

STUDIEN ZUR ROMANISCHEN
SPRACHWISSENSCHAFT UND
INTERKULTURELLEN KOMMUNIKATION

Herausgegeben von Gerd Wotjak

BAND 115



PETER LANG
EDITION

M^a José Domínguez Vázquez /
M^a Teresa Sanmarco Bande (ed.)

Lexicografía y didáctica

Diccionarios y otros recursos lexicográficos en el aula



PETER LANG
EDITION

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation
in der Deutschen Nationalbibliografie; detaillierte bibliografische
Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Library of Congress Cataloging-in-Publication Data

Names: Domínguez Vázquez, María José, 1972- | Sanmarco Bande, M^a. Teresa
(María Teresa), 1970-

Title: Lexicografía y didáctica : diccionarios y otros recursos
lexicográficos en el aula / M^a. José Domínguez Vázquez, M^a. Teresa
Sanmarco Bande (ed.).

Description: New York : Peter Lang, 2017. | Series: Studien zur romanischen
Sprachwissenschaft und interkulturellen Kommunikation ; Vol. 115

Identifiers: LCCN 2017002911 | ISBN 9783631664483

Subjects: LCSH: Language and languages—Study and teaching. | Lexicography. |
Encyclopedias and dictionaries, Spanish.

Classification: LCC P53.2963 L49 2017 | DDC 413.028—dc23

LC record available at <https://lcn.loc.gov/2017002911>

Gedruckt auf alterungsbeständigem, säurefreiem Papier.
Druck und Bindung: CPI books GmbH, Leck

ISSN 1436-1914

ISBN 978-3-631-66448-3 (Print)

E-ISBN 978-3-653-05627-3 (E-PDF)

E-ISBN 978-3-631-70879-8 (EPUB)

E-ISBN 978-3-631-70880-4 (MOBI)

DOI 10.3726/978-3-653-05627-3

© Peter Lang GmbH

Internationaler Verlag der Wissenschaften

Frankfurt am Main 2017

Alle Rechte vorbehalten.

Peter Lang Edition ist ein Imprint der Peter Lang GmbH.

Peter Lang – Frankfurt am Main · Bern · Bruxelles ·

New York · Oxford · Warszawa · Wien

Das Werk einschließlich aller seiner Teile ist urheberrechtlich
geschützt. Jede Verwertung außerhalb der engen Grenzen des

Urheberrechtsgesetzes ist ohne Zustimmung des Verlages
unzulässig und strafbar. Das gilt insbesondere für

Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die
Einspeicherung und Verarbeitung in elektronischen Systemen.

Diese Publikation wurde begutachtet.

www.peterlang.com

Índice

Introducción	7
--------------------	---

I. La lexicografía en la era digital

Rufus H. Gouws

La sociedad digital y los diccionarios	17
--	----

Mercè Lorente Casafont

La tecnología al servicio de la lexicografía y la didáctica: herramientas para la implementación de las obras lexicográficas y mejora en la presentación de contenidos	35
--	----

Carolin Müller-Spitzer

Investigación sobre el uso de diccionarios en la era digital	55
--	----

Félix San Vicente Santiago

El diccionario bilingüe: contenidos y finalidad	81
---	----

II. Los recursos lexicográficos en el aula

1. Aplicaciones del diccionario

Dolores Azorín Fernández y M^o Isabel Santamaría Pérez

El diccionario monolingüe en el aula	109
--	-----

Ana Lourdes de Hériz Ramón

El diccionario bilingüe	133
-------------------------------	-----

Concepción Maldonado González

El diccionario didáctico monolingüe y bilingüe	161
--	-----

María José Domínguez Vázquez

Portales y diccionarios multilingües electrónicos	177
---	-----

Cecilio Garriga Escribano

El diccionario de lenguajes sectoriales	203
---	-----

2. Herramientas lingüísticas con fines lexicográficos

<i>Pablo Gamallo Otero</i> Linguakit y sus aplicaciones en el aula	229
<i>Xosé María Gómez Clemente</i> Propuesta de un MOOC para la enseñanza del léxico de una segunda lengua	247
<i>Xavier Gómez Guinovart</i> Aplicaciones de WordNet a la didáctica de lenguas	267
<i>Xosé Luís Regueira Fernández</i> El aprendizaje de la pronunciación y de la lengua oral: dos herramientas para el aula de lengua gallega	283

III. Unidades didácticas: ¿Cómo mejorar las competencias del usuario?

<i>María Álvarez de la Granja</i> ¿Qué obras lexicográficas podemos utilizar en el aula? Diversidad de herramientas y posibilidades de explotación	307
<i>Pedro Antonio Fuertes-Olivera</i> Cómo abordar en el aula la equivalencia entre lenguas	329
<i>M.^a Montserrat Muriano Rodríguez</i> El diccionario como unidad didáctica en la enseñanza del español L2	343
<i>Dolores Torres Medina</i> Informaciones sobre el uso en diccionarios bilingües español-inglés: aplicaciones didácticas de las marcas	359
<i>Carlos Valcárcel Riveiro</i> Usar el <i>Wiktionnaire</i> en el aula de francés lengua extranjera: una propuesta de actividades desde el enfoque por tareas	377
<i>María J. Valero Gisbert</i> La fraseología en la L2 a través de la lexicografía bilingüe	399

Xavier Gómez Guinovart
Grupo TALG – Universidade de Vigo

Aplicaciones de WordNet a la didáctica de lenguas

Abstract: In this chapter, I will describe the characteristics of WordNet as a powerful lexicographical tool in the field of foreign language didactics, showing some illustrative examples of its applications in different tasks with Galnet, a multilingual digital lexicographical resource which extends the original WordNet lexical database of English with their correspondences in Galician, Spanish, Portuguese, Catalan and Basque.

Keywords: WordNet, computational lexicography, foreign language didactics, linguistic resources

1. Introducción

La red semántica WordNet es un recurso léxico computacional diseñado originalmente con una perspectiva psicolingüística (Fellbaum 1998), si bien orientado mayoritariamente a su aplicación en el campo del procesamiento del lenguaje, en tareas de desambiguación semántica automática (Agirre / Edmonds 2006), de recuperación de la información (Zhao *et al.* 2012), de generación de respuestas a preguntas (Cai *et al.* 2016), de traducción automática (Vintar / Fišer / Vrščaj 2012), de recuperación interlingüística de la información (Agirre *et al.* 2007), de búsqueda plurilingüe de respuestas (Ferrández *et al.* 2007), de clasificación automática de textos (Elberrichi *et al.* 2008), de expansión de consultas (Fang 2008), de corrección ortográfica (Huang 2016) o de resumen automático (Plaza *et al.* 2010). WordNet también ha sido utilizado con provecho en aplicaciones computacionales relacionadas con el aprendizaje de idiomas, por ejemplo, en la evaluación de la competencia léxica de estudiantes de inglés como segunda lengua (Hu / Graesser 1998), en la generación automática de ejercicios de vocabulario de elección múltiple para teléfonos móviles inteligentes (Knoop / Wilske 2013) o en el diseño de sistemas de aprendizaje del vocabulario asistido por ordenador (Sun / Huang / Liu 2011).

En este capítulo, describiremos las características de WordNet que lo convierten en una potente herramienta lexicográfica en el ámbito de la didáctica de lenguas, centrándonos en algunos ejemplos ilustrativos de su uso obtenidos a través de la interfaz web de Galnet (Solla Portela / Gómez Guinovart 2015a), un

recurso lexicográfico digital plurilingüe de libre acceso, elaborado en el grupo de investigación TALG de la Universidade de Vigo, que amplía la red léxico-semántica original del WordNet del inglés con sus correspondencias en gallego, español, portugués, catalán y vasco.

2. El caudal léxico de WordNet

WordNet es una base de datos léxica del inglés configurada como una red semántica donde los nodos son los conceptos representados como grupos de sinónimos, y los enlaces entre los nodos son las relaciones semánticas entre los conceptos léxicos. Los nodos de la red están formados por nombres, verbos, adjetivos y adverbios agrupados por su sinonimia. En la terminología propia de WordNet, cada grupo de sinónimos constituye un *synset*, y cada sinónimo lematizado que forma parte del grupo es una *variante* léxica del concepto. Así, un *synset* representa un concepto lexicalizado único y agrupa el conjunto de variantes sinonímicas de ese concepto. Como complemento de cada *synset*, WordNet puede incluir una breve definición distintiva (o *glosa*) del significado compartido por todas las variantes del *synset* y, en ciertos casos, ejemplos de uso de las variantes en contexto.

En el modelo de representación del léxico empleado por WordNet, los *synsets* están conectados por relaciones léxico-semánticas. En el caso de los sustantivos, algunas de las relaciones más frecuentes representadas en WordNet son las de hiperonimia / hiponimia y las de holonimia / meronimia; en el caso de los adjetivos, las de antonimia y las de cuasisinonimia; en el caso de los adverbios, las de antonimia y las derivativas; y en el caso de los verbos, las de implicación, hiperonimia / hiponimia, causatividad y oposición.

WordNet fue concebido originalmente para la lengua inglesa y, aunque hoy existen versiones de WordNet en muchas lenguas, el WordNet del inglés sigue siendo hasta ahora la versión de referencia y la más desarrollada. La elaboración del WordNet del inglés se viene llevando a cabo desde 1985 en la Universidad de Princeton. En su versión 3.0, el WordNet de esta lengua contiene 206.941 lemas o variantes sinonímicas (155.287 de los cuales son formas únicas no homógrafas) agrupados en 117.659 grupos de sinónimos o *synsets*.

En la actualidad existen versiones de WordNet en distintas fases de desarrollo para muy diversas lenguas, incluidas el hebreo (Ordan / Wintner 2007), el sánscrito (Bhingardive *et al.* 2014), el japonés (Isahara *et al.* 2008), el español (Fernández / Vázquez 2010), el catalán (Oliver / Climent 2011), el portugués (Simões / Gómez Guinovart 2014) y el vasco (Pociello / Agirre / Aldezabal

2011). La mayoría de las versiones en lenguas distintas del inglés siguen el modelo de diseño de EuroWordNet (Vossen 2002), en el que los *synsets* que forman parte del WordNet de la lengua propia están vinculados con los *synsets* del resto de las lenguas a través de un índice interlingüístico (*InterLingual Index* o ILI), que es único para cada concepto y que está basado principalmente en los *synsets* del WordNet inglés de referencia. De este modo, el conjunto de léxicos WordNet en los distintos idiomas permite la conexión entre los *synsets* de cualquier par de lenguas a través de su ILI, constituyendo así un recurso de gran utilidad en aplicaciones de las tecnologías lingüísticas que precisan el procesamiento plurilingüe del lenguaje. En el caso de Galnet, el WordNet del gallego, los *synsets* del gallego encuentran sus equivalentes interlingüísticos en decenas de idiomas como los mencionados a través del ILI generado a partir del WordNet 3.0 del inglés.

El marco de desarrollo en el que se integra Galnet es el *Multilingual Central Repository* (González Agirre / Rigau 2013), una plataforma que abarca en la actualidad los léxicos WordNet de seis lenguas (inglés, español, catalán, vasco, portugués y gallego), enlazadas interlingüísticamente por el ILI correspondiente al WordNet 3.0 y con los *synsets* categorizados en la jerarquía de dominios IRST (Bentivogli *et al.* 2004) y en las ontologías SUMO (Pease / Niles / Li 2002) y el *Top Concept Ontology* (Álvez *et al.* 2008). La aplicación web diseñada para la explotación de Galnet (<http://sli.uvigo.gal/galnet/>) amplía las funcionalidades de consulta ofrecidas por la interfaz de EuroWordNet (<http://adimen.si.ehu.es/cgi-bin/wei/public/wei.consult.perl>) con una nueva clasificación semántica de orientación terminológica basada en epinónimos (Solla Portela / Gómez Guinovart 2015a), con nuevas posibilidades de visualización de las relaciones semánticas entre los *synsets*, con una nueva presentación de la información asociada a los *synsets* en formato de datos abiertos enlazados mediante las aplicaciones *LodLive* y *Virtuoso Facets* (Solla Portela / Gómez Guinovart 2016) y con una herramienta específicamente diseñada para la extracción de campos léxico-semánticos.

En la Tabla 1 se presenta el estado actual de desarrollo de los léxicos de las seis lenguas, tal como se puede consultar en estos momentos mediante la interfaz web de Galnet. La tabla muestra la cantidad de *synsets* y variantes sinonímicas disponibles en cada lengua, distinguiendo entre las distintas categorías gramaticales, e incluye el progreso alcanzado por cada léxico respecto al WordNet 3.0 del inglés (en la fila rotulada como %EWN) y los datos actualizados de Galnet en su versión 3.0.21:

Tabla 1. Número de synsets y variantes sinónimas por idioma

	Inglés (WordNet 3.0)		Gallego (Galnet 3.0.21)	
	<i>variantes</i>	<i>synsets</i>	<i>variantes</i>	<i>synsets</i>
<i>Nombres</i>	146.312	82.115	34.175	24.323
<i>Verbos</i>	25.047	13.767	5.913	2.416
<i>Adjetivos</i>	30.002	18.156	9.508	5.795
<i>Adverbios</i>	5.580	3.621	849	586
<i>Total</i>	206.941	117.659	50.445	33.120
%EWN	100 %	100 %	24 %	28 %
	Español (MCR 3.0)		Portugués (MCR 3.0)	
	<i>variantes</i>	<i>synsets</i>	<i>variantes</i>	<i>Synsets</i>
<i>Nombres</i>	81.064	49.120	14.084	9.994
<i>Verbos</i>	16.803	9.018	6.184	3.770
<i>Adjetivos</i>	17.799	11.478	4.837	3.556
<i>Adverbios</i>	1.107	693	596	521
<i>Total</i>	116.773	70.309	25.701	17.841
%EWN	56 %	60 %	12 %	15 %
	Catalán (MCR 3.0)		Vasco (MCR 3.0)	
	<i>variantes</i>	<i>synsets</i>	<i>variantes</i>	<i>Synsets</i>
<i>Nombres</i>	52.045	36.763	40.943	27.063
<i>Verbos</i>	11.589	5.432	9.470	3.442
<i>Adjetivos</i>	7.941	4.362	148	111
<i>Adverbios</i>	2	1	0	0
<i>Total</i>	71.577	46.558	50.561	30.616
%EWN	34 %	40 %	24 %	26 %

3. WordNet como diccionario temático plurilingüe visual

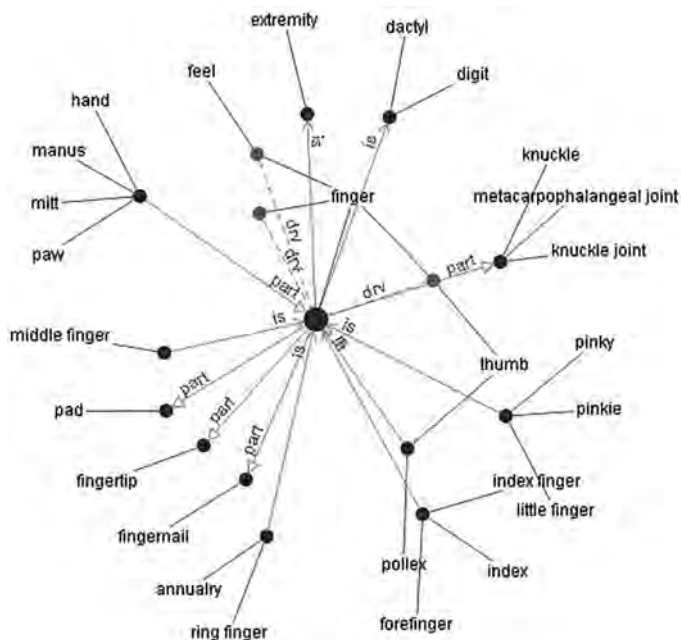
WordNet puede ser usado en el aula de idiomas como una poderosa herramienta de consulta lexicográfica plurilingüe capaz de poner en relación decenas de lenguas de muy diversas familias lingüísticas. *The Global WordNet Association* mantiene un listado por idiomas de más de cien léxicos WordNet en su página web (<http://www.globalwordnet.org>). El proyecto Open Multilingual Wordnet (Bond / Foster

2013; <http://compling.hss.ntu.edu.sg/omw/>) ofrece la posibilidad de consultar conjuntamente 34 de estos léxicos WordNet y obtener equivalencias plurilingües entre lenguas como el albanés, árabe, búlgaro, chino mandarín, danés, griego, inglés, persa, finés, francés, hebreo, croata, islandés, italiano, japonés, catalán, vasco, gallego, español, indonesio, malayo, neerlandés, noruego, polaco, portugués, rumano, lituano, eslovaco, esloveno, sueco y tailandés. La consulta conjunta de los léxicos WordNet, por tanto, permite poner en correspondencia interlingüística pares de lenguas para las que resulta difícil conseguir diccionarios bilingües, como podrían ser los pares islandés-rumano, albanés-vasco o gallego-tailandés.

En tanto que diccionario plurilingüe de consulta para el aprendizaje, WordNet complementa su amplia oferta lingüística con la posibilidad de asociar sentidos con imágenes y de trabajar con agrupaciones semánticas de palabras. La utilización de imágenes asociadas a conceptos o palabras como recurso pedagógico en la docencia de lenguas extranjeras no es en absoluto ajena a la lexicografía convencional, habiendo dado lugar a obras de gran extensión y calidad gráfica (Corbeil / Archambault 2004). El proyecto *ImageNet* (Deng *et al.* 2009) desarrollado en la Universidad de Princeton tiene como finalidad la construcción de una base de datos de imágenes asociadas a los *synsets* de WordNet. Su objetivo final es ilustrar cada *synset* de WordNet con un promedio de 1.000 imágenes cuidadosamente seleccionadas por revisores humanos. En la actualidad, la interfaz web de *ImageNet* (<http://www.image-net.org>) permite visualizar más de 14 millones de imágenes correspondientes a 32.326 *synsets*, superando ampliamente en número de imágenes cualquier otro diccionario visual realizado hasta la fecha y equiparándolo en número de conceptos. La red de relaciones léxico-semánticas entre los *synsets* de WordNet constituye otro importante elemento didáctico inherente a este recurso. Este tipo de relaciones se explota también en la lexicografía plurilingüe convencional, especialmente, en los diccionarios organizados por temas, donde los grupos de palabras suelen incluir los hipónimos de un término o sus merónimos. Por ejemplo, la entrada para *bird* del diccionario visual temático pentalingüe de Corbeil y Archambault incluye merónimos en español como *ala*, *pechuga*, *penacho* y *plumón*, e hipónimos como *estornino*, *vencejo* o *urraca* (Corbeil / Archambault 2004: 115–120). En el caso de WordNet, estas relaciones pueden visualizarse en forma de grafos dirigidos donde los vértices o nodos representan los *synsets*, y las aristas o flechas representan las distintas relaciones semánticas que se establecen entre ellos. Un ejemplo de este modo de visualización se puede observar en WordVis (<http://wordvis.com>), una aplicación web desarrollada en la NTNU (Norwegian University of Science and Technology) que permite realizar consultas léxicas sobre el WordNet del inglés y visualizar los resultados de mane-

ra interactiva en forma de conjunto de conceptos agrupados por sus relaciones semánticas. La presentación visual de estas relaciones en WordVis se ilustra en la Figura 1 con el grafo generado por esta aplicación para el concepto representado por la variante *finger* en inglés en el sentido descrito en la glosa de WordNet como «any of the terminal members of the hand (sometimes excepting the thumb)».

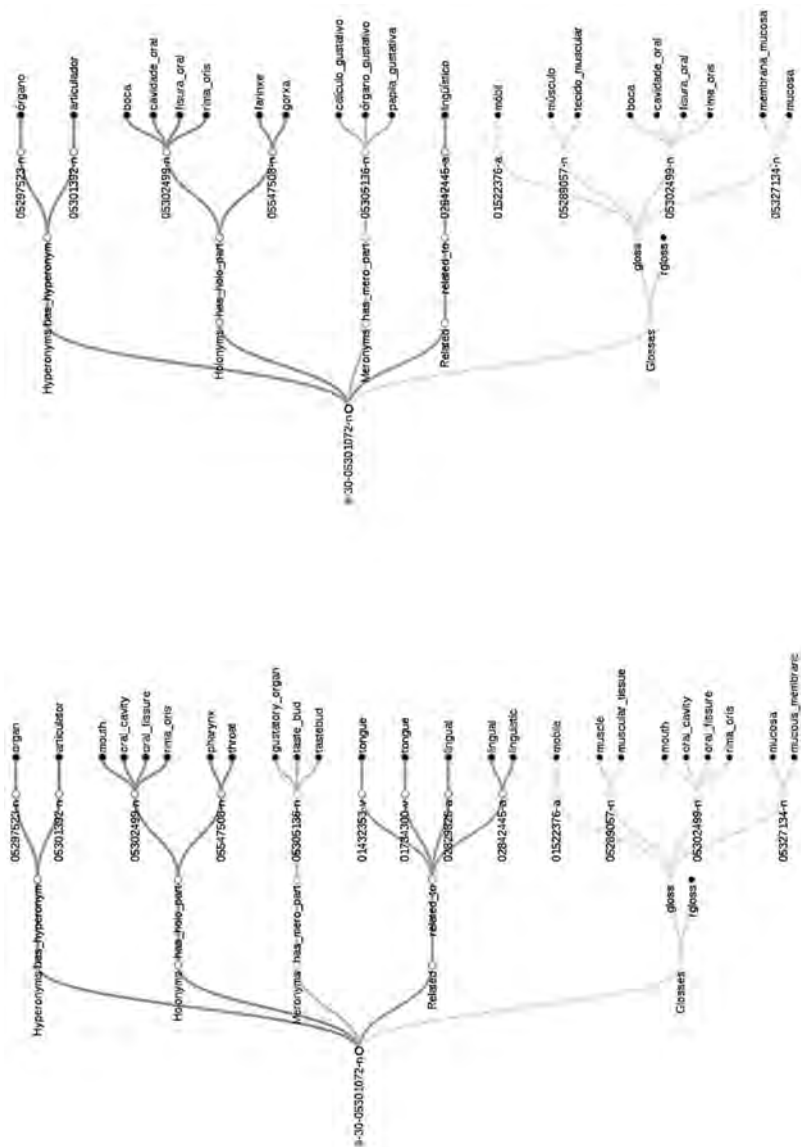
Figura 1. Visualización en forma de grafo con WordVis



También la interfaz de Galnet ofrece una herramienta de visualización de los resultados de las búsquedas léxicas, si bien en este caso la presentación visual elegida posee una marcada orientación bilingüe, centrada en las correspondencias entre inglés y gallego. Desde el resultado de cualquier consulta léxica es posible ver la lista de *synsets* que poseen una relación semántica directa con el *synset* buscado, tanto para el inglés como para el gallego. Galnet permite asimismo desplegar un grafo dirigido en forma de árbol interactivo, donde la raíz es el *synset* consultado, los nodos hijos son los *synsets* relacionados semánticamente con el nodo raíz, y las ramas o flechas que unen los nodos son las diversas relaciones semánticas entre los *synsets* hijos y el *synset* raíz, tal y como se muestra en la Figura 2 para el concepto glosado en el WordNet del inglés como «a mobile mass of muscular tissue

covered with mucous membrane and located in the oral cavity», representado por la variante inglesa *tongue* y la gallega *lingua*.

Figura 2. Visualización bilingüe en forma de grafo árbol con Galnet



4. Construcción de campos léxico-semánticos en Galnet

Los campos léxicos o semánticos son usados en el campo de la enseñanza de las lenguas con distintos fines, como por ejemplo la elaboración de material didáctico o la organización de la información léxica en obras lexicográficas destinadas al aprendizaje de un idioma. Los diccionarios temáticos, que pueden ser también visuales, agrupan las palabras y las expresiones de significado parecido en entradas basadas en campos semánticos. Tratándose de repertorios plurilingües, este tipo de organización lexicográfica posibilita que las personas usuarias del diccionario puedan elegir la traducción buscada entre los varios sinónimos de un mismo sentido, pudiendo considerar al mismo tiempo la elección de términos similares. El diccionario bilingüe inglés-español *Word Selector* de Cambridge (McCarthy 1995) constituye, sin duda, un magnífico ejemplo de diccionario temático plurilingüe y uno de los diccionarios escolares inglés-español más utilizados en la actualidad en las aulas de idiomas de nuestro entorno académico. Esta excelente obra de referencia no solo proporciona las traducciones al español de las distintas acepciones de las palabras en inglés, sino que para cada sentido de la palabra, cuando es necesario, ofrece su definición, ejemplos de uso y aclaraciones sobre su traducción. Las entradas del *Word Selector* están agrupadas en 450 categorías léxicas numeradas, por ejemplo, la categoría número 124 para *Illnesses / Enfermedad*, la 155 para *Vegetables / Hortalizas* o la 379 para *Music / Música*. Muchas de estas categorías se encuentran divididas a su vez en subcategorías. Así, la categoría 379 para *Music / Música* se encuentra subdividida en la 379.1 para *Tipos de música*, la 379.2 para *Aspectos de la música*, la 379.3 para *Agrupaciones musicales*, la 379.4 para *Instrumentos musicales*, la 379.5 para *Música vocal*, la 379.6 para *Acontecimientos musicales*, la 379.7 para *Obras de música*, la 379.8 para *Música escrita* y la 379.9 para *Música grabada*. Una subcategoría como la de *Instrumentos musicales* incluye información lexicográfica para 60 términos del inglés (*play, player, instrument, instrumental, musical instrument, stringed instruments, violin, violinist, viola, viola player, cello, cellist*, etc.), mientras que otra como *Tipos de música* puede contener solo 14 (*pop, pop music, pop star, folk music, folk, folk song, rock, reggae*, etc.). En total, el número de términos recogidos en las distintas categorías del *Word Selector* se sitúa en torno a los 7.000 para cada lengua.

Del mismo modo, pero con una extensión terminológica mucho mayor, la interfaz web de Galnet ofrece Termonet (Solla Portela / Gómez Guinovart 2015b; <http://sli.uvigo.es/galnet/termonet.php>), una poderosa herramienta para consultar los *synsets* de WordNet desde una perspectiva temática plurilingüe. La función de Termonet consiste en facilitar la extracción de variantes de WordNet relacionadas con un campo temático. A tal fin, Termonet ofrece un formulario de consulta que permite indicar el *synset* de origen que definirá el ámbito de extracción, y seleccionar luego el conjunto de relaciones semánticas que se utilizarán para la identificación de los términos del ámbito, así como la distancia o nivel de profundidad hasta donde se desea desplegar cada tipo de relación. De este modo, Termonet desplegará el árbol de relaciones desde el *synset* de origen y a través de cada relación hasta conseguir el nivel de profundidad determinado. Así mismo, el formulario ofrece la posibilidad de eliminar ciertas ramas del árbol de relaciones desplegado mediante la selección de filtros que impiden la exploración de ciertas relaciones en función de la relación precedente. A pesar de que Termonet permite un alto grado de personalización del conjunto de relaciones utilizados para la extracción del campo léxico, se pueden obtener resultados congruentes con los valores por omisión de los parámetros de la herramienta, siendo el procedimiento de extracción idéntico para ámbitos conceptuales amplios y para campos más concisos. En la Figura 3 se muestra un ejemplo de extracción del campo léxico del ámbito de la medicina a partir de la exploración llevada a cabo por Termonet del *synset* representado por la variante inglesa *medical science* con los parámetros predefinidos de la aplicación.

Figura 3. Construcción del campo léxico de la medicina en Termonet

- [0] [06045562-n](#) | **eng:** [medical_science](#) | **glg:** [ciencia_médica](#), [medicina](#) ({3} [science](#))
- [+1] 1 has_hyperonym [06037298-n](#) | **eng:** [bioscience](#), [life_science](#) ({2} [science](#))
- [+1] 2 has_hyponym [06043075-n](#) | **eng:** [medical_specialty](#), [medicine](#) | **glg:** [especialidade_médica](#), [medicina](#) ({0} [medical_specialty](#))
- [+2] 1 has_hyponym [06046745-n](#) | **eng:** [allergy](#) ({1} [medical_specialty](#))
- [+2] 2 has_hyponym [06046383-n](#) | **eng:** [anesthesiology](#) ({1} [medical_specialty](#))
- [+3] 1 related_to [09793495-n](#) | **eng:** [anaesthetist](#), [anesthesiologist](#), [anesthetist](#) | **glg:** [anestesista](#) ({1} [medical_specialist](#))
- [+4] 2 related_to [14023491-n](#) | **eng:** [anaesthesia](#), [anesthesia](#) | **glg:** [anestesia](#) ({1} [physical_condition](#))
- [+5] 1 related_to [02103052-n](#) | **eng:** [anaesthetic](#), [anesthetic](#) ({2} [drug](#) {2} [physical_condition](#))
- [+6] 1 related_to [02710786-n](#) | **eng:** [anaesthetic](#), [anaesthetic_agent](#), [anesthetic](#), [anesthetic_agent](#) | **glg:** [anestesia](#) ({1} [drug](#))
- [+5] 2 related_to [03041491-n](#) | **eng:** [anaesthetic](#), [anesthetic](#) ({2} [physical_condition](#))
- [+2] 3 has_hyponym [06046528-n](#) | **eng:** [angiology](#) ({1} [medical_specialty](#))
- [+3] 1 related_to [09793830-n](#) | **eng:** [angiologist](#) ({1} [doc](#))
- [+2] 4 has_hyponym [06046692-n](#) | **eng:** [bacteriology](#) ({1} [medical_specialty](#))
- [+3] 1 related_to [02914740-n](#) | **eng:** [bacteriologic](#), [bacteriological](#) ({2} [medical_specialty](#))
- [+3] 2 related_to [09831411-n](#) | **eng:** [bacteriologist](#) ({1} [biologist](#))
- [+3] 3 category_term [14899378-n](#) | **eng:** [culture_medium](#), [medium](#) | **glg:** [medio](#), [medio_do_cultivo](#) ({1} [food](#))
- [+4] 1 has_hyponym [14900184-n](#) | **eng:** [agar](#), [nutrient_agar](#) | **glg:** [ágar-ágar](#), [ágar_nutritivo](#), [placa_de_ágar-ágar](#) ({2} [food](#))
- [+5] 1 has_hyponym [14900342-n](#) | **eng:** [blood_agar](#) | **glg:** [sangue_ágar-ágar](#) ({3} [food](#))
- [+5] 2 has_mero_madeof [14900008-n](#) | **eng:** [agar](#), [agar-agar](#) | **glg:** [ágar-ágar](#) ({1} [num](#))
- [+4] 2 has_hyponym [80000645-n](#) | **eng:** [nutrient_broth](#) ({2} [doc](#))
- [+2] 5 has_hyponym [06046898-n](#) | **eng:** [biomedicine](#) ({1} [medical_specialty](#))
- [+3] 1 has_hyponym [06046037-n](#) | **eng:** [aeromedicine](#), [aerospace_medicine](#), [aviation_medicine](#) ({2} [medical_specialty](#))
- [+4] 1 related_to [02907187-a](#) | **eng:** [aeromedical](#) ({3} [medical_specialty](#))
- [+2] 6 has_hyponym [06047096-n](#) | **eng:** [biomedicine](#) | **glg:** [biomedicina](#) ({1} [medical_specialty](#))
- [+3] 1 related_to [02760316-a](#) | **eng:** [biomedical](#) | **glg:** [biomédico](#) ({2} [medical_specialty](#))
- [+2] 7 has_hyponym [06047275-n](#) | **eng:** [cardiology](#) ({1} [medical_specialty](#))
- [+3] 1 related_to [02914902-a](#) | **eng:** [cardiologic](#) ({2} [medical_specialty](#))
- [+3] 2 related_to [09894445-n](#) | **eng:** [cardiologist](#), [heart_specialist](#), [heart_surgeon](#) | **glg:** [cardiólogo](#) ({1} [medical_specialist](#))
- [+2] 8 has_hyponym [06047430-n](#) | **eng:** [dental_medicine](#), [dentistry](#), [odontology](#) | **glg:** [odontología](#) ({1} [medical_specialty](#))
- [+3] 1 has_hyponym [06047923-n](#) | **eng:** [cosmetic_dentistry](#) ({2} [medical_specialty](#))
- [+3] 2 has_hyponym [06048057-n](#) | **eng:** [dental_surgery](#) ({2} [medical_specialty](#))
- [+4] 1 has_hyponym [06048373-n](#) | **eng:** [exodontia](#), [exodontics](#) ({3} [medical_specialty](#))
- [+5] 1 related_to [02916403-a](#) | **eng:** [exodontic](#) ({4} [medical_specialty](#))
- [+5] 2 related_to [10071917-n](#) | **eng:** [exodontist](#) ({5} [adult](#))
- [+3] 3 has_hyponym [06048184-n](#) | **eng:** [endodontia](#), [endodontics](#) ({2} [medical_specialty](#))
- [+4] 1 related_to [02915507-a](#) | **eng:** [endodontic](#) ({3} [medical_specialty](#))
- [+4] 2 related_to [10057114-n](#) | **eng:** [endodontist](#) ({5} [adult](#))
- [+3] 4 has_hyponym [06048552-n](#) | **eng:** [dental_orthopaedics](#), [dental_orthopedics](#), [orthodontia](#), [orthodontics](#), [orthodonture](#) | **glg:** [ortodoncia](#), [ortodontología](#) ({2} [medical_specialty](#))
- [+4] 1 related_to [02915815-a](#) | **eng:** [orthodontic](#) ({3} [medical_specialty](#))
- [+4] 2 related_to [10384610-n](#) | **eng:** [orthodontist](#) ({5} [adult](#))
- [+3] 5 has_hyponym [06048851-n](#) | **eng:** [periodontia](#), [periodontics](#) ({2} [medical_specialty](#))
- [+4] 1 related_to [02916006-a](#) | **eng:** [periodontal](#), [periodontic](#) ({3} [medical_specialty](#))
- [+4] 2 related_to [10416567-n](#) | **eng:** [periodontist](#) ({5} [adult](#))

En comparación con el *Word Selector*, las posibilidades de WordNet en cuanto a la consulta de campos léxico-semánticos son mucho más extensas, ya que parte de un número de categorías mucho mayor (unas 2.000 en el *Word Selector*, teniendo en cuenta las subcategorías incluidas, frente a los más de 100.000 *synsets* de WordNet que potencialmente pueden encabezar un campo léxico en Termonet) y de un número de términos también mucho mayor (unos 7.000 por lengua en el *Word Selector*, frente a las 206.941 variantes sinonímicas del WordNet del inglés, las 116.773 del español o las 50.445 del gallego). Para ilustrar esta diferente amplitud con un ejemplo concreto, obsérvese que la extracción con Termonet del campo léxico de *Instrumentos musicales*, correspondiente al *synset* de WordNet representado por la variante sinonímica *musical instrument*, genera 1.602 *synsets* (con sus relaciones semánticas explícitas) que incluyen 2.942 variantes sinonímicas del inglés, 1.674 del español y 709 del gallego (frente a los 60 vocablos del inglés recogidos en esta categoría por el *Word Selector*). Las diferencias entre lenguas que se aprecian en la extracción se deben al diverso grado de desarrollo de los léxicos WordNet de estas tres lenguas, siendo la inglesa la única lengua que posee variantes sinonímicas para todos los *synsets*. En cualquier caso, Termonet es capaz de construir un campo léxico bastante más poblado que el recogido en *Word Selector*, tanto en relación al número de palabras o expresiones, como a la variedad de lenguas representadas, ofreciendo así un abanico de posibilidades mucho más amplias para la enseñanza y el aprendizaje del idioma.

5. Conclusión

WordNet constituye en la actualidad el recurso léxico plurilingüe más extenso en número de conceptos, de palabras y de equivalencias interlingüísticas que puede ser usado en el aula de idiomas. Esto es así especialmente para aquellas lenguas que poseen un léxico WordNet en estado de desarrollo medio o avanzado, como el inglés (117.659 *synsets*), el finés (116.763), el tailandés (73.350), el francés (59.091), el japonés (57.184), el rumano (56.026), el catalán (46.558), el esloveno (42.583), el mandarín (42.312), el indonesio (38.085), el malayo (36.911), el italiano (35.001), el polaco (33.826), el gallego (33.120), el vasco (30.616), el neerlandés (30.177) o el croata (23.120). La detallada codificación de las relaciones semánticas entre los conceptos almacenados en WordNet permite la explotación del recurso en todo tipo de actividades didácticas orientadas temáticamente. Además, la posibilidad de complementar los *synsets* de WordNet con imágenes lo convierten en un diccionario temático ilustrado plurilingüe de gran utilidad en la enseñanza del idioma.

En su forma actual, WordNet resulta un recurso computacional sumamente extenso, complejo y versátil. Concebido originalmente como un modelo del léxico mental para la experimentación psicolingüística, WordNet es sin duda, en estos momentos, el diccionario electrónico usado mayoritariamente en los ámbitos de la lingüística computacional, del procesamiento del lenguaje natural y de las tecnologías lingüísticas. Así y todo, WordNet también nos permite explotar directamente su potencial de conocimiento en aplicaciones lexicográficas y didácticas, demostrando un alto nivel de eficacia y utilidad pedagógica en el ámbito de la enseñanza de las lenguas extranjeras.

Referencias bibliográficas

- Agirre, Eneko / Edmonds, Philip (2006): *Word Sense Disambiguation*. Berlín: Springer.
- Agirre, Eneko / Alegria, Iñaki / Rigau, German / Vossen, Piek (2007): «MCR for CLIR», *Procesamiento del Lenguaje Natural*, 38, pág. 3–15.
- Álvez, Javier / Atserias, Jordi / Carrera, Jordi / Climent, Salvador / Oliver, Antoni / Rigau, German (2008): «Consistent Annotation of EuroWordNet with the Top Concept Ontology», en Attila Tanács / Dóra Csendes / Veronika Vincze / Christiane Fellbaum / Piek Vossen, eds., *Proceedings of the 4th Global WordNet Conference*. Szeged: Global WordNet Association, s.p.
- Bentivogli, Luisa / Forner, Pamela / Magnini, Bernardo / Pianta, Emanuele (2004): «Revising WordNet Domains Hierarchy: Semantics, Coverage, and Balancing», en Gilles Sérasset / Susan Armstrong / Christian Boitet / Andrei Popescu-Belis / Dan Tufis, eds., *Proceedings of COLING Workshop on Multilingual Linguistic Resources*. Ginebra: ACL, pág. 101–108.
- Bhingardive, Sudha / Ajotikar, Tanuja / Kulkarni, Irawati / Kulkarni, Malhar / Bhattacharyya, Pushpak (2014): «Semi-Automatic Extension of Sanskrit Wordnet using Bilingual Dictionary», en Heili Orav / Christiane Fellbaum / Piek Vossen, eds., *Proceedings of the Seventh Global WordNet Conference*. Tartu: Global WordNet Association, pág. 324–329.
- Bond, Francis / Foster, Ryan (2013): «Linking and Extending an Open Multilingual WordNet», en Roberto Navigli / Jing-Shin Chang / Stefano Faralli, eds., *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*. Sofía: ACL, pág. 1352–1362.
- Cai, Qingqing / Gung, James / Guan, Maochen / Kurlandski, Gerald / Pease, Adam (2016): «Word Substitution in Short Answer Extraction: A WordNet-based Approach», en Verginica Barbu Mititelu / Corina Forăscu / Christiane Fellbaum / Piek Vossen, eds., *Proceedings of the Eighth Global WordNet Conference*. Bucarest: University of Iași, pág. 66–73.

- Corbeil, Jean-Claude / Archambault, Ariane (2004): *The Firefly Five Languages Visual Dictionary: English-Spanish-French-German-Italian*. Buffalo / Richmond Hill: Firefly Books.
- Deng, Jia / Dong, Wei / Socher, Richard / Li, Li-Jia / Li, Kai / Fei-Fei, Li (2009): «ImageNet: A Large-Scale Hierarchical Image Database», en *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition CVPR 2009*. Miami: IEEE, pág. 248–255.
- Elberrichi, Zakaria / Rahmoun, Abdelattif / Bentaalah, Mohamed Amine (2008): «Using WordNet for Text Categorization», *The International Arab Journal of Information Technology*, 5/1, pág. 16–24.
- Fang, Hui (2008): «A Re-examination of Query Expansion Using Lexical Resources», en *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics*. Columbus: Association for Computational Linguistics, pág. 139–147.
- Fellbaum, Christiane, ed. (1998): *WordNet: An Electronic Lexical Database*. Cambridge: MIT Press.
- Fernández Montraveta, Ana / Vázquez, Gloria (2010): «La construcción del WordNet 3.0 en español», en María Auxiliadora Castillo / Juan Manuel García Platero, eds., *La lexicografía en su dimensión teórica*. Málaga: Universidad de Málaga, pág. 201–220.
- Ferrández, Sergio / Ferrández, Antonio / Roger, Sandra / López-Moreno, Pilar (2007): «Búsqueda de respuestas bilingüe basada en ILLI, el sistema BRILL», *Procesamiento del Lenguaje Natural*, 38, pág. 27–33.
- González Agirre, Aitor / Rigau, German (2013): «Construcción de una base de conocimiento léxico multilingüe de amplia cobertura: Multilingual Central Repository», *Linguamática*, 5/1, pág. 13–28.
- Hu, Xiangen / Graesser, Arthur C. (1998): «Using WordNet and latent semantic analysis to evaluate the conversational contributions of learners in the tutorial dialog», en *Proceedings of the International Conference on Computers in Education*, vol. 2. Beijing: Springer, pág. 337–341.
- Huang, Bill (2016): «WNSpell: a WordNet-Based Spell Corrector», en Verginica Barbu Mititelu / Corina Forăscu / Christiane Fellbaum / Piek Vossen, eds., *Proceedings of the Eighth Global WordNet Conference*. Bucarest: University of Iași, pág. 135–142.
- Isahara, Hitoshi / Bond, Francis / Uchimoto, Kiyotaka / Utiyama, Masao / Kanzaki, Kyoko (2008): «Development of the Japanese WordNet», en Nicoletta Calzolari / Khalid Choukri / Bente Maegaard / Joseph Mariani / Jan Odjik / Stelios Piperidis / Daniel Tapias, eds., *Proceedings of the Sixth International Language Resources and Evaluation*. Marrakech: ELRA, s.p.

- Knoop, Susanne / Wilske, Sabrina (2013): «WordGap – Automatic generation of gap-filling vocabulary exercises for mobile learning», en Elena Volodina / Lars Borin / Hrafn Loftsson, eds., *Proceedings of the Second Workshop on NLP for Computer-Assisted Language Learning*. Linköping: Linköpings Universitet, pág. 39–47.
- McCarthy, Michael (1995): *Cambridge Word Selector Inglés-Español. Diccionario temático del inglés contemporáneo*. Cambridge: Cambridge University Press.
- Oliver, Antoni / Climent, Salvador (2011): «Construcción de los WordNets 3.0 para castellano y catalán mediante traducción automática de corpus anotados semánticamente», *Procesamiento del Lenguaje Natural*, 47, pág. 293–300.
- Ordan, Noam / Shuly Wintner (2007): «Hebrew WordNet: a Test Case of Aligning Lexical Databases Across Languages», *International Journal of Translation*, 19/1, pág. 39–58.
- Pease, Adam / Niles, Ian / Li, John (2002): «The Suggested Upper Merged Ontology: A Large Ontology for the Semantic Web and its Applications», en *Working Notes of the AAAI-2002 Workshop on Ontologies and the Semantic Web*. Edmonton: AAAI, s.p.
- Plaza, Laura / Díaz, Alberto / Gervás, Pablo (2010): «Automatic summarization of news using WordNet concept graphs», *IADIS International Journal on Computer Science and Information Systems*, 5/1, pág. 45–57.
- Pociello, Elisabete / Agirre, Eneko / Aldezabal, Izaskun (2011): «Methodology and Construction of the Basque WordNet», *Language Resources and Evaluation*, 45/2, pág. 121–142.
- Simões, Alberto / Gómez Guinovart, Xavier (2014): «Bootstrapping a Portuguese WordNet from Galician, Spanish and English wordnets», en Juan Luis Navarro Mesa et al., eds., *Advances in Speech and Language Technologies for Iberian Languages*. Berlín: Springer, pág. 239–248.
- Solla Portela, Miguel Anxo / Gómez Guinovart, Xavier (2015a): «Galnet: o WordNet do galego. Aplicacións lexicolóxicas e terminolóxicas», *Revista Galega de Filoloxía*, 16, pág. 169–201.
- Solla Portela, Miguel Anxo / Gómez Guinovart, Xavier (2015b): «Termonet: Construcción de terminologías a partir de WordNet y corpus especializados», *Procesamiento del Lenguaje Natural*, 55, pág. 165–168.
- Solla Portela, Miguel Anxo / Gómez Guinovart, Xavier (2016): «DBpedia del gallego: recursos y aplicaciones en procesamiento del lenguaje», *Procesamiento del Lenguaje Natural*, 57, pág. 139–142.
- Sun, Koun-Tem / Huang, Yueh-Min / Liu, Ming-Chi (2011): «A WordNet-Based Near-Synonyms and Similar-Looking Word Learning System», *Educational Technology & Society*, 14/1, pág. 121–134.

- Vintar, Špela / Fišer, Darja / Vrščaj, Aljoša (2012): «Were the clocks striking or surprising?: using WSD to improve MT performance», en *Proceedings of the Joint Workshop on Exploiting Synergies between Information Retrieval and Machine Translation (ESIRMT) and Hybrid Approaches to Machine Translation (HyTra) (EACL 2012)*. Stroudsburg: ACL, pág. 87–92.
- Vossen, Piek (2002): «WordNet, EuroWordNet and Global WordNet», *Revue française de linguistique appliquée*, 7, pág. 27–38.
- Zhao, Feng / Fang, Fei / Yan, Fengwei / Jin, Hai / Zhang, Qin (2012): «Expanding approach to information retrieval using semantic similarity analysis based on WordNet and Wikipedia», *International Journal of Software Engineering and Knowledge Engineering*, 22/2, pág. 305–322.

Índice de tablas

Tabla 1. Número de *synsets* y variantes sinonímicas por idioma p. 270

Índice de gráficos

Figura 1. Visualización en forma de grafo con WordVis p. 272

Figura 2. Visualización bilingüe en forma de grafo arboreo con Galnet p. 273

Figura 3. Construcción del campo léxico de la medicina en Termonet ... p. 276