases de datos integradas en sistemas de procesamiento del lenguaje). Los ejemplos analizados son especiales porque su preocupación fundamental (por encima de los requisitos impuestos por la industria lexicográfica) consiste en presentar una visión global, científicamente fundada, del nivel léxico, que a buen seguro ha de influir en el desarrollo de la lexicografía en el futuro.

Esta revisión de las diferentes opciones teóricas parece confirmar que los modelos de tipo relacional no solo son los que más potencial teórico tienen (en la medida en que, además de llevar a cabo un estudio exhaustivo del componente léxico, analizan y formalizan su relación con la sintaxis y el discurso), sino también los de mayor repercusión en el tratamiento lexicográfico del léxico, al que dotan de los instrumentos necesarios para poder retratar la palabra con toda la riqueza de contenidos y asociaciones que implica.

### Bibliografía

- Bosque, I. (1982): «Sobre la teoría de la definición lexicográfica», *Verba*, 9, 105-124.
- (2004a): «Combinatoria y significación. Algunas reflexiones», en I. Bosque (dir.), LXX-VII-CLXXIV.
- Coseriu, E. (1977): *Principios de semántica estructural*, Madrid, Gredos.
- Cruse, D. A. (2000): *Meaning in Language. An Introduction to Semantics and Pragmatics*, Oxford, OUP.
- Fillmore, C. (1985): «Frames and the Semantics of Understanding», *Quaderni di Semantica*, 6:2, 222-254.
- Fillmore, C. y Atkins, B. (1994): «Starting where the Dictionaries Stop: The Challenge of Corpus Lexicography», en B. Atkins y A. Zampolli (eds.), 349-393.
- Haensch, G. et al. (1982): *La lexicografía. De la lingüística teórica a la lexicografía práctica*, Madrid, Gredos.
- Hausmann, J. (ed.) (1989): *Wörterbücher. Ein internationales Handbuch zur Lexikographie*, Berlin, De Gruyter.
- Kleiber, G. (1995): *La semántica de los prototipos*, Madrid, Visor Libros.
- Levin, B. (1993): *English Verb Classes and Alternations*, Chicago, The University of Chicago Press.
- Mel'čuk, I. (2001): «Fraseología y diccionario en la lingüística moderna», en I. Uzcanga Vivar et al. (eds.), *Presencia y renovación de la lingüística francesa*, Salamanca, Universidad de Salamanca, 267-310.
- Pustejovsky, J. (1995): *The Generative Lexicon*, Cambridge (Massachusetts), MIT Press.
- Rafel i Fontanals, J. (2005): *Lexicografía*, Barcelona, UOC.
- Wiegand, H. E. (ed.) (2002): *Lexikologie. Ein internationales Handbuch zur Natur und Struktur von Wörtern und Wortschätzen*, Berlín, De Gruyter.
- Wierzbicka, A. (1996): *Semantics: Primes and Universals*, Oxford, OUP.
- Zgusta, L. (2006): *Lexicography then and now*, Tübingen, Max Niemeyer.

### Referencias adicionales

- Almela Pérez, R. et al. (eds.) (2005): *Frecuencias del español. Diccionario y estudios léxicos y morfológicos*, Madrid, Editorial Universitat.
- 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, Barcelona, IULA, Documenta Universitaria, 149-162.
- Alvar Ezquerro, M. (dir.) (1995): *Diccionario ideológico de la lengua española Vox*, Barcelona, Biblograf.
- Apresjan, Y. (2000): *Systematic lexicography*, NY, OUP.
- Atkins, B. et al. (1994): «Computational Approaches to the Lexicon: An Overview», en B. Atkins y A. Zampolli (eds.), 17-45.
- Atkins, B. y Zampolli, A. (eds.) (1994): *Computational Approaches to the Lexicon*, NY, OUP.
- Azorín Fernández, D. (2003): «La lexicografía como disciplina lingüística», en A. Medina Guerra (ed.), *Lexicografía española*, Barcelona, Ariel, 31-52.
- Benson, M. et al. (1997): *The BBI Dictionary of English Word Combinations*, Amsterdam, J. Benjamins.
- Bosque, I. (2001): «Sobre el concepto de 'colocación' y sus límites», *Lingüística española actual*, 23:1, 9-40.

- (dir.) (2004b): *REDES. Diccionario combinatorio del español contemporáneo*, Madrid, SM.
- Calzolari, N. (1994): «Structure and Access in an Automated Lexicon and Related Issues», en D. Walker *et al.* (eds.), *Automating the lexicon*, Oxford, NY, OUP, 337-356.
- Casares, J. (1973): *Diccionario ideológico de la lengua española*, Barcelona, Gustavo Gili.
- (1992 [1950]): *Introducción a la lexicografía moderna*, 3.ª ed., Madrid, CSIC.
- Cruse, D. A. (1986): *Lexical semantics*, Cambridge, CUP.
- Cuervo, R. J. e Instituto Caro y Cuervo (1998): *Diccionario de construcción y régimen*, Barcelona, Herder.
- Escobedo Rodríguez, A. (2000): «Necesidad de un diccionario estructural», en *Actas del Congreso Internacional de Semántica*, Universidad de La Laguna, 1997, Madrid, Ediciones Clásicas, 379-390.
- Fellbaum, C. (1998): *WordNet: An Electronic Lexical Database*, Cambridge (Massachusetts), MIT Press.
- Fernández-Pampillón Cesteros, A. y Matesanz del Barrio, M. (2006): «Los Diccionarios Electrónicos: hacia un nuevo concepto de diccionario», *Estudios de Lingüística del Español*, 24.
- Fillmore, C. (1968): «The case for case», en E. Bach y R. Harms (eds.), *Universals in Linguistic Theory*, NY, Holt Rinehart and Winston, 1-88.
- (1982): «Frame Semantics», en *Linguistics in the Morning Calm*, The Linguistics Society of Korea, Seoul, Hanshin, 111-137.
- Fodor, J. y Lepore, E. (2002): «The Pet Fish and the Red Herring: Why Concepts Still Can't Be Prototypes», en J. Fodor y E. Lepore (eds.), *The Compositionality Papers*, Oxford, Clarendon, 27-42.
- Garrido Medina, J. (1990): «Diccionario español de discurso: fundamentos de diseño», *Revista española de lingüística aplicada*, 6, 31-42.
- Geckeler, H. (1976): *Semántica estructural y teoría del campo léxico*, Madrid, Gredos.
- (2002): «Anfänge und Ausbau des Wortfeldgedankens», en H. Wiegand (ed.), vol. I, 713-737.
- Geeraerts, D. (1985): «Les dones stéréotypiques, prototypiques et encyclopédiques dans le dictionnaire», *Cahiers de Lexicologie*, 46:1, 27-43.
- (1989): «Principles of Monolingual Lexicography», en J. Hausmann (ed.), vol. I, 287-296.
- (2001): «The definitional practice of dictionaries and the Cognitive Semantic conception of polysemy», *Lexicographica*, 17, 6-21.
- Gelpí Arroyo, C. (2003): «El estado actual de la lexicografía: los nuevos diccionarios», en A. Medina Guerra (ed.), *Lexicografía española*, Barcelona, Ariel, 307-328.
- Gross, G. (1994): «Classes d'objects et description des verbes», *Langages*, 116, 48-59.
- Gruber, J. (1965): *Studies in Lexical Relations*, Tesis doctoral, MIT. [Reimpreso en *Lexical Structures in Syntax and Semantics*, 1976, Amsterdam, North-Holland.
- Hanks, P. y Pustejovsky, J. (2005): «A Pattern Dictionary for Natural Language Processing», *Revue Française de linguistique appliquée*, 10:2, 63-82.
- Helbig, G. y Schenkel, W. (1991): *Wörterbuch zur Valenz und Distribution Deutscher Verben*, Tübingen, Max Niemeyer.
- Herbst, T. *et al.* (2004): *Valency Dictionary of English*, Berlín, De Gruyter.
- Hernández Hernández, H. (ed.) (1994): *Aspectos de lexicografía contemporánea*, Barcelona, Bibliograf.
- Hill, J. y Lewis, M. (eds.) (1997): *LTP Dictionary of Selected Collocations*, Hove, LTP.
- Hornby, A. S. (1987): *Oxford Advanced Learner's Dictionary of Current English*, Oxford, OUP.

- Jiménez Ríos, E. (2000): «Aspectos de lexicografía y traducción automática», en S. Ruhstaller y J. Prado Aragonés (eds.), *Tendencias en las investigación lexicográfica del español: El diccionario como objeto de estudio lingüístico y didáctico*, Huelva, Junta de Andalucía, Universidad de Huelva, 485-496.
- Kipper, K. et al. (2006): «A large-scale extension of VerbNet with novel verb classes», *Proceedings of the 12th EURALEX International Congress*, Turín, Italia, 2006.
- Knowles, F. E. (1989): «The Computer in Lexicography», en J. Hausmann (ed.), vol. II, 1645-1672.
- Köhler, R. (2002): «Statistische Methoden in der Lexikologie», en H. Wiegand (ed.), v. II, 953-963.
- Lara, L. F. (1989): «Problemas y métodos del significado estereotípico», en *Actes du XVIII Congrès International de Linguistique et de Philologie Romanes*, Tübingen, Max Niemeyer, 124-138.
- (1992): «El discurso del diccionario», en G. Wotjak (ed.), *Estudios de lexicología y meta-lexicografía del español actual*, Tübingen, Max Niemeyer, 1-12.
- (dir.) (1996): *Diccionario del español usual en México*, México, El Colegio de México.
- Lavid, J. (2005): *Lenguaje y nuevas tecnologías*, Madrid, Cátedra.
- Lenci, A. et al. (2000a): «SIMPLE Linguistic Specifications», Deliverable D2.2, [www.ub.es/gilcub/SIMPLE/simple.html#Specifications](http://www.ub.es/gilcub/SIMPLE/simple.html#Specifications).
- (2000b): «SIMPLE: A General Framework for the Development of Multilingual Lexicons», *International Journal of Lexicography*, 13, 249-263.
- Lenders, W. (2002): «Lexikologie und Computerlinguistik», en H. Wiegand (ed.), vol. II, 1886-1895.
- Longman Dictionaries (1987): *Longman Dictionary of Contemporary English*, Harlow.
- (1993): *Longman Language Activator: the World's First Production Dictionary*, Harlow.
- Lorente Casafont, M. (1996): «La lexicografía teòrica: els prototips de diccionaris», en *Estudis de llengua i literatura catalanes*, XXXIII, 339-363.
- Martín Mingorance, L. (1994): «La lexicología onomasiológica», en H. Hernández Hernández (ed.), 15-27.
- McArthur, T. (1982): *Longman Lexicon of Contemporary English*, Harlow, Longman.
- Mederos, H. (1994): «A propósito de la definición lexicográfica», en H. Hernández Hernández (ed.), 95-106.
- Mel'čuk, I. (1989): «Semantic primitives from the Viewpoint of the Meaning-Text Linguistic Theory», *Quaderni di semantica*, 10:1, 65-102.
- (1996): «Lexical Functions: A Tool for the Description of Lexical Relations in a Lexicon», en L. Wanner (ed.), *Lexical functions in lexicography and natural language processing*, Amsterdam, J. Benjamins, 37-102.
- Mel'čuk, I. et al. (1984-1999): *Dictionnaire explicatif et combinatoire du français contemporain*, vol. I-IV, Montréal, Les Presses de l'Université de Montréal.
- Mel'čuk, I. y Zholkovskij, A. (1985): *Tolkovo-kombinatornyj slovar' sovremennogo russkogo jazyka*, Viena, Wiener Slawistischer Almanach, Sonderband 14.
- Pascual, J. A. (1996): «La coherencia en los diccionarios de uso», en E. Forgas (coord.), *Léxico y diccionarios*, Tarragona, Universitat Rovira i Virgili, 167-198.
- (2006): «Confesiones de un aprendiz de lexicógrafo», Lliçó inaugural del curs acadèmic 2006-2007, Universitat Pompeu Fabra, Barcelona.
- Pascual, J. A. y Gutiérrez Cuadrado, J. (2007): *Diccionario Salamanca de la lengua española*, Madrid, Santillana.
- Peeters, B. (2000): «Setting the scene: Some recent milestones in the lexicon-encyclopedia debate», en B. Peeters (ed.), *The Lexicon-Encyclopedia Interface*, Amsterdam, Elsevier, 1-52.



- (ed.) (2006): *Semantic Primes and Universal Grammar*, Amsterdam, J. Benjamins.
- Petruck, M. et al. (2004): «Reframing FrameNet Data», *Proceedings of the 11th EURALEX International Congress*, Lorient, France, 405-426.
- Porto Dapena, A. (2002): *Manual de técnica lexicográfica*, Madrid, Arco Libros.
- Pustejovsky, J. (2006): «Type Theory and Lexical Decomposition», *Journal of Cognitive Science*, 6, 39-76.
- Pustejovsky, J. et al. (2004): «Automated Induction of Sense in Context», COLING 2004, Geneva.
- Putnam, H. (1975): «The meaning of 'meaning'», en *Mind, Language and Reality: Philosophical Papers*, vol. 2, Cambridge, CUP, 215-271.
- Real Academia Española (2001): *Diccionario de la Lengua Española*, 22.<sup>a</sup> ed. online: [www.rae.es/rae.html](http://www.rae.es/rae.html).
- Rosch, E. y Mervis, C. (1975): «Family Resemblances: Studies in the Internal Structure of Categories», *Cognitive Psychology*, 7, 573-605.
- Saurí, R. (2003): «Un corpus pal asturianu. Les tecnoloxíes llingüístiques na consolidación de les llingües minorizaes», *Revista de Filología Asturiana*, 3-4, 135-174.
- Shcherba, L. V. (1995) [1940]: «Towards a General Theory of Lexicography», trad. D. Farina, *International Journal of Linguistics*, 8:4, 314-350.
- Sinclair, J. (ed.) (1987): *Collins Cobuild English Language Dictionary*, Londres, Collins.
- (ed.) (1995): *Collins Cobuild English Collocations on CD-ROM*, Londres, Collins.
- Sommerfeldt, K.-E. y Schreiber, H. (1977): *Wörterbuch zur Valenz und Distribution der Substantive*, Leipzig, VEB Bibliographisches Institut.
- (1983): *Wörterbuch zur Valenz und Distribution deutscher Adjektive*, Tübingen, Max Niemeyer.
- (1983): *Wörterbuch zur Valenz und Distribution deutscher Verben*, Tübingen, Max Niemeyer.
- Subirats, C. (2004): «FrameNet Español. Una red semántica de marcos conceptuales», VI International Congress of Hispanic Linguistics, Universidad de Leipzig, octubre de 2003.
- Vossen, P. (2001): «Condensed Meaning In EuroWordNet», en P. Bouillon y F. Busa (eds.), *The Language of Word Meaning*, NY, CUP, 363-383.
- Weinreich, U. (1975): «Lexicographic Definition in Descriptive Semantics», en F. Householder y S. Saporta (eds.), *Problems in Lexicography*, Bloomington, Indiana University, 25-44.
- Wiegand, H. E. (1998): *Wörterbuchforschung. Untersuchungen zur Wörterbuchbenutzung, zur Theorie, Geschichte, Kritik und Automatisierung der Lexikographie*, Berlín, De Gruyter.
- Wierzbicka, A. (1985): *Lexicography and Conceptual Analysis*, EEUU, Ann Arbor, Karoma.
- (1987): *English Speech Act Verbs. A semantic dictionary*, Australia, Academic Press.
- Zaenen, A. y Engdahl, E. (1994): «Descriptive and Theoretical Syntax», en B. Atkins y A. Zampolli (eds.), 181-212.
- Zampolli, A. (1991): «Los bancos de datos léxicos: Bases multifuncionales de datos léxicos», en J. Vidal Beneytó (dir.), *Las industrias de la lengua*, Madrid, Fundación Germán Sánchez Ruipérez, 127-146.
- Zgusta, L. (1994): «Prototipos y lexicografía», *Voz y Letra*, 5:1, 3-13.

PRE-FINAL

*D*icento

