

## WORDNET Y LA LEXICOGRAFÍA ELECTRÓNICA

Jorge Antonio Leoni de León\*

### ABSTRACT

For its particular structure based on relationships such as synonymy, hyponymy and antonymy, WordNet, a lexicon-centered data-base supported by modern psycholinguistic theories about lexicon memory, constitutes a relational system of knowledge which can be connected to similar data-bases. From this connection, a powerful, computer-exploitable, inter-linguistic system of lexicon date-bases can be achieved.

**Key words:** computer linguistics, lexicography, WordNet, data base.

### RESUMEN

WordNet, base de datos léxico-conceptual, inspirada en teorías psicolingüísticas contemporáneas sobre la memoria léxica humana, por su particular estructura basada en relaciones tales como la sinonimia, la hiponimia y la antonimia, constituye un sistema relacional del conocimiento, el cual puede ser vinculado con otras bases de datos similares, con lo que se obtiene un poderoso sistema interlingüístico de bases de datos léxicas computacionalmente explotable.

**Palabras clave:** lingüística computacional, lexicografía, WordNet, base de datos.

### 1. Introducción

La aplicación de metodologías informáticas en la lingüística a menudo invita a repensar tanto la formalización de los elementos estudiados, como la manera más adecuada de expresar su universo de relaciones. Si nos detenemos un momento a analizar, por ejemplo, las obras lexicográficas, notaremos que en buena medida la presentación del material y su disposición en orden alfabético han respondido, históricamente, a criterios prácticos directamente relacionados con las propiedades físicas del libro y a la manera más eficiente de extraer información de una obra impresa. En otras palabras, no hay ninguna implicación en el hecho de que las entradas *recibo*, *reciclar* y *reciedumbre* estén juntas en un determinado diccionario, aparte de que las tres comparten las primeras letras que las componen.

Una de las muchas consecuencias del desarrollo de la informática en el área de las ciencias humanas es la ruptura de la forma, entendida ésta como *disposición de las cosas y figura exterior de un cuerpo*, es decir, como un *formato*. Muchas propuestas contemporáneas sobre estructura de documentos<sup>1</sup> engendran no sólo taxonomías del texto, sino también informaciones metatextuales sobre la índole del documento y su organización, entre otros aspectos<sup>2</sup>. En el caso de la lexicografía, el replanteamiento de la organización de los lexemas ha producido la idea de que el orden puede ser un elemento significativo en el diccionario, lo que por su parte responde a algunas teorías psicolingüísticas del léxico, donde, al contrario de la colocación alfabética de las entradas, la posición de los lexemas tiene implicaciones sobre algunos rasgos semánticos que las distinguen; es decir, que nos encontramos ante una noción de *diccionario* como sistema.

\* Profesora de la Escuela de Filología, Lingüística y literatura, Universidad de Costa Rica.

Desde hace varios años, la lexicografía ha tenido un gran impulso en Costa Rica gracias a la gran labor realizada en la línea de investigación *Estudios de Lexicografía Hispánico-Costarricense* (ELEXHICÓS), bajo la guía del Dr. Víctor Ml. Sánchez Corrales, en la Universidad de Costa Rica. Dentro del ámbito de actividades de dicha línea de investigación no se ha descartado en ningún momento la búsqueda de nuevas aproximaciones teóricas y tecnológicas al quehacer lexicográfico, lo cual ha llevado a la propuesta de nuevas aplicaciones de la lexicografía en nuestro país, tal y como está considerado en el proyecto *Sistema Léxico Virtual* (N° 745-A1-827 en la Vicerrectoría de Investigación, proyecto que cuenta además con el apoyo directo de la empresa Word Magic Software, que pretende encontrar propuestas teóricas y metodológicas para el tratamiento automático de la desambiguación.

Dentro de la larga serie de investigaciones sobre técnicas de desambiguación, las metodologías basadas en WordNet<sup>3</sup> ocupan un lugar especial que merece nuestra atención. Es dentro de este espíritu que tiene lugar la investigación sobre las propiedades de WordNet y su aplicación a la lengua española.

A pesar de la importancia de WordNet y del hecho de que hay algunos proyectos para la construcción de sistemas similares en español<sup>4</sup>, en nuestra lengua hay relativamente poca bibliografía al respecto por lo que se hace necesario disponer de una descripción general de tal sistema que a la vez brinde algunos indicios sobre la metodología y aplicación de los conceptos teóricos; con el presente artículo esperamos llenar, de alguna manera, ese vacío.

El texto conocido como *Five papers on Wordnet* (Miller *et al* 1990) brinda una excelente descripción general de WordNet, la cual abarca todas sus etapas de creación desde los archivos de los lexicógrafos hasta su codificación y representación en una base de datos, por lo cual se ha constituido en un manual de referencia obligatoria para todos aquellos interesados en profundizar en el tema.<sup>5</sup> Por otra parte, Fellbaum (1998) presenta una excelente recopilación de artículos divididos en tres aspectos: a) descripción de la base de datos; b) métodos y perspectivas de

WordNet; y c) aplicaciones. Sobre el trabajo que se ha realizado hasta la fecha en español y otras lenguas ibéricas, Escudero y Rigau (1999) es un buen punto de partida; de gran interés son, sin duda alguna, Díez Orzas (1999) y Climent Roca (2000), ambos tienen excelentes artículos sobre las relaciones de meronimia en la lengua española, además de un resumen de proyectos en otros idiomas en los cuales se considera dicha relación semántica.

En este artículo pretendemos una descripción general de los principios sobre los que se basa WordNet, probablemente el más conocido de los proyectos de lexicografía electrónica.

Hemos considerado tratar en la sección 2 el universo de las principales relaciones estructurales presentes en WordNet; mientras que en la 3 describimos como se configura el sistema dentro de las principales categorías gramaticales; la sección 4 está dedicada a una reflexión general sobre la utilidad de WordNet.

## 2. Aspectos generales

WordNet es una base de datos léxica estructurada a partir de las principales relaciones conceptuales que vinculan entre sí a los lexemas intra- y trans-categorialmente; algunas relaciones, como la hiponimia, tienen como consecuencia la formación de un sistema jerárquico en el que cada lexema ocupa una posición que le es propia y que además tiene consecuencias directas en la atribución de sentido<sup>6</sup>. En WordNet, el léxico está dividido en cinco categorías gramaticales: sustantivos, verbos, adjetivos, adverbios y palabras funcionales; sin embargo, por el momento, sólo están incluidos los sustantivos, los verbos, los adjetivos y los adverbios.

WordNet trata de dar respuesta práctica a algunas de las preguntas planteadas por la semántica léxica, la cual postula que debe existir una asociación convencional entre un concepto lexicalizado y una expresión que cumple una función sintáctica. Dichos planteamientos se refieren al tipo de expresiones que participan en esas asociaciones, a cuál es la naturaleza y organización de los conceptos lexicalizados que las

palabras pueden expresar y al tipo de funciones sintácticas que las palabras tienen.

Una forma de concebir la base de datos de WordNet es a través de una matriz léxica en la que se asocian las formas léxicas con los sentidos. La Tabla 1 presenta un esquema de este concepto:<sup>7</sup>

TABLA 1  
Matriz léxica

Sentidos	Formas léxicas				
	F <sub>1</sub>	F <sub>2</sub>	F <sub>3</sub>	...	F <sub>n</sub>
S <sub>1</sub>	L <sub>1,1</sub>	L <sub>1,2</sub>			
S <sub>2</sub>		L <sub>2,2</sub>			
S <sub>3</sub>			L <sub>3,3</sub>		
....				...	
S <sub>i</sub>					L <sub>i,n</sub>

En la Tabla 1, la sinonimia, representada horizontalmente, está determinada por la cantidad de F<sub>n</sub> (*formas léxicas*) para cada S<sub>i</sub> (*sentidos*); mientras que la polisemia, establecida verticalmente, se visualiza a partir del número de sentidos (S<sub>i</sub>) para cada forma léxica (F<sub>n</sub>). Así, L<sub>1</sub> y L<sub>2</sub> son sinónimos; por otra parte L<sub>2</sub> es polisémica.

De esta forma, es posible afirmar que el sentido S<sub>1</sub> puede ser representado por la lista de formas léxicas (F<sub>n</sub>) que pueden ser usadas para expresarlo dentro de lo que se denomina como *conjunto sinónimo*:<sup>8</sup> {F<sub>1</sub>, F<sub>2</sub>,...}<sup>9</sup>, en (1) tenemos un ejemplo, en el que *bed* y *bottom* son sinónimos:

- (1) {bed, bottom, depression,@ (a depression forming the ground under a body of water; "he searched for treasure on the ocean bed")}

Los conjuntos sinónimos no explican lo que los conceptos son, simplemente significan que los conceptos existen. Cuando no pueda ser encontrado un sinónimo adecuado, el problema de la polisemia puede ser resuelto con una glosa como se puede apreciar en (2):

- (2) {board,(a person's meals, provided regularly for money)}<sup>10</sup>

Las llaves "{" y "}" contienen los conjuntos de sinónimos que sirven para identificar las definiciones de los conceptos lexicalizados; la glosa se encuentra entre paréntesis.

## 2. Las relaciones estructurales

Los principales tipos de relaciones usados para la creación de WordNet son los siguientes:

**Sinonimia:** dos términos (o más) son considerados sinónimos si dado un contexto lingüístico en el que uno de los dos aparezca, su sustitución por el otro (o los otros) términos no producen alteración en los valores de verdad.<sup>11</sup> La sinonimia es entendida como un continuum a lo largo del cual la similaridad de sentido puede ser graduada. En (3) tenemos un ejemplo de un conjunto sinónimo:

- (3) {universe, existence, nature, creation, world, cosmos, macrocosm, noun.Tops:natural\_object,@ (everything that exists anywhere; "they study the evolution of the universe"; "the biggest tree in existence")}

**Antonimia:** al igual que la sinonimia, la antonimia es una relación léxica entre dos formas léxicas, no se trata, pues, de una relación semántica entre sentidos. No todos los lexemas tienen antónimos. La antonimia es entendida de la manera siguiente: el antónimo de una palabra *X* es *no-X*. La antonimia es el principio básico para la organización de los adjetivos. En (4) tenemos dos adjetivos antónimos entre corchetes con el signo final de exclamación que marca la antonimia:

- (4) {[ [ IPSILATERAL, CONTRALATERAL,! ] (on or relating to the same side (of the body))}

**Hiponimia:** la hiponimia es una relación entre sentidos. Un término *X* es un hipónimo de *Y* si un hablante nativo acepta la aserción "*X* es un (*tipo*) de *Y*". La relación puede ser representada incluyendo en {*x*, *x'*...} un puntero hacia sus hipónimos. Un ejemplo de hiponimia es el siguiente: un árbol es una planta, árbol es hipónimo de planta.

Dentro de este tipo de relaciones es válida la afirmación de que si hay una relación X/Y entonces debe haber necesariamente una relación Y/X; esta última se denomina *hiperonimia* y se refiere a los términos superordinados. Las relaciones de hiperonimia tienen varios elementos interesantes, el primero de los cuales se refiere al concepto de herencia de rasgos de los rasgos de los elementos superiores por parte de los inferiores, es decir, si árbol es una planta y si planta es un organismo vivo, entonces planta es un organismo vivo. Por otra parte, se supone que los elementos inferiores de la jerarquía así formada son los portadores de rasgos sintácticos. En (5) encontramos que los sinónimos *smoke* y *fume* son hipónimos de aerosol, lo cual se indica por cuanto éste último está seguido de la marca @; aerosol sería, como consecuencia, un hiperónimo de dichos términos:

- (5) {smoke, fume, aerosol,@ (a cloud of fine particles suspended in a gas)}

**Meronomia:** Meronomia es la relación parte-todo entre dos conceptos; un concepto Y es merónimo de un X si un hablante nativo acepta que “X tiene un Y (como parte)” o “Y es parte de X”. Al igual que en la hiponimia, si hay relación X/Y, entonces debe haber también una relación Y/X, la cual se denomina holonimia. La relación meronímica es transitiva, asimétrica y utilizable para construir ciertas jerarquías.<sup>12</sup> En WordNet códigos #m, #s y #p marcan los miembros merónimos, las

substancias merónimas y las partes merónimas respectivamente. En los ejemplos siguientes tenemos que el conjunto sinónimo {locomotive, ...} es miembro merónimo de *train*; {belting} es sustancia merónima de *belt* y de *belt1* (además de ser hipónimo de *material*) y {adit} es parte merónima de *mine1*:

- (6) *Hiperónimos:*  
 a. {locomotive, engine1, locomotive\_engine, railway\_locomotive, wheeled\_vehicle,@ train,#m (self-propelled engine used to draw trains along railway tracks)}  
 b. {belting, material,@ belt,#s belt1,#s (the material of which belts are made)}  
 c. {adit, passage,@ mine1,#p (a nearly horizontal passage from the surface into a mine)}

Es necesario recordar que la antonimia se da dentro de una misma clase de categorías, la meronomia puede ser transcategorial y la hiponimia estructura categorías; la sinonimia existe en todas las entradas léxicas, siendo la unidad el mínimo conjunto sinónimo posible, como se puede ver en (2). Por otra parte, hay más relaciones que son aclaradas en la sección siguiente.

### 3. Las categorías

En la Tabla 2 se muestran las relaciones léxicas establecidas según la categoría:<sup>13</sup>

TABLA 2  
Relaciones léxicas en WordNet según categoría

Símbolo	Relación	Pertinencia por categoría gramatical			
		S	V	A	Adv
!	Antonimia	✓	✓	✓	✓
@	Hiperonimia	✓	✓	x	x
~	Hiponimia	✓	✓	x	x
#m	Meronomia (miembro)	✓	x	x	x
#s	Meronomia (substancia)	✓	x	x	x
#p	Meronomia (parte)	✓	x	x	x
%m	Holonimia (miembro)	✓	x	x	x
%s	Holonimia (substancia)	✓	x	x	x
%p	Holonimia (parte)	✓	x	x	x
=	Atributo	✓	x	✓	x

Símbolo	Relación	Pertinencia por categoría gramatical			
		S	V	A	Adv
*	Implicación estricta	X	✓	X	X
>	Causa	X	✓	X	X
^	Ver también	X	✓	✓	X
\$	Grupo verbal	X	✓	X	X
&	Similitud	X	X	✓	X
<	Participio	X	X	✓	X
\	lexema relativo a	X	X	✓	X
\	Derivado de un adjetivo	X	X	✓	X

En la columna de símbolos tenemos los identificadores para cada relación, según están codificados en la base de datos. En la segunda columna se describe brevemente el tipo de relación; las últimas columnas corresponden al tipo de categoría. A continuación damos una breve descripción de cada categoría.

### 3.1. Sustantivos

La relación dominante entre los sustantivos es la hiponimia, la cual se forma a partir de los 25 conceptos *primitivos* mostrados en la Tabla 3:

TABLA 3  
Lista de conceptos primitivos para los sustantivos en WordNet

{act, action, activity}	{natural object}
{animal, fauna}	{natural phenomenon}
{artifact}	{person, human being}
{attribute, property}	{plant, flora}
{body, corpus}	{possession}
{cognition, knowledge}	{process}
{communication}	{quantity, amount}
{event, happening}	{relation}
{feeling, emotion}	{shape}
{food}	{state, condition}
{group, collection}	{substance}
{location, place}	{time}

Es a partir de estos conceptos que se inicia la jeraquización de los sustantivos. Estos *primitivos*, claro está, se descomponen en otros conceptos según sea necesario para afinar las relaciones

conceptuales de un término. En (7) tenemos los hiperónimos para el término inglés *roadster*:

- ```
(7) roadster
    @ → car
        @ → motor vehicle
            @ → wheeled vehicle
                @ → vehicle
                    @ → conveyance
                        @ → artifact
```

Es necesario señalar que aunque en principio a cada hipónimo no le corresponde más un hiperónimo, hay un cierto número de *hipónimos* conjuntivos, los cuales existen con el fin de incluir una entrada nueva, que únicamente se distingue de otra por una sola ramificación. Generalmente se trata de términos muy generales. A continuación un ejemplo de una entrada:

- ```
(8) {[man, woman!], person, @ ... (a male person)}
    {[woman, man!], person, @ ... (a female person)}
```

El signo de admiración marca la relación de antonimia entre dos conceptos, la arroba indica el superordinado y la frase entre paréntesis corresponde a la glosa.

Los atributos (marcados con =) señalan ciertas características propias de determinados sustantivos, tales como color o tamaño. Esta característica no está completamente introducida en WordNet, aunque en algunos casos es fuente de información importante relacionada con un sustantivo que nombra el atributo y todos los adjetivos relacionados con éste; esto corresponde al fenómeno de marcaje dentro de los adjetivos.

### 3.2. Adjetivos

Los adjetivos presentes en WordNet se clasifican de la siguiente manera:

*i. Descriptivos:* estos le atribuyen un valor a un sustantivo. Se pueden plantear también de la siguiente forma: un adjetivo es un adjetivo descriptivo si cumple con la siguiente fórmula  $A(x) = \text{Adj}$ , la cual en la práctica se expresa en la práctica como en (9):

- (9) PESO (paquete) = liviano

La organización de estos adjetivos se basa principalmente en la antonimia. En muchas ocasiones no existe un antónimo directo para un término dado, pero sí hay un grupo de términos asociados con uno que sí posee un antónimo, esta relación se denomina antonimia indirecta. Esta posee la forma de un grupo de adjetivos asociados a un término adjetivo principal (es decir que sí posee antónimo) los cuales son antónimos indirectos del término o los términos asociados con su antónimo, esto se muestra en (10):

- (10) *wet* y *dry* son antónimos directos, sin embargo sus términos asociados *watery* y *sere* (asociados respectivamente a los términos *wet* y *dry*), no son antónimos, pero se convierten en antónimos indirectos por estar asociados a dos términos que sí son directamente sinónimos.<sup>14</sup>

*ii. Modificadores de referencia:* ésta es una clase cerrada de adjetivos que pueden hacer variar la interpretación del sintagma verbal; por ejemplo, el sustantivo *amigo* en *mi viejo amigo* y *mi nuevo amigo* se encuentran modificados por las características semánticas de los sintagmas adjetivos *nuevo* y *viejo*. En WordNet los adjetivos modificadores de referencia llevan, en su mayoría, marcas de realización, las cuales indican que estos se manifiestan como atributos y no como predicados; el marcaje indica que siempre se encuentran dentro del sintagma nominal. Es necesario resaltar el hecho de que este tipo de adjetivos funciona a menudo como adverbios:

- (11) *My former teacher...*  
*He was formerly my teacher.*

*iii. Relacionales:* estos, al igual que los adjetivos modificadores de referencia, se manifiestan generalmente como atributos y no como predicados. Este es el caso de adjetivos como *fraternal* y *dental*:

- (12) a. La higiene dental es necesaria  
b. \* La higiene es dental

Estos adjetivos a menudo se derivan de sustantivos, no tienen gradación ni tampoco antónimos directos, por lo que no pueden ser incorporados en los grupos de adjetivos que caracterizan a los adjetivos descriptivos. En WordNet estos adjetivos tienen un puntero que los vincula con los sustantivos correspondientes.

En el marcaje de los adjetivos, la base de datos cuenta con las marcas siguientes de restricción sintáctica:

- (p): indica que sólo se realiza como predicado.  
(a): pronominal, estos adjetivos siempre se encuentran junto al sustantivo que modifican.  
(ip): postnominales, es decir, que los adjetivos que llevan esta marca siempre están en posición postnominal

Para hacer un paralelismo con el español, podemos señalar que, por ejemplo, en un WordNet español, debería haber dos entradas para el adjetivo cierto, una con la marca (a) y otra con (ip), lo cual se ve claramente en (13):

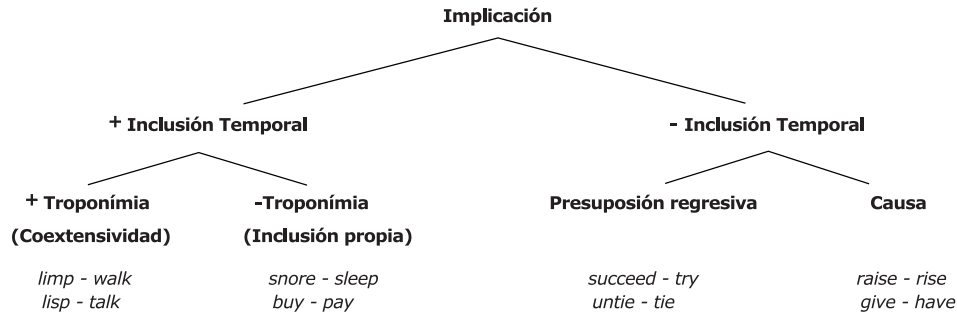
- (13) a. Un cierto asunto  
b. Un asunto cierto

En ellos se ve la variación del sentido según la posición del lexema cierto.

### 3.3. Verbos

El sistema léxico verbal se basa en la implicación que cada verbo tiene dentro de un universo de relaciones, el cual se presenta resumido en la Figura 1:

FIGURA 1  
Relaciones de Implicación en los Verbos



La implicación (estricta) está definida para las proposiciones dentro de los términos de (14):

- (14) *Implicación estricta:*  
Una proposición P implica a una proposición Q si y sólo si no existe un estado concebible de las cosas que puedan hacer verdadero a P y falso a Q.

(14) es una relación semántica porque implica referencia a un estado de las cosas que P y Q representan. El término fue generalizado para referirse a la relación que dos verbos V1 y V2 tienen cuando la oración *alguien V1* implica lógicamente la oración *alguien V2*; esto se denomina implicación léxica.

La implicación léxica es una relación unilateral: si un verbo V1 implica a otro verbo V2, entonces, no puede existir el caso en el que V2 implique V1. La excepción es que cuando dos verbos pueden implicarse mutuamente, entonces son sinónimos. En (15) tenemos otro aspecto por considerar, el cual se refiere a la inclusión temporal:

- (15) *Inclusión temporal:*  
Un verbo V1 incluye un verbo V2 si existe algún lapso durante el cual las actividades denotadas por los dos verbos co-ocurren, pero no existe un lapso durante el cual V2 puede darse y V1 no. Si existiere un lapso durante el cual V1 tiene lugar, pero V2 no, entonces V1 incluye propiamente a V2.

Esto nos lleva a la generalización en (16):

- (16) *Inclusión temporal:*  
Si V1 implica a V2, y si una inclusión temporal se lleva a cabo entre ellos, entonces los hablantes aceptarán una relación de parte-todo que relaciona a V2 y a V1.

Tenemos, por ejemplo, el caso de la relación entre las palabras roncar/dormir, en la cual hay inclusión temporal:

- (17) roncar / dormir: roncar (V1) implica dormir (V2), puesto que no se puede roncar sin estar durmiendo, por tanto, hay una relación temporal parte-todo entre roncar (V1) y (V2).

La relación de hiponimia en los verbos está planteada más bien en términos de troponimia (del griego *tropos*, manera); ésta se define en los siguientes términos:

- (18) *Troponimia:*  
V1 es V2 de una manera particular.

La troponimia es un tipo muy particular de implicación en el que cada tropónimo V1 de un verbo V2 más general que también implica a V2. El verbo inglés *to lisp* implica *to talk*, en (19) vemos, por ejemplo, la relación de *to understand* y *to lisp*, alguien habla de una manera particular difícil de entender:

- (19) I can't understand her because she lisps.

Un aspecto importante en la taxonomía verbal es la relación de *presuposición regresiva*, planteada en (20):

- (20) *Presuposición regresiva:*  
La actividad denotada por el verbo implicado siempre precede, en tiempo, a la actividad denotada por el verbo que implicante.

Un verbo V1 que es implicado por otro verbo V2 por presuposición regresiva no puede ser concebido como parte de V2. La relación parte-todo entre verbos es posible sólo cuando exista una inclusión temporal entre estos verbos. La presuposición regresiva es complementada, de cierta manera, por las relaciones de causalidad definidas en (21):

- (21) *Causalidad:*  
Si V1 necesariamente causa V2, entonces V1 también implica a V2.

La relación de causalidad es unidireccional:

- (22) *dar / tener:* dar (V1) necesariamente causa que el benefactor *tenga* (V2), pero tener no implica, de ninguna manera, *dar*. Sin embargo dar (V1) implica a tener (V2).

WordNet fue diseñado como memoria léxica, más que como una representación léxica del conocimiento, por lo que mucho del conocimiento del hablante sobre propiedades semánticas y sintácticas de los verbos fue excluido. Sin embargo, hay que notar que la subcategorización se encuentra presente como una referencia a un cuadro de reglas sintácticas<sup>15</sup>.

#### 4. A propósito de WordNet

WordNet, como se ha podido ver hasta el momento, constituye una base de datos léxico-relacional que simula, parcialmente, la memoria humana. Es muy tentador referirse a las posibles aplicaciones de este instrumento. Entre las más obvias, tenemos la creación de diccionarios electrónicos multilingües, lo que sin duda sería una herramienta de primera mano en el aprendizaje de lenguas extranjeras. Sin embargo, no hay que dejar de lado que también es posible explotar el conjunto de relaciones dentro de la base datos misma y luego generalizarlas a otros idiomas.

Las aplicaciones en este caso, serían múltiples y sin duda aún no hemos alcanzado a imaginar todo lo que sería posible con WordNet.

Entre las primeras posibilidades que se han previsto, se encuentra la desambiguación automática de textos, como una herramienta que posee información semántica pertinente sobre los sintagmas. Otras opciones no deben dejarse de lado, como por ejemplo, el hecho de que WordNet ya ofrece una taxonomía del conocimiento, la cual puede utilizarse como referencia para labores de indexación, por ejemplo. Tampoco hay que olvidar las posibles aplicaciones en la educación y el aprendizaje de lenguas extranjeras. Sobre estas y otras formas de utilizar WordNet esperamos referirnos en un artículo futuro.

#### 6. Notas

1. Literarios, científicos, comerciales, etc.
2. Esto será tratado con más detalle en un artículo futuro.
3. Para más detalles: Claudia Leacock y Martin Chodorow (1998) y Resnik (1998).
4. Ver Vossen (2001).
5. El Profesor George A. Miller es el *investigador principal* de este proyecto en el *Laboratorio de Ciencias Cognoscitivas* de la Universidad de Princeton.
6. Esto sin duda trae a la memoria la concepción de la lengua como sistema de signos según de Saussure.
7. Basado en Miller *et al.*(1990: 4)
8. Los creadores de WordNet alegan que el uso de conjuntos sinonímicos para representar los sentidos de las palabras es consistente con evidencia sobre la organización, en memoria léxica, de los conjuntos de sustantivos, verbos y modificadores (Miller *et al.* 1990).
9. Las llaves “{“ y “}” contienen los conjuntos de sinónimos que sirven para identificar las definiciones de los conceptos lexicalizados.
10. Miller (1990: 6).



11. Esto implica que términos de categorías gramaticales diferentes no pueden ser sinónimos puesto que no pueden ser (sintácticamente) intercambiables.
12. Para obtener mayor información sobre la relación meronímica en el léxico español, ver Climent Roca (2000) y Díez Orzas (1999).
13. ✓ es una relación presente en esa categoría, ✕ es una relación ausente.
14. Ver Miller *et al.* (1990:29)
15. Se trata básicamente de la subcategorización y del tipo de sujeto.

## 7. Bibliografía

- Claudia Leacock y Martin Chodorow. 1998. "Combining Local Context and WordNet Similarity for Word Sense Identification". En: Fellbaum, Christiane (ed.). *WordNet: and Electronic Lexical Database*. Cambridge, Massachusetts: MIT Press: pp. 265-274.
- Climent Roca, Salvador. 2000. "Individuación e información Parte-Todo: Representación para el procesamiento computacional del lenguaje". En: *Estudios de lingüística española* Vol. 8 <http://elies.rediris.es/elies8/> [24 de mayo de 2002]
- Díez Orzas, Pedro Luis. 1999. "La relación de meronimia en los sustantivos del léxico español: contribución a la semántica computacional". En: *Estudios de lingüística española*. Vol.2. <http://elies.rediris.es/elies2/> [24 de mayo de 2002]
- Escudero, Gerard y Rigau, German. 1999. *Web EuroWordNet Interface* (version 0.1): <http://nipadio.lsi.upc.es/wei.html> [21 de mayo de 2002]
- Fellbaum, Christiane. 1998. *WordNet: An Electronic Lexical Database*. Cambridge, Massachusetts: MIT Press.
- Miller, George A., Richard Beckwith, Christiane Fellbaum, Derek Gross and Katherine J. Miller. "Introduction to WordNet: an online lexical database". En: *International Journal of Lexicography* 3 (4), 1990, pp. 235 - 244. <ftp://ftp.cogsci.princeton.edu/pub/wordnet/5papers.pdf> [24 de mayo de 2002]
- Resnik, Philip. 1998. "WordNet and Class-Based Probabilities". En: Fellbaum, Christiane (ed.). *WordNet: and Electronic Lexical Database*. Cambridge, Massachusetts: MIT Press: pp. 240-263.
- Vossen, Piek. 2001. *EuroWordNet: Building a multilingual database with wordnets for several European languages*: <http://www.hum.uva.nl/~ewn/index.html> [25 de marzo de 2002]