

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE FILOLOGÍA



TESIS DOCTORAL

Combinatoria léxica, polisemia y polisemia regular en una base de conocimiento léxico conceptual : el caso de *Redes diccionario combinatorio del español contemporáneo* y functional grammar knowledge base

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Maia Sherwood Droz

Director

Ricardo Mairal Usón

Madrid, 2014



FUNDACIÓN
INSTITUTO UNIVERSITARIO DE INVESTIGACIÓN
JOSÉ ORTEGA Y GASSET

**COMBINATORIA LÉXICA, POLISEMIA Y POLISEMIA REGULAR
EN UNA BASE DE CONOCIMIENTO LÉXICO CONCEPTUAL:
EL CASO DE *REDES DICCIONARIO COMBINATORIO DEL ESPAÑOL
CONTEMPORÁNEO* Y FUNCTIONAL GRAMMAR KNOWLEDGE BASE**

Doctoranda

Maia Sherwood Droz

Director de la Tesis

Dr. Ricardo Mairal Usón

Universidad Complutense de Madrid

Instituto Universitario de Investigación Ortega y Gasset

Doctorado en Lingüística Teórica y sus Aplicaciones

Madrid 2013

ABSTRACT

This dissertation attempts to link the linguistic information contained in *REDES Diccionario combinatorio del español contemporáneo* (Bosque, 2004), or REDES, to the ontological framework of *Functional Grammar Knowledge Base*, or FunGramKB. REDES is a dictionary that gathers systematic restrictions imposed by some 4,000 Spanish predicates to their selection of lexical arguments, while FunGramKB is a multilingual and multipurpose lexical conceptual knowledge base (KB) designed for Natural Language Processing (NLP). This work is thus part of the field known as electronic lexicography for the XXIst century or the third millennium (Fuertes and Tarp, 2011).

This new lexicography comprises electronic lexicographical resources which are much more complex than the commonly used digitalized versions of traditional dictionaries. It is composed of lexical databases or lexical knowledge bases, of variable complexity and depth, built on electronic platforms. Even though electronic lexicography can also serve the typical dictionary queries posed by humans, the best-known NLP applications include machine translation (MT), question and answer systems, information extraction, and voice recognition programs. However, the main problem facing NLP continues to be Word Sense Disambiguation (WSD). Most words in a language are polysemic and a successful NLP application will depend on its ability to assign the correct sense to each word in a given context. A great deal of the work that is taking place in NLP is therefore focused on finding strategies to effectively disambiguate words in their context.

In working with REDES and FunGramKB, it could seem at first glance that we are dealing with two distant fields –print lexicography and knowledge engineering, respectively–, but this thesis assumes that linking these resources would yield significant benefits for both. Furthermore, it would combine two sources of valuable linguistic and theoretical data: on the one hand, REDES contributes thousands of patterns of systematic predicate-argument word combinations, taken from real use in Spanish-language corpora. These combinations are not presented as “collocations”, that is, combinations between a single predicate and a single argument, or vice versa, but rather, between predicates and “lexical classes”, groups of arguments that share a common semantic basis. On the other hand, FunGramKB contributes a multilevel electronic platform designed for NLP, which includes a conceptual, a lexical and a grammatical model, and has been built around a hierarchical and taxonomical ontology of universal cognitive concepts.

The research is therefore motivated by two general questions: How can FunGramKB serve REDES? and How can REDES serve FunGramKB?, as well as various sub questions. Our general hypotheses are that a lexical conceptual KB such as FunGramKB could help structure, standardize and interrelate the content in REDES in ways that are impossible in the current print version, which

would allow for a greater use of its data. It could also set the groundwork for an electronic version of REDES –not just a digitalized version, but a potentially “intelligent” dictionary–, which would show more clearly the systematic relationship between certain predicates and arguments, or rather, between certain predicate senses and argument senses. At the same time, the content of REDES – which has not been collected in any other work– could help FunGramKB populate both its conceptual and lexical models with exhaustive and semantically-classified information on Spanish word combinations. This data could also potentially aide FunGramKB in the management of WSD, through associating particular word combinations to particular word senses, as well as in natural language production.

In this study, we were also interested in regular or systematic polysemy, that is, regular sense alternation, the phenomenon by which words like *school*, *college* or *university* can alternate regularly between meanings such as ‘building’ and ‘institution’. Regular polysemy is valuable to NLP because once a pattern is identified, it can be extended to other lexical units that may have one or the other meaning. For instance, a NLP program may assign the ‘institution’ meaning to any word that already has the ‘building’ meaning, substantially reducing the processing effort. Most of the work done on regular polysemy has focused on nouns, and the work on verbs has focused on aspectual variation, such as the causative/inchoative alternation. We look at regular polysemy in a group of verbs in semantic terms, to see how the alternating meanings, and their lexical selection, can be represented in FunGramKB.

To answer our questions, the thesis explores the representation in FunGramKB of 15 polysemic verbs, in terms of their senses, their systematic word combinations –as provided by REDES–, and the relation between the two. First, the verb *dictar*, which selects a total of eight “lexical classes”, is analyzed independently. Then nine verbs that share at least one general meaning (‘waste’), and which altogether select a total of 46 “lexical classes”, are analyzed independently and comparatively: *malgastar*, *desaprovechar*, *dilapidar*, *derrochar*, *malograr*, *disipar*, *perder*, *despilfarrar*, *desperdiciar*. Finally, five verbs that show evidence of regular polysemy –and which altogether select 32 “lexical classes”– are analyzed independently and comparatively: *tejer*, *tramar*, *urdir*, *cocinar* and *maquinar*.

During our research, we met various challenges. Some have to do with the nature of the information contained in each resource, and the way that information is represented. In REDES, linguistic information is gathered from corpora and inductively organized in “lexical classes”, each of which is described semantically using natural language. Every “lexical class” is therefore unique. FunGramKB organizes its architecture around a taxonomical ontology of universal cognitive (non linguistic) concepts, which are defined using a formal language called Conceptual Representation Language (COREL). The matching of detailed semantic descriptions of “lexical

classes” in REDES with a limited number of formally defined concepts in FunGramKB required reducing and standardizing the linguistic particularities of the data provided by REDES. The final challenge had to do with figurative language: “lexical classes” included nouns that expressed a sense due to their inherent semantics, but also others that acquired that sense in combination with a given verb, through metaphoric, metonymic or pragmatic processes. The former were linked to concepts in FunGramKB, whereas some of the latter were placed in the predicate’s lexical template, as “collocations”.

The work concludes the linguistic information presented in REDES can be adequately represented in FunGramKB’s ontological framework: verb meanings were linked to conceptual events and “lexical class” meanings were linked to conceptual entities. Lexical selection restrictions captured in REDES were expressed as conceptual selection preferences in FunGramKB, as part of an existing event’s properties or through the creation of new terminal concepts or subconcepts. Subconcepts are a useful resource provided by FunGramKB that allow for the inclusion of selection preferences in existing concepts, without having to add new concepts. These strategies allowed us to standardize and reduce REDES’ linguistic data, and made patterns more visible. For example, the nine verbs that shared the sense of ‘waste’ were linked satisfactorily to the event +WASTE_00. The 46 “lexical classes” which they selected were reduced to 12 conceptual entities in FunGramKB. Our results suggest that the regular selection preferences for +WASTE_00 include the following entities: +ABILITY_00, +TIME_00, +MONEY_00, +STRENGTH_00 and +BULLET_00. At the same time, the particular meanings and lexical selection restrictions of each verb were expressed by terminal concepts or subconcepts. For instance, *derrochar* merited reinforcing the meaning of +WASTE_00 with the quantifier *m* (*much*) through the creation of terminal concept \$\$SQUANDER_00. Also, even though *dilapidar*, *disipar*, *despilfarrar* and *desperdiciar* were linked to +WASTE_00, their systematic lexical selections justify the creation of subconcepts -WASTE_MONEY (*dilapidar*, *disipar dinero*), and -WASTE_FOOD (*despilfarrar*, *desperdiciar comida*).

We used FunGramKB’s lexical model to accommodate as “collocations” the lexical selections that could not be codified conceptually, in particular, not established metaphoric or metonymic uses, pragmatic uses, and idiomatic uses. For example, *desperdiciar* (‘waste’) selects a lexical class that we linked to the concept +PERIOD_00. Some of the lexical units were linked to the concept (*tiempo*, *día*, *semana*, *juventud*), but some had to be located as collocations of *desperdiciar* in its lexical template (*porvenir*, *vacaciones*, *vida*). These nouns acquire the value of ‘time period’ in combination with the verb *desperdiciar* (*desperdiciar las vacaciones* would be *to waste your vacation*). Therefore, FunGramKB can house all of the information provided by REDES, but it requires that the members of the “lexical classes” be divided and redistributed in

different places in the KB, according to whether its meaning in the combination responds to its inherent meaning or to an adjustment process. Given that the meanings of each verb, including their lexical selections (as conceptual selection preferences or “collocations”), were satisfactorily represented by FunGramKB, we believe that this data can aide WSD processes; nevertheless, that test falls outside the scope of the present study.

FunGramKB was also able to represent a direct pattern of regular polysemy, in the case of verbs *tejer*, *tramar* and *urdir*, whose senses alternate between +WEAVE_00 and \$CONSPIRE_00. We were also able to codify an indirect pattern of regular polysemy through higher level concepts: +CREATE_00 is the superordinate concept to +WEAVE_00 as well as to +COOK_00, so we can say *cocinar(se)*, *tejer*, *tramar*, *urdir* alternate between +CREATE_00 and \$CONSPIRE_00.

Another one of our conclusions is that the data in REDES can help populate and enrich both the conceptual and lexical models in FunGramKB. Lexical selection restrictions captured in REDES can be expressed as conceptual selection preferences that may not yet be included in FunGramKB’s ontology, and which require the addition of new terminal concepts or subconcepts. Our work proposes the addition of four new terminal concepts (\$DICTATE_00, \$MISS_00, \$DISSIPATE_00, and \$SQUANDER_00), as well as 37 new subconcepts. The data in REDES can also help to populate FunGramKB’s lexical model: the verbs become lexicalizations for conceptual events, and the noun members of the “lexical classes” become lexicalizations for conceptual entities. The members of the “lexical classes” that cannot be considered a lexicalization of a conceptual entity, populate the verb’s lexical template as “collocations” of the verb.

The thesis is organized as follows: Chapter I describes the data contained in REDES, emphasizing its importance for any model that aims to represent linguistic knowledge, specifically systematic selection restrictions. It argues that REDES gives priority to meaning, rather than to words, an outlook that is shared by XXIst century electronic lexicography, which makes possible its transference to an electronic platform. We also detail the challenges entailed by this process. In Chapter II, we review the most salient accomplishments of current electronic lexicography, as well as its pending goals, focusing on the WordNet, EuroWordNet and FrameNet lexical data bases. In Chapter III, we present FunGramKB, our theoretical and practical framework, giving particular attention to the ontology, the concepts and their properties, and the management of selection preferences and collocations. In Chapter IV, we discuss the problem of polysemy and WSD in traditional and electronic lexicography. We focus on the main WSD strategies used in NLP, specifically on the role of selection preferences in WSD. In Chapter V, we discuss regular polysemy from a theoretical perspective and from the standpoint of electronic lexicography. We comment on the main studies conducted on regular polysemy in electronic lexical data bases. Chapter VI explains the methodology followed on this thesis. Chapter VII presents and analyzes

our results for polysemic verbs *dictar* and *derrochar*, *desaprovechar*, *desperdiciar*, *despilfarrar*, *dilapidar*, *disipar*, *malgastar*, *malograr*, *perder*. Chapter VIII presents and analyzes results for regular polysemy verbs *tejer*, *tramar*, *urdir*, *cocinar(se)*, *maquinar*. Chapter IX presents our conclusions, and establishes future areas of research.

AGRADECIMIENTOS

A mis padres, Edna Droz y Tim Sherwood, sin cuyo apoyo incondicional no hubiera podido terminar esta tesis, y a mi hermano, Nicolás, por servir siempre de inspiración.

A la Academia Puertorriqueña de la Lengua Española, por su apoyo concreto a esta tesis, y a cada uno de sus miembros, por su respaldo y cariño través de los años. También a la memoria de la Dra. María Vaquero de Ramírez, cuyo ejemplo me sigue inspirando al día de hoy.

A Ricardo Mairal, por dirigir este trabajo con ánimo y confianza constantes, a Carlos Periñán, por su asesoría técnica, y a Ignacio Bosque, por la inspiración de su trabajo y su disposición continua de formular y contestar preguntas. También a Patrick Hanks, por su generosidad intelectual.

A la Universidad de Puerto Rico, por el uso de sus recursos bibliotecarios, y en especial a Manuel Martínez, por su diligencia, eficiencia y buen humor en la gestión de préstamos interbibliotecarios.

A mis amigas del alma, Alexandra e Inés, y a todas mis otras amigas cercanas y queridas del mundo de la lingüística.

A Rogelio, por su compañía a través de este camino.

A todos mis otros amigos que de tantas maneras me dieron la mano desde el comienzo del programa de doctorado en Madrid hasta el final de la redacción en San Juan.

¡A TODOS, GRACIAS!

A mis padres, Edna y Tim.

ÍNDICE

| | |
|--|--------------|
| Abstract | i |
| Agradecimientos | vi |
| Índice de Tablas | xi |
| Índice de Imágenes | xx |
| Siglarío | xxi |
| Abreviaturas | xxiii |
| Introducción | xxiv |
| Capítulo I. REDES. Diccionario combinatorio del español contemporáneo | 1 |
| Introducción | 1 |
| 1. Presentación general de REDES | 2 |
| 2. Los predicados | 4 |
| 2.1 Clases de predicados..... | 7 |
| 3. Las clases léxicas | 8 |
| 4. Los argumentos | 11 |
| 5. Los índices conceptuales | 13 |
| 6. Las restricciones de selección léxica | 17 |
| 6.1 Modelos teóricos..... | 17 |
| 6.2 Diccionarios de colocaciones..... | 19 |
| Conclusión | 20 |
| Capítulo II. La lexicografía informatizada del siglo XXI | 22 |
| Introducción | 22 |
| 1. Relación entre lexicografía e informática | 24 |
| 2. Diccionarios electrónicos | 27 |
| 2.1 Inglés..... | 27 |
| 2.2 Español..... | 32 |
| 3. Lexicones electrónicos | 33 |
| 3.1 WordNet..... | 35 |
| 3.2 EuroWordNet..... | 38 |
| 3.3 FrameNet..... | 40 |
| 4. Ontologías | 46 |
| Conclusión | 49 |
| Capítulo III. Functional Grammar Knowledge Base (FunGramKB) | 51 |
| Introducción | 51 |
| 1. Trasfondo teórico | 53 |
| 2. Niveles de FunGramKB | 56 |
| 2.1 Nivel léxico..... | 58 |
| 2.2 Nivel gramatical..... | 60 |
| 2.3 Nivel conceptual..... | 61 |
| 3. La ontología | 62 |
| 3.1 Los metaconceptos..... | 64 |
| 3.2 Los conceptos básicos..... | 66 |
| 3.3 Los conceptos terminales..... | 68 |
| 3.4 Los subconceptos..... | 69 |
| 4. Propiedades de los conceptos | 70 |
| 4.1 El sistema de notación COREL..... | 72 |
| 4.2 Los marcos temáticos (MT)..... | 74 |
| 4.3 Los postulados de significado (PS)..... | 75 |

| | |
|---|------------|
| 5. Las preferencias de selección en FunGramKB | 78 |
| 5.1 Preferencias de selección en los conceptos básicos..... | 79 |
| 5.2 Preferencias de selección en los conceptos terminales..... | 80 |
| 5.3 Preferencias de selección en los subconceptos | 80 |
| 6. Las colocaciones en FunGramKB | 82 |
| Conclusión..... | 84 |
| CAPÍTULO IV. La polisemia en la lexicografía informatizada | 85 |
| Introducción..... | 85 |
| 1. Polisemia, significado y sentidos | 86 |
| 2. La polisemia en el PLN: panorama histórico..... | 88 |
| 3. Legado de los diccionarios tradicionales..... | 90 |
| 4. Granularidad en los lexicones electrónicos | 93 |
| 5. La WSD: metodologías básicas..... | 97 |
| 6. El rol del contexto | 100 |
| 7. Las preferencias de selección | 102 |
| Conclusión..... | 105 |
| CAPÍTULO V. La polisemia regular | 106 |
| Introducción..... | 106 |
| 1. Acercamientos generales al estudio de la polisemia regular..... | 109 |
| 2. Relación entre los sentidos..... | 111 |
| 3. Clases de palabras..... | 114 |
| 4. Activación de sentidos..... | 115 |
| 5. La polisemia regular en la lexicografía electrónica..... | 118 |
| 5.1 Buitelaar (1998a, 1998b)..... | 119 |
| 5.2 Tomuro (1998) | 121 |
| 5.3 Peters y Peters (2000)..... | 122 |
| 5.4 Peters y Wilks (2001) | 124 |
| 5.5 Wing (2002) | 125 |
| 5.6 Peters (2004)..... | 127 |
| 5.7 Peters (2006)..... | 128 |
| 5.8 Barque y Chaumartin (2006)..... | 129 |
| 5.9 Barque (2007) | 131 |
| Conclusión..... | 132 |
| CAPÍTULO VI. Metodología..... | 134 |
| Introducción..... | 134 |
| 1. Significado y combinatoria | 135 |
| 1.1 Elegir predicados de REDES relacionados semánticamente..... | 136 |
| 1.2 Vincular cada clase léxica con un concepto de la ontología | 136 |
| 1.3 Clasificar cada clase léxica según la tipología establecida | 145 |
| 1.4 Analizar colectivamente las clases léxicas similares | 152 |
| 1.5 Vincular cada predicado con un concepto de la ontología..... | 153 |
| 1.6 Ubicar como colocaciones los argumentos que no puedan ser descritos conceptualmente .. | 159 |
| 2. Polisemia regular y combinatoria | 160 |
| 2.1 Elegir predicados relacionados semánticamente que sugieran un patrón de polisemia regular | 161 |
| 2.2 Vincular cada clase léxica con un concepto de la ontología | 161 |
| 2.3 Vincular cada predicado con un concepto de la ontología..... | 162 |
| 2.4 Ubicar como colocaciones los argumentos que no pueden ser descritos conceptualmente .. | 163 |
| Conclusión..... | 163 |
| Capítulo VII. Resultados y análisis: Polisemia y combinatoria léxica | 164 |
| Introducción..... | 164 |
| 1. DICTAR..... | 165 |
| 1.1 Resumen <i>DICTAR</i> | 174 |

| | |
|--|------------|
| 2. PERDER, DISIPAR, DERROCHAR, MALOGRAR, MALGASTAR, DILAPIDAR, DESAPROVECHAR, DESPILFARRAR, DESPERDICAR | 176 |
| 2.1 PERDER..... | 176 |
| 2.2 DISIPAR..... | 195 |
| 2.3 DERROCHAR | 210 |
| 2.4 MALOGRAR..... | 225 |
| 2.5 MALGASTAR..... | 232 |
| 2.6 DILAPIDAR..... | 241 |
| 2.7 DESAPROVECHAR..... | 248 |
| 2.8 DESPILFARRAR..... | 254 |
| 2.9 DESPERDICAR..... | 259 |
| 3. Análisis | 266 |
| 3.1 Preferencias de selección de +WASTE_00..... | 267 |
| Capítulo VIII. Resultados y análisis: Polisemia regular y combinatoria léxica | 278 |
| Introducción | 278 |
| 1. TEJER, TRAMAR, URDIR, COCINAR(SE), MAQUINAR | 278 |
| 1.1 TEJER..... | 278 |
| 1.2. TRAMAR..... | 292 |
| 1.3. URDIR | 298 |
| 1.4 COCINAR(SE) | 307 |
| 1.5 MAQUINAR | 314 |
| 2. Análisis | 317 |
| 2.1 Sentidos verbales..... | 317 |
| 2.2 Preferencias de selección | 319 |
| 2.3 Clases léxicas y entidades conceptuales | 320 |
| Conclusiones..... | 322 |
| 1. Cómo puede servir FunGramKB a REDES | 322 |
| 2. Cómo puede servir REDES a FunGramKB | 333 |
| 3. Futuras líneas de investigación..... | 339 |
| Bibliografía | 341 |
| Apéndice 1. Los metaconceptos y la representación semántica de los roles temáticos | 359 |
| Apéndice 2. La interpretación semántica de los papeles temáticos de los satélites..... | 361 |

ÍNDICE DE TABLAS

| | |
|---|--------|
| Tabla 1. MT y PS en FunGramKB de +OBTAIN_00 | xxviii |
| Tabla 2. Entrada de <i>malgastar</i> en REDES..... | 3 |
| Tabla 3. Distribución de clases léxicas por sentidos de <i>perder</i> en REDES..... | 5 |
| Tabla 4. El concepto <i>ALIMENTO</i> en REDES..... | 13 |
| Tabla 5. El concepto ALIMENTO en REDES | 14 |
| Tabla 6. Vínculos de las clases léxicas de <i>malgastar</i> al índice conceptual de REDES..... | 15 |
| Tabla 7. Principales diccionarios en inglés disponibles en Internet..... | 28 |
| Tabla 8. Diccionarios en español disponibles en Internet..... | 33 |
| Tabla 9. Relación de meronimia en EuroWordNet (tomado de Perrián y Arcas, 2007: 7)..... | 39 |
| Tabla 10. Marcos semánticos de <i>bake</i> en FrameNet..... | 41 |
| Tabla 11. El verbo <i>to give</i> en FrameNet | 42 |
| Tabla 12. Patrón de valencias de <i>to give</i> en FrameNet | 43 |
| Tabla 13. Ejemplo de significado relacional en EuroWordNet para el inglés..... | 71 |
| Tabla 14. Ejemplo de significado relacional en EuroWordNet para varias lenguas..... | 71 |
| Tabla 15. PS de +BIRD_00..... | 72 |
| Tabla 16. Descripción de +STORE_00 en FunGramKB | 76 |
| Tabla 17. Descripción de +PRESERVE_00 en FunGramKB..... | 77 |
| Tabla 18. Descripción de +BREED_00 en FunGramKB..... | 77 |
| Tabla 19. Descripción de +WEAR_00 en FunGramKB | 79 |
| Tabla 20. Descripción de \$SPORT_00 en FunGramKB | 80 |
| Tabla 21. Descripción de -TAKE_SHOES en FunGramKB | 81 |
| Tabla 22. Descripción de -WIELD en FunGramKB..... | 81 |
| Tabla 23. Descripción de +STORE en FunGramKB | 83 |
| Tabla 24. Perfil sintagmático de <i>to urge</i> (tomado de Hanks, 2013) | 101 |
| Tabla 25. Ejemplos de polisemia regular | 106 |
| Tabla 26. Par hiperonímico ‘Person’/ ‘Quality’ en EuroWordNet (Wing, 2002)..... | 126 |
| Tabla 27. Par hiperonímico ‘Control’/ ‘Trait’ en EuroWordNet (Wing, 2002)..... | 126 |
| Tabla 28. Par hiperonímico ‘profession’/‘discipline’ en WordNet..... | 127 |
| Tabla 29. Par hiperonímico ‘game’/‘equipment’ en WordNet | 127 |
| Tabla 30. Descripción de clase léxica en REDES que se combina con <i>cobrar</i> ‘adquirir’ | 134 |
| Tabla 31. Verbos estudiados y cantidad de clases léxicas con que se combinan..... | 137 |
| Tabla 32. Descripción de +STRENGTH_00 en FunGramKB..... | 138 |
| Tabla 33. Descripción de clase léxica en REDES que se combina con <i>malgastar</i> | 142 |
| Tabla 34. Descripción de clase léxica en REDES que se combina con <i>malgastar</i> | 143 |
| Tabla 35. Descripción de clase léxica en REDES que se combina con <i>disipar</i> (‘hacer desaparecer’)..... | 144 |
| Tabla 36. Descripción de clase léxica en REDES que se combina con <i>malgastar</i> | 145 |
| Tabla 37. Descripción de clase léxica en REDES que se combina con <i>malgastar</i> | 145 |

| | |
|---|-----|
| Tabla 38. Comparación entre clase léxica de REDES y +COURAGE_00 en FunGramKB | 146 |
| Tabla 39. Comparación entre clase léxica de REDES y +TRIAL_00 en FunGramKB | 147 |
| Tabla 40. Miembros de clase léxica que se combina con <i>dilapidar</i> y su vinculación con entidades de FunGramKB | 149 |
| Tabla 41. Comparación entre clase léxica de REDES y +ABILITY_00 en FunGramKB | 151 |
| Tabla 42. Resumen de información lexicográfica de <i>perder</i> ('desperdiciar') | 154 |
| Tabla 43. Descripción de +WASTE_00 en FunGramKB | 155 |
| Tabla 44. Descripción de +OBTAIN_00 en FunGramKB..... | 156 |
| Tabla 45. Comparación entre clase léxica que se combina con <i>adquirir</i> y +STRENGHT_00 en FunGramKB | 156 |
| Tabla 46. Comparación entre clases léxicas de REDES y +PLAN_00 en FunGramKB..... | 159 |
| Tabla 47. Sección de plantilla léxica de <i>dictar</i> [-COMMAND_PLAN] en FunGramKB | 160 |
| Tabla 48. Sentidos y combinatoria léxica de <i>dictar</i> , en REDES | 166 |
| Tabla 49. Descripción de +SAY_00 en FunGramKB..... | 166 |
| Tabla 50. Comparación entre clase léxica de REDES y +WRITING_00 en FunGramKB..... | 167 |
| Tabla 51. Descripción de \$DICTATE_00 en FunGramKB (concepto terminal propuesto)..... | 168 |
| Tabla 52. Descripción de +TEACH_00 en FunGramKB | 168 |
| Tabla 53. Comparación entre clase léxica de REDES y +SPEECH_00 en FunGramKB | 169 |
| Tabla 54. Descripción de +COMMAND_00 en FunGramKB | 169 |
| Tabla 55. Comparación entre clase léxica de REDES y +LAW_00 en FunGramKB | 170 |
| Tabla 56. Descripción de -COMMAND_LAW (subconcepto propuesto)..... | 171 |
| Tabla 57. Comparación entre clase léxica de REDES y +DIRECTION_00 en FunGramKB..... | 171 |
| Tabla 58. Descripción de -COMMAND_DIRECTION (subconcepto propuesto) | 172 |
| Tabla 59. Comparación entre clase léxica de REDES y +PLAN_00 en FunGramKB..... | 172 |
| Tabla 60. Descripción de -COMMAND_PLAN (subconcepto propuesto) | 173 |
| Tabla 61. Comparación entre clase léxica de REDES y +OCCURRENCE_00 en FunGramKB | 173 |
| Tabla 62. Comparación entre clase léxica de REDES y +MUSIC_00 en FunGramKB..... | 174 |
| Tabla 63. Rutas conceptuales de conceptos vinculados a clases léxicas seleccionadas por <i>dictar</i> [+COMMAND_00] | 174 |
| Tabla 64. Resumen: <i>DICTAR</i> en FunGramKB (codificación propuesta) | 175 |
| Tabla 65. Tipología de clases léxicas que se combinan con <i>dictar</i> con respecto a FunGramKB..... | 175 |
| Tabla 66. Sentidos y combinatoria léxica de <i>perder</i> , en REDES | 177 |
| Tabla 67. Comparación entre clase léxica de REDES y +VEHICLE_00 en FunGramKB | 178 |
| Tabla 68. Comparación entre clase léxica de REDES y +TRAVEL_00 en FunGramKB | 178 |
| Tabla 69. Descripción de +FAIL_00 en FunGramKB..... | 179 |
| Tabla 70. Descripción de \$MISS_00 (concepto terminal propuesto) | 179 |
| Tabla 71. Descripción de +WASTE_00 en FunGramKB..... | 180 |
| Tabla 72. Comparación entre clase léxica de REDES y +PERIOD_00 en FunGramKB..... | 180 |
| Tabla 73. Descripción de -WASTE_PERIOD (subconcepto propuesto)..... | 181 |
| Tabla 74. Descripción de +FAIL_00 en FunGramKB..... | 181 |

| | |
|--|-----|
| Tabla 75. Comparación entre clase léxica de REDES y +TRIAL_00 en FunGramKB | 182 |
| Tabla 76. Descripción de clase léxica en REDES que se combina con <i>perder</i> ‘resultar vencido’ | 182 |
| Tabla 77. Descripción de +BATTLE_00 en FunGramKB | 183 |
| Tabla 78. Descripción de +COMPETITION_00 en FunGramKB..... | 183 |
| Tabla 79. Descripción de -FAIL_TRIAL (subconcepto propuesto) | 184 |
| Tabla 80. Descripción de -FAIL_BATTLE (subconcepto propuesto)..... | 184 |
| Tabla 81. Descripción de -FAIL_COMPETITION (subconcepto propuesto)..... | 184 |
| Tabla 82. Descripción de +LOSE_00 en FunGramKB..... | 185 |
| Tabla 83. Comparación entre clase léxica de REDES y +SOLID_00 en FunGramKB | 185 |
| Tabla 84. Descripción de +LACK_00 en FunGramKB..... | 186 |
| Tabla 85. Comparación entre clase léxica de REDES y +PHYSICAL_ATT_00 en FunGramKB | 186 |
| Tabla 86. Comparación entre clase léxica de REDES y +SENSE_00 en FunGramKB | 187 |
| Tabla 87. Comparación entre clase léxica de REDES y +DESIRE_00 en FunGramKB | 187 |
| Tabla 88. Comparación entre clase léxica de REDES y +COGNITIVE_ATT en FunGramKB..... | 188 |
| Tabla 89. Comparación entre clase léxica de REDES y +FEELING_00 en FunGramKB..... | 189 |
| Tabla 90. Comparación entre clase léxica de REDES y +BEHAVIOUR_00 en FunGramKB..... | 189 |
| Tabla 91. Comparación entre clase léxica de REDES y +CUSTOM_00 en FunGramKB..... | 190 |
| Tabla 92. Comparación entre clase léxica de REDES y +SUBSTANCE_00 en FunGramKB..... | 190 |
| Tabla 93. Comparación entre clase léxica de REDES y +DIRECTION_00 en FunGramKB..... | 191 |
| Tabla 94. Rutas conceptuales de conceptos vinculados a clases léxicas que se combinan con <i>perder</i> [+LACK_00]..... | 191 |
| Tabla 95. Descripción de -LACK__PSYCHOLOGICAL_ATT (concepto propuesto) | 192 |
| Tabla 96. Descripción de -LACK__SUBSTANCE (concepto propuesto) | 192 |
| Tabla 97. Descripción de -LACK__DIRECTION (concepto propuesto)..... | 192 |
| Tabla 98. Descripción de -LACK__CUSTOM (concepto propuesto)..... | 192 |
| Tabla 99. Resumen: <i>PERDER</i> en FunGramKB (vinculación propuesta) | 193 |
| Tabla 100. Tipología de clases léxicas que se combinan con <i>perder</i> con respecto a FunGramKB.... | 194 |
| Tabla 101. Sentidos y combinatoria léxica de <i>disipar</i> | 195 |
| Tabla 102. Descripción de +SPLIT_00 en FunGramKB | 196 |
| Tabla 103. Comparación clase léxica de REDES y +GAS_00 en FunGramKB | 196 |
| Tabla 104. Comparación clase léxica de REDES y +ENERGY_00 en FunGramKB | 197 |
| Tabla 105. Descripción de \$DISSIPATE_00 en FunGramKB (concepto propuesto)..... | 197 |
| Tabla 106. Descripción de +WASTE_00 en FunGramKB | 198 |
| Tabla 107. Comparación entre clase léxica de REDES y +MONEY_00 en FunGramKB..... | 198 |
| Tabla 108. Descripción de -WASTE_MONEY (subconcepto propuesto)..... | 199 |
| Tabla 109. Descripción de +FINISH_00 en FunGramKB..... | 199 |
| Tabla 110. Comparación entre clase léxica de REDES y +BELIEF_00 en FunGramKB..... | 200 |
| Tabla 111. Comparación entre clase léxica de REDES y +THOUGHT_00 en FunGramKB | 200 |
| Tabla 112. Comparación entre clase léxica de REDES y +DESIRE_00 en FunGramKB | 201 |

| | |
|--|-----|
| Tabla 113. Comparación entre clase léxica de REDES y +MYSTERY_00 en FunGramKB | 201 |
| Tabla 114. Comparación entre clase léxica de REDES y +DANGER_00 en FunGramKB..... | 202 |
| Tabla 115. Comparación entre clase léxica de REDES y +PROBLEM_00 en FunGramKB | 203 |
| Tabla 116. Comparación entre clase léxica de REDES y +FEAR_00 en FunGramKB | 203 |
| Tabla 117. Comparación entre clase léxica de REDES y +SADNESS_00 en FunGramKB | 204 |
| Tabla 118. Comparación entre clase léxica de REDES y +ANGER_00 en FunGramKB | 205 |
| Tabla 119. Comparación entre clase léxica de REDES y \$ENTHUSIASM_00 en FunGramKB..... | 205 |
| Tabla 120. Rutas conceptuales de conceptos vinculados a clases léxicas que se combinan con <i>disipar</i> [+FINISH_00] | 206 |
| Tabla 121. Descripción de -FINISH_FEELING (subconcepto propuesto) | 206 |
| Tabla 122. Descripción de -FINISH_DANGER (subconcepto propuesto) | 206 |
| Tabla 123. Descripción de -FINISH_PROBLEM (subconcepto propuesto) | 207 |
| Tabla 124. Descripción de -FINISH_THOUGHT (subconcepto propuesto)..... | 207 |
| Tabla 125. Descripción de -FINISH_BELIEF (subconcepto propuesto) | 207 |
| Tabla 126. Resumen: <i>DISIPAR</i> en FunGramKB (vinculación propuesta) | 207 |
| Tabla 127. Tipología de clases léxicas que se combinan con <i>disipar</i> con respecto a FunGramKB ... | 209 |
| Tabla 128. Sentidos y combinatoria léxica de <i>derrochar</i> | 210 |
| Tabla 129. Descripción de +WASTE_00 en FunGramKB | 211 |
| Tabla 130. Comparación entre clase léxica de REDES y +MONEY_00 en FunGramKB..... | 212 |
| Tabla 131. Información lexicográfica de <i>derrochar</i> ‘malgastar en exceso’ | 212 |
| Tabla 132. Descripción de \$SQUANDER_00 (concepto propuesto)..... | 213 |
| Tabla 133. Descripción de \$ABOUND_00 | 213 |
| Tabla 134. Comparación entre clase léxica de REDES y +PLEASURE_00 en FunGramKB | 214 |
| Tabla 135. Comparación entre clase léxica de REDES y +COURAGE_00 en FunGramKB | 214 |
| Tabla 136. Comparación entre clase léxica de REDES y +EFFORT_00 en FunGramKB | 215 |
| Tabla 137. Comparación entre clase léxica de REDES y +WISDOM_00 en FunGramKB..... | 215 |
| Tabla 138. Comparación entre clase léxica de REDES y +STRENGHT_00 en FunGramKB | 216 |
| Tabla 139. Comparación entre clase léxica de REDES y +ABILITY_00 en FunGramKB | 217 |
| Tabla 140. Comparación entre clase léxica de REDES y +IMAGINATION_00 en FunGramKB | 217 |
| Tabla 141. Comparación entre clase léxica de REDES y \$FEATURE_00 en FunGramKB | 218 |
| Tabla 142. Comparación entre clase léxica de REDES y +HUMOUR_00 en FunGramKB | 219 |
| Tabla 143. Comparación entre clase léxica de REDES y +PSYCHOLOGICAL_ATT_00 en FunGramKB | 220 |
| Tabla 144. Comparación entre clase léxica de REDES y +ANGER_00 en FunGramKB | 220 |
| Tabla 145. Comparación entre clase léxica de REDES y +PRIDE_00 en FunGramKB..... | 221 |
| Tabla 146. Rutas conceptuales de conceptos vinculados a clases léxicas que se combinan con <i>derrochar</i> [\$ABOUND_00] | 221 |
| Tabla 147. Descripción de -ABOUND_PSYCHOLOGICAL_ATT (subconcepto propuesto)..... | 222 |
| Tabla 148. Descripción de -ABOUND_STRENGHT (subconcepto propuesto)..... | 222 |
| Tabla 149. Descripción de -ABOUND_EFFORT (subconcepto propuesto)..... | 222 |

| | |
|--|-----|
| Tabla 150. Resumen: <i>DERROCHAR</i> en FunGramKB (vinculación propuesta)..... | 223 |
| Tabla 151. Tipología de clases léxicas que se combinan con <i>derrochar</i> con respecto a FunGramKB224 | |
| Tabla152. Sentidos y combinatoria léxica de <i>malograr</i> | 225 |
| Tabla 153. Descripción de +WASTE_00 en FunGramKB | 225 |
| Tabla 154. Comparación entre clase léxica de REDES y +FUTURE_00 en FunGramKB..... | 226 |
| Tabla 155. Comparación de clase léxica de REDES y +ABILITY_00 en FunGramKB | 226 |
| Tabla 156. Descripción de -WASTE_FUTURE en FunGramKB (subconcepto propuesto)..... | 227 |
| Tabla 157. Descripción de -WASTE_ABILITY en FunGramKB (subconcepto propuesto) | 227 |
| Tabla 158. Descripción de +FAIL_00 en FunGramKB..... | 227 |
| Tabla 159. Comparación entre clase léxica de REDES y +OCCURRENCE_00 en FunGramKB | 228 |
| Tabla 160. Comparación entre clase léxica de REDES y +PLAN_00 en FunGramKB..... | 229 |
| Tabla 161. Comparación entre clase léxica de REDES y +EFFORT_00 en FunGramKB | 229 |
| Tabla 162. Comparación entre clase léxica de REDES y +VICTORY_00 en FunGramKB | 230 |
| Tabla 163. Rutas conceptuales de conceptos vinculados a clases léxicas que se combinan con <i>malograr</i> [FAIL+_00] | 230 |
| Tabla 164. Descripción de -FAIL_PLAN (subconcepto propuesto) | 231 |
| Tabla 165. Descripción de -FAIL_EFFORT (subconcepto propuesto)..... | 231 |
| Tabla 166. Descripción de -FAIL_VICTORY (subconcepto propuesto) | 231 |
| Tabla 167. Resumen: <i>MALOGRAR</i> en FunGramKB (vinculación propuesta)..... | 232 |
| Tabla 168. Tipología de clases léxicas que se combinan con <i>malograr</i> con respecto a FunGramKB232 | |
| Tabla 169. Descripción de +WASTE_00 en FunGramKB | 233 |
| Tabla 170. Combinatoria léxica de <i>malgastar</i> , en REDES..... | 233 |
| Tabla 171. Descripción de clase léxica que se combina con <i>malgastar</i> | 234 |
| Tabla 172. Comparación entre clase léxica de REDES y +PERIOD_00 en FunGramKB..... | 234 |
| Tabla 173. Comparación entre clase léxica de REDES y +STRENGHT_00 en FunGramKB | 235 |
| Tabla 174. Comparación entre clase léxica de REDES y +ABILITY_00 en FunGramKB | 235 |
| Tabla 175. Comparación entre clase léxica de REDES y +VICTORY_00 en FunGramKB | 236 |
| Tabla 176. Comparación entre clase léxica de REDES y +INFORMATION_OBJECT_00 en FunGramKB | 237 |
| Tabla 177. Comparación entre clase léxica de REDES y +BULLET_00 en FunGramKB..... | 237 |
| Tabla 178. Rutas conceptuales de conceptos vinculados a clases léxicas que se combinan con <i>malgastar</i> [WASTE+_00] | 238 |
| Tabla 179. Descripción de -WASTE_PERIOD (subconcepto propuesto)..... | 238 |
| Tabla 180. Descripción de -WASTE_STRENGTH (subconcepto propuesto) | 238 |
| Tabla 181. Descripción de -WASTE_ABILITY (subconcepto propuesto) | 238 |
| Tabla 182. Descripción de -WASTE_VICTORY (subconcepto propuesto) | 239 |
| Tabla 183. Descripción de -WASTE_INFORMATION_OBJECT (subconcepto propuesto) | 239 |
| Tabla 184. Descripción de -WASTE_BULLET (subconcepto propuesto)..... | 239 |
| Tabla 185. Resumen: <i>MALGASTAR</i> en FunGramKB (vinculación propuesta)..... | 239 |
| Tabla 186. Tipología de clases léxicas que se combinan con <i>malgastar</i> con respecto a FunGramKB240 | |

| | |
|---|-----|
| Tabla 187. Descripción de +WASTE_00 en FunGramKB | 241 |
| Tabla 188. Combinatoria léxica de <i>dilapidar</i> , en REDES | 241 |
| Tabla 189. Comparación entre clase léxica de REDES y +MONEY_00 en FunGramKB..... | 242 |
| Tabla 190. Comparación entre clase léxica de REDES y +EFFORT_00 en FunGramKB | 243 |
| Tabla 191. Comparación entre clase léxica de REDES y +PERIOD_00 en FunGramKB..... | 243 |
| Tabla 192. Comparación entre clase léxica de REDES y +FAME_00 en FunGramKB | 244 |
| Tabla 193. Comparación entre clase léxica de REDES y +ABILITY_00 en FunGramKB | 244 |
| Tabla 194. Comparación entre clase léxica de REDES y +DESIRE_00 en FunGramKB | 245 |
| Tabla 195. Comparación entre clase léxica de REDES y +FEELING_00 en FunGramKB..... | 245 |
| Tabla 196. Rutas conceptuales de conceptos vinculados a clases léxicas que se combinan con <i>dilapidar</i> [+WASTE]..... | 246 |
| Tabla 197. Descripción de -WASTE_ABILITY (subconcepto propuesto) | 246 |
| Tabla 198. Descripción de -WASTE_FEELING (subconcepto propuesto) | 246 |
| Tabla 199. Descripción de -WASTE_PERIOD (subconcepto propuesto)..... | 247 |
| Tabla 200. Descripción de -WASTE_FAME (subconcepto propuesto)..... | 247 |
| Tabla 201. Descripción de -WASTE_MONEY (subconcepto propuesto) | 247 |
| Tabla 202. Descripción de -WASTE_EFFORT (subconcepto propuesto) | 247 |
| Tabla 203. Resumen: <i>DILAPIDAR</i> en FunGramKB (vinculación propuesta)..... | 247 |
| Tabla 204. Tipología de clases léxicas que se combinan con <i>dilapidar</i> con respecto a FunGramKB | 248 |
| Tabla 205. Descripción de +WASTE_00 en FunGramKB | 249 |
| Tabla 206. Combinatoria léxica de <i>desaprovechar</i> , en PRÁCTICO | 249 |
| Tabla 207. Comparación entre clase léxica de REDES y +ABILITY_00 en FunGramKB | 250 |
| Tabla 208. Comparación entre clase léxica de REDES y +STRENGHT_00 en FunGramKB | 250 |
| Tabla 209. Comparación entre clase léxica de REDES y +TIME_00 en FunGramKB | 251 |
| Tabla 210. Comparación entre clase léxica de REDES y +INFORMATION_00 en FunGramKB ... | 251 |
| Tabla 211. Rutas conceptuales de conceptos vinculados a clases léxicas que se combinan con <i>desaprovechar</i> [WASTE+_00]..... | 252 |
| Tabla 212. Descripción de -WASTE_TIME (subconcepto propuesto) | 252 |
| Tabla 213. Descripción de -WASTE_STRENGTH (subconcepto propuesto) | 252 |
| Tabla 214. Descripción de -WASTE_ABILITY (subconcepto propuesto) | 252 |
| Tabla 215. Descripción de -WASTE_INFORMATION (subconcepto propuesto) | 253 |
| Tabla 216. Resumen: <i>DESAPROVECHAR</i> en FunGramKB (vinculación propuesta) | 253 |
| Tabla 217. Tipología de clases léxicas que se combinan con <i>desaprovechar</i> con respecto a FunGramKB | 254 |
| Tabla 218. Combinatoria léxica de <i>despilfarrar</i> , en PRÁCTICO | 254 |
| Tabla 219. Descripción de +WASTE_00 en FunGramKB | 255 |
| Tabla 220. Información lexicográfica de <i>despilfarrar</i> ‘malgastar en exceso’ | 255 |
| Tabla 221. Descripción de \$\$QUANDER_00 (concepto terminal propuesto)..... | 255 |
| Tabla 222. Comparación entre clase léxica de REDES y +MONEY_00 en FunGramKB..... | 256 |
| Tabla 223. Comparación entre clase léxica de REDES y +FOOD_00 en FunGramKB | 256 |

| | |
|--|-----|
| Tabla 224. Comparación entre clase léxica de REDES y +ABILITY_00 en FunGramKB | 257 |
| Tabla 225. Comparación entre clase léxica de REDES y +BULLET_00 en FunGramKB | 257 |
| Tabla 226. Rutas conceptuales de conceptos vinculados a clases léxicas que se combinan con <i>despilfarrar</i> [\$\$QUANDER_00] | 258 |
| Tabla 227. Descripción de -SQUANDER_ABILITY (subconcepto propuesto) | 258 |
| Tabla 228. Descripción de -SQUANDER_MONEY (subconcepto propuesto) | 258 |
| Tabla 229. Descripción de -SQUANDER_FOOD (subconcepto propuesto) | 258 |
| Tabla 230. Descripción de -SQUANDER_BULLET (subconcepto propuesto)..... | 258 |
| Tabla 231. Resumen: <i>DESPILFARRAR</i> en FunGramKB (vinculación propuesta) | 259 |
| Tabla 232. Tipología de clases léxicas que se combinan con <i>despilfarrar</i> | 259 |
| Tabla 233. Combinatoria léxica de <i>desperdiciar</i> en PRÁCTICO..... | 260 |
| Tabla 234. Comparación entre clase léxica de REDES y +PERIOD_00 en FunGramKB..... | 260 |
| Tabla 235. Comparación entre clase léxica de REDES y +ABILITY_00 en FunGramKB | 261 |
| Tabla 236. Comparación entre clase léxica de REDES y +MONEY_00 en FunGramKB..... | 261 |
| Tabla 237. Comparación entre clase léxica de REDES y +STRENGHT_00 en FunGramKB | 262 |
| Tabla 238. Comparación entre clase léxica de REDES y +BULLET_00 en FunGramKB..... | 262 |
| Tabla 239. Comparación entre clase léxica de REDES y +FOOD_00 en FunGramKB | 263 |
| Tabla 240. Rutas conceptuales de conceptos vinculados a clases léxicas que se combinan con <i>desperdiciar</i> [+WASTE_00] | 263 |
| Tabla 241. Descripción de -WASTE_ABILITY (subconcepto propuesto) | 264 |
| Tabla 242. Descripción de -WASTE_STRENGTH (subconcepto propuesto) | 264 |
| Tabla 243. Descripción de -WASTE_PERIOD (subconcepto propuesto)..... | 264 |
| Tabla 244. Descripción de -WASTE_FOOD (subconcepto propuesto) | 264 |
| Tabla 245. Descripción de -WASTE_BULLET (subconcepto propuesto)..... | 264 |
| Tabla 246. Descripción de -WASTE_MONEY (subconcepto propuesto) | 264 |
| Tabla 247. Resumen: <i>DESPERDICAR</i> en FunGramKB (vinculación propuesta) | 265 |
| Tabla 248. Tipología de clases léxicas que se combinan con <i>desperdiciar</i> | 265 |
| Tabla 249. Vinculación de los sentidos de los predicados a conceptos de FunGramKB | 266 |
| Tabla 250. Polisemia regular entre <i>PERDER</i> y <i>MALOGRAR</i> | 267 |
| Tabla 251. Distribución de clases léxicas entre verbos monosémicos y polisémicos en su vínculo con +WASTE_00 | 268 |
| Tabla 252. Selección de entidades por predicados vinculados con +WASTE_00, y subconceptos propuestos | 269 |
| Tabla 253. Colocaciones en FunGramKB de verbos vinculados con -WASTE_ABILITY | 270 |
| Tabla 254. Descripción de +ABILITY_00 en FunGramKB..... | 271 |
| Tabla 255. Clases de REDES (y PRÁCTICO) vinculadas a +ABILITY_00 | 272 |
| Tabla 256. Clases léxicas de REDES (y PRÁCTICO) vinculadas a +TIME_00/+PERIOD_00/+FUTURE_00 | 273 |
| Tabla 257. Clases de REDES (y PRÁCTICO) vinculadas a +MONEY_00..... | 274 |
| Tabla 258. Clases de REDES (y PRÁCTICO) vinculadas a +STRENGTH_00 | 274 |

| | |
|---|-----|
| Tabla 259. Clases de REDES (y PRÁCTICO) vinculadas a +BULLET_00..... | 275 |
| Tabla 260. Clases de REDES (y PRÁCTICO) vinculadas a +FOOD_00 | 275 |
| Tabla 261. Clases de REDES (y PRÁCTICO) vinculadas a la noción de ‘recurso’ | 276 |
| Tabla 262. Clases de REDES (y PRÁCTICO) vinculadas a la noción de ‘posibilidad’ | 276 |
| Tabla 263. Sentidos y combinatoria léxica de <i>tejer</i> , en REDES..... | 279 |
| Tabla 264. Descripción de +WEAVE_00 en FunGramKB | 280 |
| Tabla 265. Comparación entre clase léxica de REDES y +CLOTH_00 en FunGramKB..... | 280 |
| Tabla 266. Descripción de \$CONSPIRE_00 en FunGramKB..... | 281 |
| Tabla 267. Comparación entre clase léxica de REDES y +PLAN_00 en FunGramKB..... | 282 |
| Tabla 268. Descripción de +WRITE_00 en FunGramKB | 282 |
| Tabla 269. Comparación entre clase léxica de REDES y +WRITING_00 en FunGramKB | 283 |
| Tabla 270. Comparación entre clase léxica de REDES y +STORY_00 en FunGramKB | 284 |
| Tabla 271. Descripción de \$EXPLAIN_00 en FunGramKB..... | 285 |
| Tabla 272. Comparación entre clase léxica de REDES y +EXPLANATION_00 en FunGramKB .. | 285 |
| Tabla 273. Descripción de +DESIRE_01 en FunGramKB..... | 286 |
| Tabla 274. Comparación entre clase léxica de REDES y +DESIRE_00 en FunGramKB | 286 |
| Tabla 275. Descripción de +FEEL_00 en FunGramKB | 287 |
| Tabla 276. Comparación entre clase léxica de REDES y +FEELING_00 en FunGramKB..... | 287 |
| Tabla 277. Descripción de +CREATE_00 en FunGramKB | 288 |
| Tabla 278. Comparación entre clase léxica de REDES y +PROBLEM_00 en FunGramKB | 289 |
| Tabla 279. Resumen: <i>TEJER</i> en FunGramKB (vinculación propuesta)..... | 290 |
| Tabla 280. Tipología de clases léxicas que se combinan con <i>tejer</i> con respecto a FunGramKB..... | 291 |
| Tabla 281. Sentidos y combinatoria léxica de <i>tramar</i> | 292 |
| Tabla 282. Descripción de +WEAVE_00 en FunGramKB | 293 |
| Tabla 283. Descripción de +THINK_00 en FunGramKB | 293 |
| Tabla 284. Comparación entre clase léxica de REDES y +THOUGHT_00 en FunGramKB..... | 294 |
| Tabla 285. Descripción de \$CONSPIRE_00 en FunGramKB..... | 294 |
| Tabla 286. Comparación entre clase léxica de REDES y +CRIME_00 en FunGramKB..... | 295 |
| Tabla 287. Comparación entre clase léxica de REDES y +PLAN_00 en FunGramKB..... | 296 |
| Tabla 288. Comparación entre clase léxica de REDES y +VIOLENCE_00 en FunGramKB | 297 |
| Tabla 289. Resumen: <i>TRAMAR</i> en FunGramKB (vinculación propuesta)..... | 297 |
| Tabla 290. Tipología de clases léxicas que se combinan con <i>tramar</i> con respecto a FunGramKB .. | 298 |
| Tabla 291. Combinatoria léxica de <i>urdir</i> , en REDES..... | 299 |
| Tabla 292. Descripción de +WEAVE_00 en FunGramKB | 300 |
| Tabla 293. Descripción de +WRITE_00 en FunGramKB..... | 300 |
| Tabla 294. Comparación entre clase léxica de REDES y +STORY_00 en FunGramKB | 301 |
| Tabla 295. Descripción de \$EXPLAIN_00 en FunGramKB..... | 301 |
| Tabla 296. Comparación entre clase léxica de REDES y +EXPLANATION_00 en FunGramKB .. | 302 |
| Tabla 297. Descripción de +AGREE_00 en FunGramKB | 303 |

| | |
|---|-----|
| Tabla 298. Descripción de \$CONSPIRE_00 en FunGramKB..... | 303 |
| Tabla 299. Comparación entre clase léxica de REDES y +PLAN_00 en FunGramKB..... | 304 |
| Tabla 300. Comparación entre clase léxica de REDES y +VIOLENCE_00 en FunGramKB | 305 |
| Tabla 301. Resumen: <i>URDIR</i> en FunGramKB (vinculación propuesta)..... | 306 |
| Tabla 302. Tipología de clases léxicas que se combinan con <i>urdir</i> con respecto a FunGramKB | 307 |
| Tabla 303. Combinatoria léxica de <i>cocinar(se)</i> , en REDES..... | 308 |
| Tabla 304. Descripción de +COOK_00 en FunGramKB | 308 |
| Tabla 305. Comparación entre clase léxica de REDES y +FOOD_00 en FunGramKB | 309 |
| Tabla 306. Descripción de +AGREE_00 en FunGramKB | 309 |
| Tabla 307. Descripción de +WRITE_00 en FunGramKB | 310 |
| Tabla 308. Comparación entre clase léxica de REDES y +WRITING_00 en FunGramKB..... | 310 |
| Tabla 309. Descripción de \$CONSPIRE_00 en FunGramKB..... | 311 |
| Tabla 310. Descripción de +CREATE_00 en FunGramKB | 311 |
| Tabla 311. Comparación entre clase léxica de REDES y +PLAN_00 en FunGramKB..... | 312 |
| Tabla 312. Comparación entre clase léxica de REDES y +LAW_00 en FunGramKB..... | 312 |
| Tabla 313. Resumen: <i>COCINAR(SE)</i> en FunGramKB (vinculación propuesta)..... | 313 |
| Tabla 314. Tipología de clases léxicas que se combinan con <i>cocinar(se)</i> con respecto a FunGramKB..... | 314 |
| Tabla 315. Combinatoria léxica de <i>maquinar</i> , en REDES | 314 |
| Tabla 316. Descripción de +OPERATE_00 en FunGramKB..... | 315 |
| Tabla 317. Descripción de \$CONSPIRE_00 en FunGramKB..... | 315 |
| Tabla 318. Comparación entre clase léxica de REDES y +PLAN_00 en FunGramKB..... | 316 |
| Tabla 319. Comparación entre clase léxica de REDES y +CRIME_00 en FunGramKB..... | 316 |
| Tabla 320. Resumen: <i>MAQUINAR</i> en FunGramKB (vinculación propuesta)..... | 317 |
| Tabla 321. Tipología de clases léxicas que se combinan con <i>maquinar</i> con respecto a FunGramKB..... | 317 |
| Tabla 322. Vínculos de los sentidos predicativos con eventos conceptuales en FunGramKB..... | 318 |
| Tabla 323. Polisemia regular (directa e indirecta) de cuatro verbos..... | 319 |
| Tabla 324. Subconceptos que reflejan preferencias de selección | 320 |
| Tabla 325. Distribución de clases léxicas por tipo..... | 320 |
| Tabla 326. Preferencias de selección de predicados | 321 |
| Tabla 327. Cantidad de eventos conceptuales vinculados con cada verbo | 325 |
| Tabla 328. Cantidad de subconceptos para cada verbo..... | 326 |
| Tabla 329. Cantidad de clases léxicas por sentido por predicado..... | 336 |

ÍNDICE DE IMÁGENES

| | |
|--|-----|
| Imagen 1. Entrada de <i>lose</i> en el <i>Historical Thesaurus of English</i> (vista parcial) | 29 |
| Imagen 2. Categoría de ‘Slim shape/physique’ en el <i>Historical Thesaurus of English</i> (vista parcial). 29 | |
| Imagen 3. Entrada de <i>lose</i> en http://dictionary.cambridge.org | 30 |
| Imagen 4. Definiciones de <i>lose</i> (NO LONGER POSSESS) en http://dictionary.cambridge.org | 31 |
| Imagen 5. Tesouro visual de la palabra <i>lose</i> (http://www.visualthesaurus.com) | 31 |
| Imagen 6. Entrada de <i>lose</i> en WordNet (http://wordnet.princeton.edu/) | 37 |
| Imagen 7. Tesouro visual de la palabra <i>lose</i> en WordNet (http://www.visuwords.com)..... | 38 |
| Imagen 8. Marco semántico de ‘Finish_competition’ en FrameNet (vista parcial)..... | 45 |
| Imagen 9. Vista en FunGramKB de \$INCLUDE_00 | 53 |
| Imagen 15. Plantilla léxica de <i>atesorar</i> en el lexicón | 83 |
| Imagen 16. +PERIOD_00 en FunGramKB (vista parcial) | 143 |
| Imagen 17. +ANGER_00 en FunGramKB (vista parcial)..... | 144 |

SIGLARIO
(Diccionarios, lexicones, ontologías)

- BBI** Benson, M., Benson, E. y Ilson, R. (1986). *The BBI Combinatory Dictionary of English. A Guide to Word Combinations*. Amsterdam: John Benjamins.
- CCEC** *Collins Cobuild English Collocations on CD ROM. A comprehensive Database of Common Word Patterns from the Bank of English*. (1995). Londres: Harper Collins.
- CLAVE** *CLAVE. Diccionario de Uso del Español Actual*. (2000). Madrid: SM. En <http://clave.librosvivos.net>.
- CyC** Ontología de CyCorp. En <http://cyc.com>.
- DEA** Seco, M., Andrés, O y Ramos, G. (1999). *Diccionario del Español Actual*. Madrid: Aguilar.
- DEC** I. Mel'čuk et al. *Dictionnaire explicatif et combinatoire du français contemporain, Recherches lexico-sémantiques*, Vol 1 (1984), Vol 2 (1988), Vol 3 (1992), Vol 4 (1999). Montréal: Les Presses de L'Université de Montréal.
- DiCE** Alonso Ramos, M. (directora). *Diccionario de colocaciones del español*. Universidad de A Coruña. En <http://www.dicesp.com/paginas>.
- DOLCE** *Descriptive Ontology for Linguistic and Cognitive Engineering*. En <http://www.loa.istc.cnr.it/DOLCE.html>.
- DRAE** *Diccionario de la Real Academia Española*, XXII edición actualizada en Internet. (2001). Madrid: Espasa. En <http://rae.es/rae.html>.
- DUE** *María Moliner. Diccionario de Uso del Español*. (2002). Madrid: Gredos.
- FrameNet** *FrameNet Project*. Berkeley, California: International Computer Science Institute. En: <https://framenet.icsi.berkeley.edu/fndrupal>.
- FunGramKB** Funcional Grammar Knowledge Base. En <http://www.fungramkb.com>.
- LAF** Polguère, A. *Lexique active du français*. (En preparación). Universidad de Montréal.
- LDOCE** Procter, P. (1978). *Longman's Dictionary of Contemporary English* (actualizado en Internet). En <http://www.ldoceonline.com/dictionary>.
- LTP** Hill, J. y Lewis, E. (Eds). (1997). *LTP Dictionary of Selected Collocations*. Londres: English Teaching Publications.
- OCD** *Oxford Collocations Dictionary for Students of English*. (2002). Oxford University Press.
- OED** *Oxford English Dictionary*. En <http://www.oed.com>.
- SUMO** *Suggested Upper Merged Ontology* En <http://www.ontologyportal.org>.

VOX *Diccionarios VOX y Larousse.* En <http://www.diccionarios.com>.

WEBode Ontología de Ontology Engineering Group. En <http://mayor2.dia.fi.upm.es/oeg>.

ABREVIATURAS¹

| | |
|-----------------|--|
| CL | Clase léxica (REDES) |
| COHERENT | Conceptualization Hierarchization Remodeling Refinement (FunGramKB) |
| COREL | Conceptual Representation Language (FunGramKB) |
| CPA | Corpus Pattern Analysis (TNE) |
| EA | Estructura argumental (TLG) |
| EE | Estructura eventiva (TLG) |
| ETL | Estructura de tipificación léxica (TLG) |
| EQ | Estructura de <i>Qualia</i> (TLG) |
| FE | Frame Element(s) (FrameNet) |
| GF | Grammatical Function (FrameNet) |
| IA | Inteligencia artificial |
| LU | Lexical unit (FrameNet) |
| MLC | Modelo Léxico Construccional (Lexicom) |
| MRD | Machine Readable Dictionaries |
| MT | Marco temático (FunGramKB) |
| MMTT | Machine Translation |
| NSM | Natural Semantic Metalanguage (Wierzbicka, 1972 y posterior) |
| PDM | Proyecto de Definiciones Mínimas (Bosque y Mairal, 2012ab) |
| PLN | Procesamiento del lenguaje natural |
| POS | Part of speech |
| PS | Postulado de significado (FunGramKB) |
| PT | Phrase Type (FrameNet) |
| RRG | Role and Reference Grammar (Van Valin y LaPolla, 1997; Van Valin 2005) |
| TLG | Teoría del Lexicón Generativo (Pustejovsky, 1995 y posterior) |
| TNE | Theory of Norms and Exploitations (Hanks, 2013) |
| WSD | Word Sense Disambiguation |

¹ Proveemos entre paréntesis el modelo al que pertenece la sigla o abreviatura.

INTRODUCCIÓN

Esta tesis se plantea como tarea general vincular la información lingüística contenida en *REDES Diccionario combinatorio del español contemporáneo* (Bosque, 2004), en adelante REDES, con el marco ontológico de *Functional Grammar Knowledge Base*, en adelante FunGramKB (Periñán y Arcas, 2004, 2007ab, 2010ab; Periñán y Mairal, 2010, entre otros), una base de conocimiento léxico conceptual diseñada para el procesamiento del lenguaje natural (PLN). El ánimo de combinar a REDES y FunGramKB, dos recursos con base lingüística, pero pertenecientes a campos diferentes –la lexicografía y la ingeniería del conocimiento, respectivamente–, enmarca este trabajo en lo que se ha llamado la lexicografía informatizada del siglo XXI o lexicografía del tercer milenio (Fuertes y Tarp, 2011).

Esta nueva lexicografía se compone de recursos lexicográficos informatizados mucho más complejos que las conocidas versiones digitales de los diccionarios tradicionales. Se trata de bases de datos léxicas o bases de conocimiento léxico, de mayor o menor complejidad y profundidad, construidas en plataformas electrónicas. El diseño particular de cada recurso, junto con el tipo de datos que lo pueble, permitirá la explotación y aplicación de la información lingüística de maneras determinadas. Aunque la lexicografía informatizada también busca servir a la consulta humana, su mira está más enfocada hacia el PLN, una subdisciplina del campo de la Inteligencia Artificial (IA) que propone mecanismos computacionales para la comprensión del lenguaje natural entre personas y máquinas. Entre las aplicaciones más conocidas del PLN, figuran la traducción automática, la respuesta a preguntas, la extracción de información y el reconocimiento del habla.

Se reconoce que el problema fundamental del PLN es la desambiguación semántica automática de las palabras polisémicas (*Word Sense Disambiguation*, en adelante WSD). Si todas las palabras del idioma tuvieran un significado único e inequívoco, los programas de PLN simplemente asignarían ese significado o su equivalente en otras lenguas. Sin embargo, la mayoría de las palabras son polisémicas, y una gestión exitosa de PLN dependerá de poder asignar a cada palabra el sentido correcto en cada contexto. La asignación de sentidos que ocurre automáticamente en la mente del hispanohablante cuando enfrenta el verbo *perder* en combinaciones como *perder el partido*, *perder la cartera* o *perder el tren*, tendrá que ser programada explícitamente para efectos de una máquina. Gran parte del trabajo que se realiza en el ámbito del PLN está enfocado, por lo tanto, en encontrar estrategias para la desambiguación efectiva de palabras en su contexto.

FunGramKB forma parte de la lexicografía informatizada del siglo XXI: es una base de conocimiento léxico conceptual multilingüe y multipropósito, diseñada para servir tareas del PLN, así como consultas de lingüistas, traductores o estudiantes de idiomas. FunGramKB busca

representar la totalidad del conocimiento lingüístico humano en una plataforma electrónica, por lo que su construcción es resultado de la colaboración entre lingüistas e ingenieros del conocimiento. El rol de la ingeniería del conocimiento se refleja en el hecho de que el pivote central de la arquitectura de FunGramKB no es un registro de palabras, sino una ontología de conceptos universales de naturaleza cognitiva, que no pertenecen a ninguna lengua y que fueron razonados *a priori* de cualquier tarea o aplicación. Al igual que otros sistemas de PLN, FunGramKB busca reducir información sin perder poder expresivo; la ontología incluye un total de 1,634 conceptos básicos, divididos en tres subontologías: entidades, eventos y atributos. A la ontología conceptual central se conectan los módulos léxico gramaticales particulares de cada lengua; se dice, pues, que las unidades léxicas particulares “lexicalizan” los conceptos de la ontología. Los conceptos de cada subontología se lexicalizan por diferentes categorías gramaticales: la entidades son lexicalizadas por nombres, los eventos por verbos y las cualidades por adjetivos.

Del otro lado tenemos al diccionario REDES que, a pesar de llamarse “diccionario”, no pertenece a la lexicografía tradicional. Aunque el formato impreso y la macroestructura de lemas ordenados alfabéticamente lo acercan a un diccionario típico, su contenido atañe más bien a la investigación lingüística de orden léxico semántico. REDES se dedica a presentar las restricciones semánticas que imponen los predicados a su selección léxica argumental. En concreto, REDES nos informa de que ciertos predicados se combinan sistemáticamente con ciertos argumentos. Esta selección no ocurre de predicado a argumento, sino de predicado a “clase léxica”, un conjunto de argumentos que expresa una misma noción semántica². Por ejemplo, el verbo *cobrar*, en su sentido de ‘adquirir’, se combina con una clase léxica de sustantivos que denotan ‘vida o vivacidad’, que incluye los argumentos *vida*, *viveza*, *vitalidad* y otra que denota ‘importancia, prominencia o actualidad’, que está compuesta por *importancia*, *protagonismo*, *relevancia*, *relieve*, *entidad*, *vigencia*, *actualidad*, *trascendencia*, *primacía*³. En total, REDES estudia la combinatoria de unos 4,000 predicados, organizándola en cada entrada por clases léxicas. Siguiendo el principio metodológico de la “navaja de Occam”⁴, el diccionario identifica todas las clases léxicas que considera imprescindibles, aunque reconoce lo deseable de que “se puedan reducir o redefinir en trabajos posteriores” (Bosque, 2004b: CL). Esto significa que los 4,000 predicados se combinan con varios miles de clases léxicas, que están compuestas a su vez por varios miles más de argumentos.

Esta investigación está inspirada por la idea de que la estructura de una base de conocimiento léxico conceptual como FunGramKB permitiría sistematizar el contenido de REDES

² El concepto de ‘clase léxica’, según propuesto por REDES, se discutirá en detalle en el capítulo I.

³ Estas son dos de las 15 clases léxicas con que se combina *cobrar* en REDES (en sus diferentes sentidos), lo que da una idea de la notable casuística de datos lingüísticos contenida en este diccionario.

⁴ Que dicta que “las entidades no deben multiplicarse innecesariamente”.

para explotarlo a cabalidad, de maneras que son imposibles en el formato impreso actual. El vínculo de REDES con FunGramKB podría dar pie a una versión informatizada de este diccionario, que nos permitiría ver más claramente las relaciones entre predicados y argumentos, o mejor, entre sentidos predicativos y sentidos argumentales. Al mismo tiempo, el contenido de REDES –que no se ha recogido en ningún otro diccionario o investigación– resultaría utilísimo a una base de conocimiento léxico conceptual como FunGramKB y, en particular, a la gestión de WSD, si se logran vincular los sentidos de una palabra con su combinatoria léxica. Un programa de WSD se guiaría por la combinatoria de la palabra para asignarle un sentido u otro. Este conocimiento sería ventajoso también para aplicaciones de producción de lenguaje natural, porque ayudaría a representar con más naturalidad la combinatoria regular en la lengua.

La investigación está motivada, pues, por dos preguntas generales de trabajo, que precisaremos a lo largo de esta introducción:

- (1) ¿Cómo puede servir FunGramKB a REDES?
- (2) ¿Cómo puede servir REDES a FunGramKB?

Aunque podría parecer que REDES y FunGramKB pertenecen a territorios distantes, la tesis parte de la premisa de que la vinculación entre ambos recursos no solo es posible, sino que supone grandes beneficios para cada uno. La conexión entre ambos pone en contacto, además, dos cúmulos de información valiosa que implican un extenso trabajo previo: REDES aporta patrones de la combinatoria léxica sistemática de palabras predicativas, tomadas del uso real en corpus y analizadas minuciosamente, y FunGramKB aporta una plataforma informática pensada para el PLN, diseñada en torno a una ontología jerarquizada de conceptos universales cognitivos.

Una respuesta preliminar a la pregunta (1), ¿Cómo puede servir FunGramKB a REDES?, y la primera hipótesis general del trabajo, es que la ontología conceptual en torno a la cual se articula FunGramKB puede servir de eje o pivote para uniformar, reducir y poder relacionar más explícitamente los datos lingüísticos de REDES. El vincular las diferentes unidades e informaciones de REDES a un mismo eje conceptual nos permitiría ver patrones y hacer generalizaciones sobre la relación entre significado y selección léxica que no son perceptibles en el formato actual del diccionario, en otras palabras, nos permitiría explorar y explotar más cabalmente las *redes* de REDES. Como dijimos antes, esta vinculación puede constituir un primer paso hacia convertir a REDES en un diccionario electrónico: no una réplica digital de la versión escrita del diccionario, sino un diccionario “inteligente” diseñado en un entorno y un lenguaje que pueda comprender, gestionar y explotar una máquina. La posibilidad de sacar el máximo provecho a los datos de REDES radica en gran medida en las herramientas de la lexicografía informatizada del siglo XXI.

Una respuesta preliminar a la pregunta (2), ¿Cómo puede servir REDES a FunGramKB?, y la segunda hipótesis general del trabajo, es que los datos de REDES pueden ayudar a poblar las estructuras de FunGramKB, enriqueciendo tanto el nivel conceptual como el léxico. La información empírica de REDES sobre las restricciones sistemáticas de selección léxica de palabras predicativas –tomada de corpus y clasificada semánticamente– será utilísima para cualquier recurso informatizado que busque representar, procesar –o, incluso, entender, como es el caso de FunGramKB (Periñán y Arcas, 2010a)– el conocimiento léxico conceptual de la mente humana. De ese conocimiento, una parcela particular e importante está contenida en REDES: la selección léxica semántica sistemática. Aunque los datos léxicos de REDES procedan del español, un vínculo satisfactorio con el aparato de FunGramKB confirmaría que algunas de estas restricciones tienen una motivación semántica y cognitiva de carácter universal; por otro lado, los casos que no permitan el vínculo conceptual constituirían evidencia de lexicalizaciones idiosincrásicas del español.

En este estudio, nos interesa también la alternancia regular de sentidos o polisemia regular. Nos referimos al fenómeno por el cual palabras como *escuela*, *colegio* o *universidad* pueden alternar regularmente entre los sentidos de ‘edificio’ o ‘institución’, por ejemplo. La polisemia regular es muy valiosa para el PLN, pues una vez identificado un patrón, se puede extender a todas las unidades léxicas que tengan uno u otro sentido, lo cual significa un ahorro significativo en el etiquetado o procesamiento de las mismas. Por ejemplo, en el caso anterior, se podría establecer un proceso automático mediante el cual a toda palabra que signifique ‘edificio’ se le asigne también el significado de ‘institución’. La mayoría de los estudios hechos sobre la polisemia regular se han enfocado en sustantivos, mientras que las investigaciones sobre verbos han estado orientadas a alternancias aspectuales, del tipo causativo/incoativo. En nuestro estudio, analizamos la polisemia regular en un grupo de verbos en términos estrictamente semánticos, a la luz de la codificación en FunGramKB de los sentidos que alternan.

De lo anterior se desprenden algunos de los desafíos que presenta el trabajo. El primero tiene que ver con la naturaleza de la información que contiene cada recurso, y las maneras en que se representa esa información. Como señala Velardi (1991), en la lingüística, el conocimiento semántico se modela con expresiones profundas, más o menos formales, y se enfoca en aspectos específicos de la comunicación⁵, mientras que en las bases de datos para el PLN, las entradas léxicas típicamente expresan el conocimiento lingüístico de sentido común que entienden y usan los humanos, codificado en un lenguaje de representación de conocimiento manipulable por una máquina. En este caso, REDES no usa formalismos, sino que presenta y describe la información

⁵ Como hemos comentado antes, aunque REDES se presenta como un producto lexicográfico, en gran medida responde a un trabajo de investigación lingüística.

semántica en lenguaje natural, como vimos arriba en los descriptores de las clases léxicas que se combinan con *cobrar*: “sustantivos que denotan vida o vivacidad” y “sustantivos que denotan importancia, prominencia o actualidad”. Por su parte, FunGramKB codifica la información en COREL (*Conceptual Representation Language*), un lenguaje formal de representación semántico conceptual. Por ejemplo, *cobrar* se vincularía en FunGramKB al evento conceptual +OBTAIN_00, que tiene un marco temático (MT) que dice que hay dos participantes: un Tema ‘humano’⁶ y un Referente ‘artefactual’, y un postulado de significado (PS) que equivale aproximadamente a ‘entrar en posesión de algo’⁷:

Tabla 1. MT y PS en FunGramKB de +OBTAIN_00

| | |
|------------|--|
| MT: | (x1: +HUMAN_00)Theme (x2: +ARTEFACT_00)Referent |
| PS: | +(e1: +DO_00 (x1)Theme (x3)Referent (f1: (e2: +HAVE_00 (x1)Theme (x2)Referent))Result) +(e3: past n +HAVE_00 (x1)Theme (x2)Referent) |

Un segundo reto, relacionado con el anterior, es que los datos de REDES tienen un punto de partida lingüístico, mientras que los de FunGramKB parten de lo conceptual. Los datos de restricción de selección de REDES son primeramente léxicos; de ellos se abstraen, inductivamente y *a posteriori*, las motivaciones semánticas subyacentes. De este proceso, emergen las clases léxicas, que están constituidas por miembros que cumplen dos criterios: (a) se combinan sistemáticamente con el mismo predicado y (b) comparten un mismo sentido en la combinación. Se trata pues de clases léxico semánticas “sensibles a la combinatoria” (Bosque, 2004b), no de clases establecidas por los rasgos inherentes de significado de sus miembros. El modelo de FunGramKB, por su parte, es primeramente conceptual. La ontología de conceptos se diseña según sus propios criterios, que toman en cuenta los conceptos de nivel alto establecidos por otras ontologías actuales reconocidas, y posteriormente se valida lingüísticamente. La ontología organiza los conceptos según sus significados inherentes, en una jerarquía taxonómica con herencia de rasgos semánticos de los conceptos superordinados a los subordinados. A pesar de la diferencia en la manera en que orientan la información, tanto REDES como FunGramKB presentan un contenido semántico que permite la comparación. En el caso de REDES, el contenido semántico se codifica en el descriptor de cada clase léxica y, en menor grado, en el “índice conceptual”⁸, mientras que en FunGramKB, el contenido semántico figura en tanto en el MT como en el PS de cada concepto.

⁶ A lo largo de la tesis, utilizaremos las comillas simples para demarcar sentidos o significados. En algunas ocasiones, para facilitar la lectura, las usamos para referirnos en español a un concepto de FunGramKB, por ejemplo, ‘humano’ por +HUMAN_00. Sin embargo, es importante recordar que no se trata de una equivalencia semántica y que cada concepto de FunGramKB tiene una definición formal codificada en su PS.

⁷ El lenguaje COREL y las nociones de marco temático (MT) y postulado de significado (PS) se discutirán a fondo en el Capítulo II.

⁸ Los índices conceptuales de REDES se discutirán en el Capítulo I.

El tercer reto es la reducción de los datos lingüísticos de REDES en el proceso de vinculación con el aparato conceptual de FunGramKB. Al comparar la cantidad de conceptos de la ontología de FunGramKB con el número de clases léxicas en REDES –o de entradas léxicas y acepciones contenidas en cualquier diccionario–, se hace claro que la vinculación de datos lingüísticos a una ontología conceptual supone una reducción significativa⁹. Al contemplar la posibilidad de reducir el inventario de clases léxicas de REDES, Bosque destaca que “tal vez la estrategia reductora se pueda abordar como tarea de investigación, es decir, como búsqueda encaminada a precisar la ontología que nos permita explicar la forma en que llevamos a cabo la selección léxica” (2004b: CLI). Sin duda, la reducción es atractiva y, de hecho, necesaria, para la vinculación con FunGramKB, pero como el mismo autor advierte: “se trata de una cuestión empírica, no de una cuestión teórica” (Bosque, 2004b: CLI), que implicaría la comprobación de que el comportamiento combinatorio de las palabras que constituyen una clase léxica es igual en todos los contextos.

A continuación, precisamos las dos preguntas iniciales. Dado que REDES contiene tres unidades de información (los predicados, los argumentos seleccionados por los predicados y las clases léxicas que agrupan a los argumentos), exploraremos cómo responde cada una a la representación conceptual en FunGramKB, y cómo aporta cada una al enriquecimiento de FunGramKB. Guiaremos el trabajo con una serie de preguntas abiertas que enfoquen diferentes aspectos de la interacción entre el diccionario combinatorio REDES y la base de conocimiento léxico conceptual FunGramKB.

Desde la primera perspectiva, mirando los datos de REDES desde el marco de FunGramKB, formulamos las siguientes interrogantes:

(1.1) ¿Cómo sirve el entorno ontológico de FunGramKB para sacar el máximo provecho de los datos de REDES, es decir, para reducir, organizar y relacionar conceptualmente la casuística lingüística de REDES?

(1.2) ¿Pueden vincularse los predicados de REDES –o, mejor, los sentidos que asumen los predicados– a la ontología conceptual de FunGramKB?

(1.3) ¿Puede reducirse el inventario de clases léxicas de REDES usando la ontología conceptual de FunGramKB? ¿Se mantiene la clase léxica como unidad ante la ontología de FunGramKB?

⁹ Por ejemplo, frente a los 1,634 conceptos básicos de FunGramKB, el DRAE (2001) tiene 88,431 entradas de lemas simples con 161,962 acepciones. Para vincular los datos del DRAE a FunGramKB, tendríamos que “traducir” esas 161,962 acepciones a 1,634 conceptos, o sea, asignar unas 100 acepciones a cada concepto. Cuando hablamos de “reducción”, nos referimos a ese proceso justamente. REDES tiene un leuario de 4,000 predicados, que se combinan con decenas de miles de clases léxicas, por lo que su vinculación con FunGramKB supone una reducción de los datos también.

(1.4) ¿Pueden estos vínculos revelar patrones o generalizaciones sobre las relaciones sistemáticas entre los sentidos de un predicado y su combinatoria?

- ¿Puede tener un predicado más de un sentido con la misma combinatoria?

- ¿Puede tener un predicado el mismo sentido con diferentes combinaciones?

(1.5) ¿Pueden diversos predicados compartir la misma alternancia de sentidos, es decir, participar en patrones de polisemia regular? ¿Qué relación tiene la polisemia regular de los predicados con su selección léxica sistemática?

(1.6) ¿Cómo nos ayudan las nociones de ‘preferencia de selección’ –de naturaleza conceptual– y ‘colocación’¹⁰ –de naturaleza léxica–, establecidas por FunGramKB, a entender mejor los datos de REDES?

(1.7) ¿Constituye el vínculo de REDES a FunGramKB un primer paso para convertir a REDES en un diccionario informatizado inteligente?

Desde la segunda perspectiva, al considerar el beneficio de incorporar la información lingüística de REDES a FunGramKB, las preguntas que intentará responder esta tesis son:

(2.1) ¿Cómo pueden los datos de REDES enriquecer a FunGramKB, tanto en el nivel conceptual, por medio de nuevos conceptos y preferencias de selección, como en el nivel léxico, con lexicalizaciones de conceptos y colocaciones?

(2.2) ¿Tiene la relación entre un predicado y su combinatoria léxica en REDES un correlato conceptual en la ontología de FunGramKB?

(2.3) ¿Pueden los datos de combinatoria léxica de REDES aportar al tratamiento de la polisemia en FunGramKB?

(2.4) ¿Puede la estructura de FunGramKB identificar patrones de polisemia regular?

(2.5) ¿Nos permiten los datos de REDES identificar principios y patrones regulares que den consistencia a las nociones clave para el PLN de ‘preferencia de selección’ y ‘colocación’, establecidas por FunGramKB?

Para contestar estas preguntas, nuestro trabajo explora la representación conceptual en FunGramKB de las relaciones entre significado y selección léxica de 15 verbos descritos en REDES¹¹. La primera parte de la investigación examina la relación entre los sentidos y la combinatoria léxica de 10 verbos polisémicos, y analiza su vinculación con el aparato de FunGramKB. Este conjunto incluye el verbo *dictar*, que se analiza independientemente, y un grupo de nueve verbos relacionados semánticamente entre sí: *derrochar*, *desaprovechar*, *desperdiciar*, *despilfarrar*, *dilapidar*, *disipar*, *malgastar*, *malograr* y *perder*. La segunda parte de

¹⁰ Las nociones de ‘preferencia de selección’ y ‘colocación’ serán discutidas en detalle en el Capítulo 3.

¹¹ En el trabajo se usa también el *Diccionario combinatorio práctico del español contemporáneo* (Bosque, 2006), en adelante PRÁCTICO, basado en REDES y publicado dos años después.

la investigación se enfoca en el fenómeno de la polisemia regular. Estudia la relación entre los sentidos y combinatoria léxica de cinco verbos que evidencian el fenómeno de la alternancia regular de sentidos, y su representación conceptual en FunGramKB. Estos verbos son: *tejer*, *tramar*, *urdir*, *cocinar(se)*, *maquinar*.

La tesis se organiza como sigue: el capítulo I describe, caracteriza y valora el contenido de REDES, destacando su importancia para cualquier modelo que busque representar el conocimiento lingüístico humano, específicamente el fenómeno de las restricciones de selección sistemáticas. Argumenta que REDES revela afinidades con las tendencias que organizan la lexicografía informatizada del siglo XXI, lo que posibilita una transferencia a la plataforma electrónica; asimismo, detalla los retos que supone ese proceso.

El capítulo II repasa los rasgos más sobresalientes de la lexicografía informatizada actual, tanto los logros alcanzados como las metas no concretadas aún, para situar en este panorama el modelo de FunGramKB, el marco teórico y práctico de este estudio. Pone énfasis especial en las bases de datos léxicas WordNet, EuroWordNet y FrameNet.

El capítulo III describe y valora la base de conocimiento FunGramKB, otorgando atención particular a la ontología, a las propiedades de los conceptos que la pueblan y al manejo de las preferencias de selección y las colocaciones.

El capítulo IV discute el problema de la polisemia y de la desambiguación de sentidos en los ámbitos de la lexicografía tradicional y electrónica. Da atención particular a las estrategias de WSD usadas en el PLN, y al rol de las preferencias de selección para la WSD.

El capítulo V discute el fenómeno de la polisemia regular, primero desde la perspectiva teórica y luego en el ámbito de la lexicografía electrónica. También se discuten los principales estudios realizados sobre polisemia regular en diferentes lexicones electrónicos.

El capítulo VI explica la metodología usada en la investigación.

El capítulo VII presenta y analiza los resultados. Primero se discute la vinculación a FunGramKB del verbo *dictar* y su combinatoria léxica, y luego la de los nueve verbos relacionados semánticamente entre sí: *derrochar*, *desaprovechar*, *desperdiciar*, *despilfarrar*, *dilapidar*, *disipar*, *malgastar*, *malograr* y *perder*.

El capítulo VIII discute la representación en FunGramKB de la polisemia regular evidenciada por cinco verbos relacionados semánticamente entre sí: *tejer*, *tramar*, *urdir*, *cocinar(se)*, *maquinar*.

El capítulo IX presenta las conclusiones, contestando las preguntas de trabajo que guiaron el estudio, y establece futuras líneas de investigación.

CAPÍTULO I. REDES. DICCIONARIO COMBINATORIO DEL ESPAÑOL CONTEMPORÁNEO

Introducción

Nuestro trabajo busca volcar el contenido de REDES en la base de conocimiento léxico conceptual FunGramKB con el objetivo general de enriquecer ambos recursos. REDES es un diccionario impreso en papel cuyo novel contenido sobre combinatoria léxica es idóneo –por las razones que elaboraremos en este capítulo– para ser traducido a una base de datos léxica en el marco de la lexicografía informatizada moderna. FunGramKB presenta las mayores posibilidades de explotación de los datos de REDES, por ser una base de conocimiento representativa del nivel más ambicioso de la lexicografía informatizada.

En la sección 1 de este capítulo, haremos una presentación general de los contenidos del diccionario REDES. Como sugiere el nombre REDES, los componentes de este diccionario –predicados, argumentos y clases léxicas– se interrelacionan en un entramado de cierta complejidad, que se manifiesta en los niveles léxico y semántico de la lengua. Ilustraremos estas interrelaciones a la luz de una entrada representativa del diccionario. Luego dedicaremos una sección a cada uno de los componentes del diccionario: en la sección 2 discutiremos los predicados, y en la 2.1, las posibles clases de predicados que se desprenden de los datos de REDES; en la sección 3 comentaremos las clases léxicas que agrupan a los argumentos; y en la sección 4, los argumentos mismos. En la sección 5 discutiremos el índice conceptual de REDES, un componente valioso para nuestro trabajo. En la sección 6, discutiremos la noción de restricción de selección léxica que presenta REDES. En la sección 6.1, ubicaremos la aportación de la noción de restricción de selección léxica de REDES en el marco de otras líneas de investigación teóricas que han abordado el fenómeno de la combinatoria léxica sistemática, como las restricciones selectivas y el *Aksionart*, y en la sección 6.2, la consideraremos ante otros acercamientos más prácticos a este fenómeno, como los diccionarios de colocaciones. A lo largo del capítulo, señalaremos los rasgos que hacen al diccionario REDES único en la lexicografía y algunos de los retos que presenta la traducción de su contenido a un entorno ontológico como el de FunGramKB. Terminaremos el capítulo con una conclusión.

1. Presentación general de REDES¹²

REDES es un diccionario de restricciones de selección léxica¹³ sin precedentes en la lexicografía hispánica o general. Se basa en un corpus de 250 millones de palabras, provenientes de 68 publicaciones periódicas españolas e hispanoamericanas de los últimos 20 años. La combinatoria de cada predicado se extrae del corpus, se organiza inductivamente por criterios semánticos y finalmente se presenta en un formato lexicográfico.

REDES incluye 4,000 lemas predicativos –principalmente verbos, adjetivos y adverbios, pero también algunas locuciones–, y los argumentos con los que estos se combinan sistemáticamente. En cada artículo lexicográfico, los argumentos se organizan en clases léxicas, las cuales se establecen inductivamente considerando la cercanía semántica de los argumentos. Los componentes semánticos que dan cohesión a la clase léxica son capturados en el “descriptor”, una descripción en lenguaje natural del significado compartido por todos los argumentos en una combinación con un predicado dado, como, por ejemplo, el grupo de sustantivos “que denotan fuerza o energía, usados en su sentido figurado o en el físico. También con otros que se refieren a algunos de sus efectos”. Las entradas “analíticas” o entradas “largas” de REDES¹⁴ contienen, por lo tanto, tres unidades de información: predicados, argumentos y clases léxicas. Estas informaciones se interrelacionan en el diccionario, a través de una serie de índices numéricos y alfabéticos y de referencias cruzadas.

Para ilustrar el contenido de una entrada analítica de REDES, en la siguiente tabla reproducimos la entrada del verbo *malgastar*:

¹² Se recomienda la lectura del ensayo titulado “Combinatoria y significación. Algunas reflexiones”, de Ignacio Bosque, que se presenta como prólogo a REDES (2004b), al igual que el artículo “Bases para un diccionario de restricciones léxicas” (2001b). Asimismo es útil la lectura de la “Presentación” al PRÁCTICO (Bosque, 2006)

¹³ Su nombre original fue *Diccionario de Restricciones Léxicas* (DRL), y más tarde, por razones editoriales, cambió a REDES.

¹⁴ Estas entradas incluyen la combinatoria detallada de los predicados, y se oponen a las “entradas abreviadas” o “entradas cortas”, que se componen exclusivamente de referencias cruzadas a las primeras.

Tabla 2. Entrada de *malgastar* en REDES¹⁵

| |
|--|
| <p>malgastar. v.</p> <p>Se combina con sustantivos que designan magnitudes, bienes materiales, recursos: <i>dinero, agua, corcho, petróleo, gasolina, papel</i> y unidades temporales, incluidas las que se refieren a la existencia: <i>vida, juventud, día</i>.</p> |
| <p>Se combina especialmente con:</p> |
| <p>A. Sustantivos que denotan fuerza o energía, usados en su sentido figurado o en el físico. También con otros que se refieren a algunos de sus efectos: 1 <i>fuerza</i>, 2 <i>energía</i>, 3 <i>esfuerzo</i>, 4 <i>salud</i>, 5 <i>sudor</i></p> |
| <p>B. Sustantivos que designan facultades, aptitudes o cualidades humanas, especialmente las de carácter intelectual. También con otros que designan diversos estados de reconocimiento personal: 6 <i>talento</i>, 7 <i>capacidad</i>, 8 <i>don</i>, 9 <i>cualidad</i>, 10 <i>generosidad</i>, 11 <i>comicidad</i>, 12 <i>valor</i>, 13 <i>credibilidad</i>, 14 <i>prestigio</i>, 15 <i>experiencia</i></p> |
| <p>C. Sustantivos que denotan oportunidad u ocasión, así como ventaja u otro tipo de circunstancia favorable. Se usan muy frecuentemente en el ámbito deportivo: 16 <i>oportunidad</i>, 17 <i>ocasión</i>, 18 <i>balón</i>, 19 <i>ventaja</i>, 20 <i>punto</i>, 21 <i>opción</i>, 22 <i>privilegio</i>, 23 <i>posibilidad</i></p> |
| <p>D. Sustantivos que denotan resultado feliz de algo: 24 <i>éxito</i>, 25 <i>victoria</i>, 26 <i>triunfo</i></p> |
| <p>E. Sustantivos que designan unidades, materias y recursos informativos, expresivos o artísticos: 27 <i>saliva</i>, 28 <i>línea</i>, 29 <i>palabra</i>, 30 <i>página</i>, 31 <i>óleo</i>, 32 <i>tinta</i>, 33 <i>cinta</i></p> |
| <p>F. Sustantivos que denotan munición. También con otros que designan algunas acciones en las que se usa o formas en que se almacena: 34 <i>disparo</i>, 35 <i>munición</i>, 36 <i>cartuchera</i>, 37 <i>bala</i></p> |

Para interpretar y valorar el contenido de esta entrada, conviene aclarar primero el uso de las letras mayúsculas y de los números. Las letras mayúsculas identifican las clases léxicas en propiedad, en este caso, de la A a la F. Estas se distinguen de las “clases léxicas abiertas”, que aparecen en la “entradilla”, al comienzo del artículo, sin letra mayúscula. Ambos tipos de clases léxicas se describirán y comentarán en este capítulo. Los números, por su parte, identifican cada argumento que se incluye en el artículo lexicográfico. Por ejemplo, la clase léxica A se compone de: 1 *fuerza*, 2 *energía*, 3 *esfuerzo*, 4 *salud*, 5 *sudor*. La utilidad de estos números radica en las “referencias cruzadas a las voces”, que son entradas donde los argumentos seleccionados son los lemas y el artículo se compone de los predicados seleccionadores. La palabra *fuerza*, por ejemplo, tiene una entrada como lema en REDES, que incluye todos los predicados que la seleccionan, entre ellos *malgastar*¹. Esto expresa que en la entrada de *malgastar*, el argumento #1 será *fuerza*, como comprobamos arriba. Cada componente de la entrada analítica será discutida en detalle a continuación.

¹⁵ Hemos eliminado los ejemplos de uso por cuestión de espacio. Los ejemplos de uso provienen del corpus e ilustran la combinación de cada argumento con el predicado.

2. Los predicados

Los lemas de REDES son únicamente palabras predicativas, e incluyen verbos, adjetivos, adverbios y algunas locuciones, por ejemplo, *delimitar*, *blando*, *enérgicamente* o *de capa caída*. No se incluyen todos los predicados, solo una selección de predicados que presentan restricciones de combinación por criterios estrictamente lingüísticos. Se excluyen pues predicados como *comer*, cuya combinatoria en el sentido literal es predecible de nuestro conocimiento de mundo, o *estupendo*, que se podría aplicar prácticamente a todas las cosas.

Aunque el propósito de REDES no es definir estas palabras, el diccionario reconoce que hay un vínculo entre el significado de una palabra y su combinatoria: “las palabras no significan algo y ADEMÁS se combinan de cierta manera, sino que en gran medida se combinan de cierta manera PORQUE expresan precisamente esos significados” (Bosque, 2004b: XVIII-XIX, énfasis del autor). Este hecho es de particular interés para nuestro trabajo, pues una de las preguntas que investigamos es si el vínculo entre REDES y FunGramKB puede revelar patrones sobre las relaciones sistemáticas entre los sentidos de un predicado y su combinatoria. Sin embargo, la conexión entre significado y combinatoria léxica queda implícita en el diccionario; solo cuando es estrictamente necesario, REDES proporciona información explícita sobre las diferentes acepciones de los lemas. En el caso de *malgastar*, por ejemplo, REDES no da ninguna indicación de su significado, probablemente porque se trata de un verbo monosémico. En otros casos, REDES hace una distinción entre el “sentido físico” y el “sentido figurado”, sin aportar contenido semántico adicional, como en *derrochar*. En el caso de las palabras polisémicas, cuando es necesario establecer distinciones de sentido para poder comprender la combinatoria que se presenta, REDES usa una marca léxica sencilla. Por ejemplo, para el predicado *perder*, REDES distingue cuatro sentidos: (a) ‘quedar desprovisto de’; (b) ‘desaprovechar’; (c) ‘resultar vencido en’; (d) ‘no alcanzar a tiempo’.

Estas indicaciones sobre los sentidos de los predicados polisémicos son mucho más generales que las definiciones de un diccionario tradicional e incluso diluyen deliberadamente las diferencias entre acepciones. Por ejemplo, REDES ofrece las cuatro nociones semánticas simples que vimos antes para *perder*, mientras que la versión actual del DRAE presenta 27 acepciones para ese verbo. Bosque (2006) explica que una definición demasiado restrictiva podría excluir las combinaciones léxicas reales de un predicado. Por ejemplo, la acepción #3 de *espolear* en el DRAE (2001, versión electrónica) es ‘avivar, incitar, estimular a alguien para que haga algo’, una definición más detallada que la pista de

sentido que se ofrece en PRÁCTICO: ‘animar, estimular’. De hecho, los sustantivos que se combinan con *espolear* en PRÁCTICO no tienen fácil acomodo en la definición del DRAE, pues, además de las *personas*, que están incluidas en el contorno “a alguien”, también se pueden espolear las *necesidades*, el *miedo*, la *práctica*, el *crecimiento*, etc. Dado lo anterior, podemos decir que la granularidad semántica de REDES para sus lemas predicativos se acerca mucho más a la granularidad de la lexicografía electrónica –que maneja inventarios relativamente reducidos de sentido, como veremos en el capítulo IV–, que a la granularidad de la lexicografía tradicional, cuyos límites son de naturaleza principalmente editorial.

La combinatoria sistemática de cada predicado se convierte, pues, en el gran indicador potencial de sus sentidos, aunque reiteramos que no es ese el objetivo principal del trabajo de REDES. En esa línea de pensamiento, un hecho de interés es que, cuando una palabra tiene varios sentidos en REDES, los grupos combinatorios tienden a distribuirse claramente entre ellas. Conocer la correspondencia entre sentidos y combinatoria es fundamental, naturalmente, para un mecanismo de WSD fundamentado en la combinatoria léxica. En el caso de *perder*, señalado arriba, las 17 clases léxicas se dividen nítidamente entre los cuatro sentidos. Queda claro, sin embargo, que la asignación de clases léxicas a los sentidos implica una agrupación de las clases en conjuntos mayores, como vemos en la tabla a continuación:

Tabla 3. Distribución de clases léxicas por sentidos de *perder* en REDES

| Sentido | Número de clases léxicas |
|-----------------------------|--------------------------|
| (a) ‘quedar desprovisto de’ | 11 clases léxicas |
| (b) ‘desaprovechar’ | 2 clases léxicas |
| (c) ‘resultar vencido en’ | 2 clases léxicas |
| (d) ‘no alcanzar a tiempo’ | 2 clases léxicas |

Hay que advertir, por otro lado, que no siempre es perfecta la correspondencia entre sentidos y combinatoria. Bosque (2006) cita el sustantivo *goteo*, cuyos sentidos ‘acción de caer gotas’ y ‘manifestación intermitente de algo’ sí tienen combinaciones casi idénticas con algunos predicados (*goteo continuo*, *cesar el goteo*, etc.). Añade que, por el contrario, el verbo *cancelar* se combina con sustantivos diferentes (en rol de objeto directo) en cada uno de sus dos sentidos: ‘saldar, liquidar’ (*cancelar una deuda*) y ‘anular, dejar sin efecto’ (*cancelar una inversión*). Sin embargo, en ambos sentidos se combina con los mismos adverbios (*cancelar una deuda/una inversión repentinamente*, *abruptamente*, etc.).

Aunque REDES no proponga definiciones para los predicados, las clases léxicas aportan información intensional y extensional sobre estos. La información intensional la

proporcionan las clases léxicas como noción semántica, mientras que la extensional la proporcionan las listas de argumentos. Aunque las clases léxicas no caracterizan el significado del lema, ayudan a comprender el hecho de que determinadas unidades léxicas sean las que proporcionan la extensión de cada concepto. La pregunta que formula Bosque en el ensayo introductorio de REDES es, pues, fundamental: ¿cuál es la relación entre la intensión de un concepto y su extensión? O vista con un ejemplo específico: ¿qué nos dice sobre el significado de *malgastar* el hecho de que se combine sistemáticamente con sustantivos que denotan ‘fuerza o energía’, como *fuerza* o *energía*? Como vimos en la Tabla 2, los datos de REDES sugieren que existe una relación entre sentidos y clases léxicas, pero no es de 1:1; más bien, un mismo sentido puede ocurrir con múltiples clases léxicas. Bosque (2004b) propone que en muchos casos la intensión de la palabra se queda igual mientras que la extensión es lo que varía.

Esto supone cuestionar fundamentalmente la noción de ‘significado’. Por ejemplo, Bosque plantea que aunque el paso de los sentidos físicos a los figurados pueda proporcionar nuevas paráfrasis del significado de una palabra, “no es evidente que estas paráfrasis constituyan necesariamente [...] nuevos significados de la voz que se intenta analizar” (2004b: CXIII). Veamos algunos ejemplos: el verbo *acarrear* significa ‘transportar’ (*el trigo*) u ‘ocasionar’ (*disgustos*). Un sentido más general, como ‘llevar consigo’, daría cabida a los dos sentidos previos, si se detalla cómo varía la extensión. Por su parte, el predicado *diluirse* puede aplicarse al *azúcar* pero también a las *responsabilidades*; por lo tanto, este verbo tal vez no necesita dos definiciones, sino dos paradigmas de combinatoria lingüística. Bosque (2004b) defiende que es más lógico añadir extensiones en la mente que intensiones. Además, señala que el añadir significados puede nublar la conexión entre ellos. Por ejemplo, *destapar* puede usarse en *destapar una caja* (‘quitar la tapa’) o en *destapar un secreto* (‘sacar a la luz pública’). En este caso, el deslindar los dos sentidos “ayudará [al lector] a descifrar los textos en los que aparezca *destapar*, pero habrá de ser él quien llegue personalmente a la conclusión de que, en realidad, *destapar* es un sola cosa” (Bosque, 2004b: CXIV). Esta visión coincide con el acercamiento al significado de FunGramKB, como veremos en el capítulo III.

REDES aboga entonces por una reducción en la representación del significado, pero también por una ampliación de las extensiones, información que no es deducible de las definiciones. Esto presenta un reto para la lexicografía tradicional, pero se acomoda bastante bien a los axiomas de una base de conocimiento con centro conceptualista como FunGramKB. Como veremos en el capítulo III, FunGramKB incluye una unidad muy

apropiada para este objetivo, el “subconcepto”, que permite codificar diferentes preferencias de selección sin cambiar la definición de un concepto dado.

2.1 Clases de predicados

Los predicados no se organizan explícitamente como clases en REDES, pero no es difícil constituir clases posibles. En nuestro trabajo, de hecho, hemos establecido dos grupos de predicados relacionados semánticamente entre sí, con el objetivo de observar patrones en sus significados y combinatorias. El primero es: *derrochar*, *desaprovechar*, *desperdiciar*, *despilfarrar*, *dilapidar*, *disipar*, *malgastar*, *malograr* y *perder*, y el segundo: *tramar*, *tejer*, *urdir*, *cocinar(se)*, y *maquinar*.

Un mecanismo para identificar posibles clases de predicados en REDES es hacer referencia a las palabras relacionadas que aparecen al final de muchas entradas, tras la indicación “Véase también”. Por ejemplo, la entrada de *malgastar* remite a *derrochar*, *dilapidar*, *escatimar*, *gastar*, entre otros verbos. Según se rastrean esas referencias, se pueden ir formulando clases de predicados con rasgos semánticos compartidos. Una segunda vía para plantear clases de predicados es usar el índice conceptual de REDES, que discutiremos en la sección 6 de este capítulo. El estar asociados a un mismo concepto garantiza que existen rasgos semánticos compartidos. Por ejemplo, el concepto *ATENCIÓN Y VIGILANCIA* remite a verbos como *atender*, *controlar*, *cuidar*, *encargar(se)*, *guiar*, *inspeccionar*, entre otros, que son lemas en el diccionario. En tercer lugar, los predicados se pueden agrupar a partir de las clases léxicas con que se combinan: los predicados que seleccionen clases léxicas similares podrían constituir un grupo de predicados. Podemos explorar los predicados que seleccionan clases léxicas asociadas a un mismo concepto; por ejemplo, al concepto *POSIBILIDAD*, se asocian clases léxicas seleccionadas por *dilapidar*, *malgastar*, *malograr(se)*, que son verbos relacionados semánticamente. Bosque (2004b) ofrece un ejemplo de cómo se pueden agrupar los predicados según los otros predicados que los escogen: *leer* pertenece a las siguientes clases verbales, con las que comparte rasgos: verbos de lengua, como *pronunciar* (*leer en voz alta*), verbos de percepción, como *ver* (*leer entre líneas*) y verbos de consumición, como *engullir* (*leer vorazmente*). En estos casos, las propiedades selectivas de los adverbios son las que proporcionan clasificaciones múltiples de los verbos.

Una clase de verbos cuyos significados y combinatorias sean afines tendrá una representación conceptual muy clara en una base de conocimiento como FunGramKB: los verbos se vincularán al mismo evento conceptual y las clases léxicas a las mismas entidades

conceptuales. Ello será prueba de que la restricción de selección léxica puede ser una manifestación lingüística de una preferencia de selección entre un evento y una entidad conceptual.

3. Las clases léxicas

Las clases léxicas de REDES se obtienen mediante la siguiente metodología general: primero, se extraen del corpus todas las combinaciones en las que participa el predicado y, segundo, se establecen diferentes grupos de argumentos inductivamente según criterios semánticos. Las clases léxicas son, pues, grupos de argumentos que comparten ciertos rasgos semánticos en el contexto de una misma combinatoria con un mismo predicado. Esta metodología es bastante similar a la que utiliza Hanks (2013) para el inglés, que va encaminada hacia su *Pattern Dictionary of English Verbs*.

REDES constituye el primer intento de identificar estas clases léxicas a través de esta metodología, lo que plantea el problema de cómo delimitarlas. Según Bosque, “este problema resalta especialmente por el hecho de que el diccionario ha sido elaborado sin partir de una ontología, al menos explícita, de conceptos establecidos *a priori*” (2004b: CIV). De hecho, el diccionario se ha construido con la intención de ofrecer informaciones que permitan establecer vínculos parciales entre la intensión y la extensión de las palabras, como posible ayuda a la construcción de una ontología de conceptos. Sin embargo, hay que recordar que las numerosas clasificaciones, taxonomías u ontologías de términos y conceptos que existen – incluida la ontología de FunGramKB– han sido establecidas atendiendo a los rasgos que comparten los miembros de cada grupo en función de una tipología preexistente, no en función de la combinatoria sintáctica de las palabras que los designan, como es el caso de las clases léxicas de REDES.

REDES presenta dos tipos de clases léxicas, la primera de las cuales se llama “clase léxica abierta”. Las clases léxicas abiertas aparecen en la entradilla, o primera sección del artículo, y no están identificadas con una letra mayúscula. En el caso de *malgastar*, en la Tabla 2 al comienzo del capítulo, las clases léxicas abiertas son dos: “sustantivos que designan magnitudes, bienes materiales, recursos” (compuesta por: *dinero, agua, corcho, petróleo, gasolina, papel*) y “sustantivos que designan unidades temporales, incluidas las que se refieren a la existencia” (compuesta por: *vida, juventud, día*). Las clases léxicas abiertas incluyen paradigmas literales y designativos de argumentos, los cuales pueden no ser enteramente lingüísticos. Como dijimos antes, estas clases pueden contener información sobre el mundo, más que sobre el idioma. Son clases fácilmente ampliables y pueden ser

similares en diferentes lenguas. Estos rasgos le dan cierta estabilidad semántica a la clase léxica abierta, lo que sugiere que el vínculo con un concepto ontológico sería viable. Sin embargo, algunas clases léxicas abiertas son más fáciles de vincular con conceptos que otras; por ejemplo, la clase que designa ‘unidades temporales, incluidas las que se refieren a la existencia’ (*vida, juventud, día*) podría vincularse con un concepto relacionado con el ‘tiempo’, pero la que designa ‘magnitudes, bienes materiales, recursos’ (*dinero, agua, corcho, petróleo, gasolina, papel*) presentará un reto por lo heterogéneo de su intensión y extensión.

Las clases léxicas que se identifican con una letra mayúscula son las clases léxicas en propiedad, y tienen el mayor interés en el diccionario. Estas clases léxicas son conjuntos relativamente cerrados, porque sus miembros no se aprenden uno a uno, ni se pueden ampliar fácilmente. Las llamaremos clases léxicas “en propiedad” o “cerradas”, para distinguirlas de las clases abiertas, reconociendo que siempre hay la posibilidad de integrar nuevos miembros, aunque con frecuencia menor que en las abiertas. Bosque (2004b: C) caracteriza estas clases léxicas con los siguientes cuatro rasgos:

- (1) no se obtienen del conocimiento de la realidad, sino del análisis del idioma;
- (2) no se alargan indefinidamente, por lo que pueden ser descritas, restringidas y caracterizadas adecuadamente;
- (3) no se deducen directamente de la definición de la palabra, es decir, no proceden de la información denotativa o de la intensión de los conceptos;
- (4) deben ser aprendidas específicamente por el que adquiere el idioma como primera o segunda lengua.

Veamos algunos ejemplos de las clases léxicas cerradas. En la Tabla 2, el verbo *malgastar* se combina con seis clases léxicas cerradas, identificadas con las letras A, B, C, D, E y F. Cada una está descrita semánticamente por un descriptor, que caracteriza el significado compartido por todos los miembros de la clase. La clase E de *malgastar*, por ejemplo, tiene el siguiente descriptor: “sustantivos que designan unidades, materias y recursos informativos, expresivos o artísticos” y los argumentos miembros de la clase son: *saliva, línea, palabra, página, óleo, tinta, cinta*. Nótese que el sentido expresado por el descriptor puede reflejar al significado recto de una palabra dada (*palabra*), pero también un sentido figurado asumido por una palabra en esa combinación particular (*saliva*).

Para ilustrar el hecho de que las clases léxicas cerradas no se pueden deducir necesariamente de la definición del predicado, consideremos el caso de *nublar*. Si conocemos el significado de *nublar* en español, podemos predecir que *se nubla el cielo* –

sustantivo que figura una clase léxica abierta–, pero no podemos predecir que *se nubla un rostro* –sustantivo que figura en una clase léxica cerrada–.

Cabe señalar que las clases léxicas cerradas que se describen en REDES son casi siempre clases semánticas. Algunas de ellas coinciden, de hecho, con clases tradicionales, por ejemplo, los verbos tradicionalmente llamados de ‘percepción’ se agrupan de igual modo en REDES: *aguzar, apreciar(se), avistar, comprender, contemplar observar, escuchar, mirar, contemplar, fijarse, espiar, oír, escudriñar, ver, visionar*.

Las clases léxicas cerradas podrían ser particulares del español, porque proceden del análisis del idioma y no de la realidad, según señala Bosque en la caracterización de las clases (rasgo 1, arriba). Sin embargo, no se puede descartar que tengan un correlato conceptual cognitivo universal. Si las clases léxicas se pueden describir, restringir y caracterizar adecuadamente (rasgo 2, arriba), es probable que tengan la estabilidad necesaria para vincularse a un concepto, que es una entidad estable que está descrita, restringida y caracterizada en una ontología. Por lo mismo, tenemos que considerar que la restricción de selección léxica de un predicado por una clase léxica cerrada podría reflejar una preferencia de selección de concepto a concepto.

Al gestionar el vínculo de los datos de REDES a FunGramKB, podremos comprobar si las restricciones de selección léxica de REDES son manifestaciones de un fenómeno conceptual o si son fenómenos léxicos idiosincrásicos, o si, tal vez, se desdoblan entre lo conceptual y lo léxico. Si la restricción de selección léxica se puede validar en la ontología de conceptos, confirmaríamos que se trata de un fenómeno cognitivo universal y estableceríamos la expectativa de que se manifestara también en otras lenguas. En el ejemplo de la combinación *nublarse el cielo*, que vimos arriba, lo lógico es que haya una preferencia de selección conceptual entre los conceptos ‘nublar(se)’ y ‘cielo’, que se manifieste en cada lengua. En combinaciones como *nublarse el rostro*, podríamos pensar que esta es una selección única del español. Sin embargo, no es necesariamente así; un cotejo del verbo *to cloud* en el diccionario Merriam-Webster¹⁶ confirma que también ocurre en inglés: “*of facial features: to become troubled, apprehensive, or distressed in appearance <her face clouded with worry>*”. Dada esta evidencia bilingüe, esta restricción de selección léxica de REDES podría tener una representación conceptual en FunGramKB.

¹⁶ Versión electrónica, en <http://www.merriam-webster.com>

4. Los argumentos

Las clases léxicas se componen de argumentos con los que el predicado se combina sistemáticamente y que, en esa combinación, comparten un significado entre sí. Como hemos visto en la Tabla 2, el verbo *malgastar* se combina con dos clases léxicas abiertas y seis clases léxicas en propiedad. La clase B se describe como “sustantivos que designan facultades, aptitudes o cualidades humanas, especialmente las de carácter intelectual; también con otros que designan diversos estados de reconocimiento personal”; los argumentos que la componen son: *talento, capacidad, don, cualidad, generosidad, comicidad, valor, credibilidad, prestigio, experiencia*.

Los argumentos miembros de las clases léxicas presentan dos problemas cuando intentamos reducir las clases léxicas de REDES a un inventario estándar o a una ontología conceptual. En primer lugar, los argumentos pueden advenir a las clases léxicas no solo en virtud de su sentido recto, sino también a través de sentidos figurados. Esto refleja el hecho de que las palabras pueden adquirir un valor especial en la combinación. Se trata de lo que Pustejovsky (1995) llama “coerción”, Chomsky “recategorización”, Bosque (2004) “operación atributiva” y Hanks (2013) “explotación”. Por ejemplo, en REDES, *saltar(se)* se combina con palabras que designan obstáculos (*barrera, protección, impedimento*), pero también con otras, como *páginas, reunión y clase*, que se convierten en un tipo de obstáculo al combinarse con *saltar*. La diferencia fundamental estriba en que los rasgos intensionales del primer grupo “vienen con las palabras”, es decir, forman parte de su naturaleza léxica, mientras que los que caracterizan al segundo los asigna el hablante en un proceso de predicación implícita; a los primeros podemos llamarlos “paradigmas de base semántica”, mientras que a los segundos, podemos llamarlos “paradigma[s] de base pragmática o discursiva” (Bosque, 2004b: CXXXV). Por lo tanto, la pertenencia a una clase léxica no equivale a poseer inherentemente un significado, sino a la manifestación de ese significado en una pieza léxica en los contextos que los activan. Esto contrasta profundamente con la manera que se establecen los conceptos en ontologías como las de FunGramKB, como veremos en el capítulo III.

El segundo problema es que el comportamiento de los argumentos miembros de una clase léxica puede ser idiosincrásico o incluso arbitrario. Por ejemplo, con *atávico* se combinan sustantivos que denotan ‘sensación de miedo en diversos grados’, que incluyen

*miedo, terror, temor, pavor*¹⁷. Sin embargo, nótese que no aparecen otros sustantivos de ‘miedo’, que en principio deberían ser parte de la clase léxica, como *susto, horror, espanto*. Otros ejemplo lo tenemos en *exiguo*, que se combina con sustantivos que designan ‘las cantidades que se perciben como retribución o emolumento’, como *nómina, renta, sueldo, salario y pensión*, pero no *precio*. Por su parte, el verbo *estrechar* se combina con algunos sustantivos que denotan ‘ayuda’, como *cooperación y colaboración*, pero no con *apoyo*. Asimismo, podríamos pensar que los argumentos del verbo *leer* son sustantivos que designan textos, pero “no es evidente, sin embargo, que en el paradigma de los ‘sustantivos que designan textos en español’ hayan de aparecer los sustantivos *dato, disculpa y número*” (Bosque, 2004b: LXXXIV, nota al calce), que sí aparecen. Para manejar estos casos, tendremos que recurrir en FunGramKB no solo al módulo conceptual, que da cuenta de las regularidades del nivel semántico, sino también al módulo léxico, que recoge las idiosincrasias de cada lengua.

Veamos otro ejemplo de este problema con dos clases léxicas vinculadas al concepto TEMOR en el índice conceptual de REDES. Bajo TEMOR se incluye la clase léxica que se combina con *atávico* que vimos arriba, descrita como “sustantivos que denotan sensación de miedo en diversos grados” y compuesta por *miedo, terror, temor, pavor*. También se incluye una clase léxica que se combina con *arraigado*, descrita como “sustantivos que designan sentimientos provocados por la incertidumbre” y compuesta por *preocupación, temor, miedo*. A pesar de estar vinculadas al mismo concepto en el índice conceptual de REDES, cada clase tiene rasgos particulares, que se reflejan tanto en los argumentos miembros como en los descriptores. Esto nos da un indicio de las particularidades de la combinatoria léxica de cada predicado y del reto que ello supone para la reducción de las clases léxicas en su vínculo con los conceptos ontológicos.

Bosque (2006) argumenta que la pertenencia a un grupo conceptual de una palabra y su combinatoria no tienen por qué coincidir: por ejemplo, la *guitarra* y el *piano* son ‘instrumentos de cuerda’, y aunque se puede *tocar* o *afinar* cualquiera de los dos, uno *rasguea* la primera y *se sienta* al segundo, no al contrario. Ninguna tipología existente de los sustantivos que designan instrumentos musicales tiene en cuenta estas diferencias, que Bosque llama “problemas de dispersión de una entrada genérica” (2006: XXXI). Sin embargo, es precisamente esa información la que resulta esencial para un diccionario combinatorio. Si queremos representar en una base de conocimiento la información

¹⁷ Elaboramos los ejemplos de esta sección a partir de las referencias que da Bosque (2004b: LXX).

contenida en las clases léxicas de REDES, necesitaremos los mecanismos para recoger tanto lo que es semánticamente regular, como lo que es semánticamente impredecible.

5. Los índices conceptuales

El diccionario REDES incluye dos índices de conceptos, cuyos miembros están integrados alfabéticamente al leuario del diccionario y se destacan con diferente tipografía. El “índice conceptual” se vincula con los lemas de REDES y se compone de conceptos en letra versalita cursiva, por ejemplo, *ALIMENTO*. El “índice de clases léxicas” se vincula, justamente, con las clases léxicas de REDES y se compone de conceptos en letra versalita regular, por ejemplo, ALIMENTO.

Los conceptos que pueblan el índice conceptual son “creados inductivamente a partir de las entradas del leuario general de REDES” (Bosque 2012, comunicación personal). Se trata, pues, de un inventario de conceptos que provee las nociones que permiten clasificar el leuario de predicados de REDES, para beneficio del lector. Aunque los conceptos no se definen, organizan o relacionan entre sí, resultan útiles como pistas para vincular los predicados a la ontología de FunGramKB¹⁸. Para ilustrar el índice conceptual de REDES, reproducimos a continuación la entrada del concepto *ALIMENTO*:

Tabla 4. El concepto *ALIMENTO* en REDES

| |
|--|
| <i>ALIMENTO</i> Véase: |
| ◆ <i>frugal, pantagruélico, suculento, voraz</i> |
| ◆ <i>alimento, ayuno, bebida, comida, fruta, manjar, sustento, zumo</i> |
| ◆ <i>beber, comer, desayunar, devorar, ingerir, mamar, merendar, pacer</i> |

La entrada del concepto *ALIMENTO* expresa que los predicados incluidos, que son lemas del diccionario, se pueden asociar al concepto *ALIMENTO*. Los lemas se organizan por categoría gramatical: se asocian al concepto *ALIMENTO* los adjetivos *frugal, pantagruélico, suculento, voraz*; los sustantivos *alimento, ayuno, bebida, comida, fruta, manjar, sustento, zumo*; y los verbos *beber, comer, desayunar, devorar, ingerir, mamar, merendar, pacer*.

¹⁸ Por otro lado, tanto el índice conceptual como el índice de clases léxicas de REDES resultan un poco difíciles de consultar. La referencia cruzada solo se da de concepto a predicados y clases léxicas, pero no en la dirección contraria. Esto quiere decir que para conocer con qué concepto se vincula una clase léxica, el lector tiene que abstraer diferentes posibles conceptos a partir de la información de la clase, y luego buscarlos en el diccionario, para confirmar si se incluyen o no. Si aparece el concepto, entonces hay que cotejar si incluye la clase léxica en particular.

Otra entrada del índice conceptual es *PERCEPCIÓN*. Bajo *PERCEPCIÓN*, figuran los verbos *aguzar, apreciar(se), avistar, comprender, contemplar*; los adjetivos *contemplativo, escrutador, inescrutable*; los sustantivos: *apreciación, aroma, atención, contemplación*; y los adverbios *apreciablemente, de vista, fijamente*. Todas las palabras identificadas bajo *PERCEPCIÓN* se consideran verbos, adjetivos y sustantivos de ‘percepción’ en REDES, los cuales, como señalamos antes, podrían coincidir con clases léxico semánticas tradicionales de ‘percepción’. Todos estos predicados son lemas en el diccionario.

El “índice de clases léxicas” o “referencias cruzadas a los conceptos” constituye un segundo índice conceptual de REDES. Estos conceptos se registran en versalita, por ejemplo, *PERCEPCIÓN*, y remiten a clases léxicas particulares, como en el ejemplo siguiente. Bajo *PERCEPCIÓN* figuran, entre los adjetivos: *abigarrado^B, abismal^H, acusado^J*, y entre los adverbios: *al vuelo^A, atentamente^A, a toda pastilla^B*. En el caso de *atentamente*, la marcación se debe interpretar como sigue: la clase léxica A seleccionada por *atentamente* es una clase léxica de verbos de ‘percepción’. En efecto, la clase léxica A de *atentamente* está compuesta de “verbos que designan la aplicación de la vista, el oído o el entendimiento a algún asunto” e incluye los miembros *observar, escuchar, mirar, contemplar, fijarse, espiar, oír, escudriñar, ver, visionar*.

Veamos como segundo ejemplo del índice de clases léxicas el concepto *ALIMENTO*, cuya entrada reproducimos en la siguiente tabla. La entrada del concepto *ALIMENTO* comunica que la clase léxica B de *casero* y la clase léxica C de *celestial* se asocian al concepto *ALIMENTO*:

Tabla 5. El concepto *ALIMENTO* en REDES

| |
|---|
| ALIMENTO |
| ◆ (SUSTANTIVOS) Véase: <i>casero^B, celestial^C, compulsivo^B, saborear^A</i> |
| ◆ (VERBOS) Véase: <i>como un cosaco^A, debidamente^J, en frío^B, vorazmente^A</i> |
| Véase también: <i>INGESTIÓN</i> |

Cuando vamos a la entrada del predicado *casero*, encontramos que la clase léxica B (con que se combina *casero*) se describe como “el sustantivo *comida* y [...] otros que designan algunas de las formas en que se presenta o los lugares en que se sirve o prepara” y se compone de: *comida, restaurante, receta, plato, guiso, postre, licor*. Por su parte, bajo *celestial*, encontramos la clase léxica C, que se describe como “sustantivos que designan comestibles o viandas, frecuentemente de agradable degustación” y se compone de: *comida, manjar*. Al

comparar ambas clases, podemos apreciar un centro semántico común relacionado con la noción de ALIMENTO, pero también las particularidades de cada clase léxica. Como hemos dicho ya, estas particularidades combinatorias constituirán un reto al intentar vincular las clases léxicas a los conceptos de FunGramKB.

Veamos un ejemplo final de la utilidad del índice de clases léxicas. Hemos creado una tabla que muestra los conceptos que se asignan a las clases léxicas que se combinan con *malgastar*, las cuales vimos en detalle la Tabla 2. Los conceptos de REDES no están pensados para proveer información combinatoria (Bosque 2012, comunicación personal); sin embargo, pueden dar una visión panorámica conceptual de la combinatoria de los predicados, en este caso, *malgastar*.

Tabla 6. Vínculos de las clases léxicas de *malgastar* al índice conceptual de REDES

| Clase léxica | Concepto |
|---|----------------------|
| <i>dinero, agua, corcho, petróleo, gasolina, papel</i> | RECURSO |
| <i>vida, juventud, día</i> | TIEMPO ¹⁹ |
| <i>A. fuerza, energía, esfuerzo, salud</i> | FUERZA |
| <i>B. talento, capacidad, don, cualidad, generosidad, valor, credibilidad, prestigio, experiencia</i> | CAPACIDAD, FACULTAD |
| <i>C. oportunidad, ocasión, balón, ventaja, punto, opción, privilegio, posibilidad</i> | POSIBILIDAD |
| <i>D. éxito, victoria, triunfo</i> | ÉXITO |
| <i>E. saliva, línea, palabra, página, óleo, tinta, cinta</i> | INFORMACIÓN |
| <i>F. disparo, munición, cartuchera, bala</i> | ARMA |

El asociar las clases léxicas a un índice conceptual puede verse como un paso hacia la reducción del inventario de clases. Bosque (2006) señala, sin embargo, que en sus intentos de hacerlo se ha topado con numerosos problemas de dispersión como el de *piano y guitarra*, citado arriba, que ilustra que aunque las palabras sean parte de una clase semántica, su combinatoria léxica no es siempre igual. En PRÁCTICO, publicado dos años después que REDES, los índices conceptuales se convierten en “entradas genéricas”. Se trata de hiperónimos de las nociones expresadas por las clases léxicas, y suman solo 18 en todo el diccionario²⁰. Según Bosque (2012, comunicación personal), las entradas genéricas de PRÁCTICO, a diferencia de las de REDES, sí tienen interés combinatorio, pues se

¹⁹ A las clases léxicas abiertas, REDES no les asigna un concepto explícitamente, por lo que la autora ha asignado estos conceptos (RECURSO, TIEMPO) a partir del Índice conceptual.

²⁰ Las entradas genéricas de PRÁCTICO son las siguientes: ÁRBOL, BEBIDA, CALZADO, COLOR, CREYENTE, DEPORTE, DÍA, DISCIPLINA, DROGA, ESTABLECIMIENTO, GOLPE, IDIOMA, INSTRUMENTO MUSICAL, MES, MONEDA, RELIGIÓN, ROPA, TÍTULO NOBILIARIO.

constituyeron con ese requisito: “las entradas genéricas de PRÁCTICO solo tienen sentido si las combinaciones de cada una de ellas son idénticas”. Hay que recordar además que para que un concepto verdaderamente abarque la clase léxica existe otra consideración: “lo fundamental es determinar cuáles de ellas se caracterizan por extender las relaciones combinatorias a los hipónimos que contienen” (Bosque, 2006: XXXI). En una ontología taxonómica como la de FunGramKB, los conceptos mantienen relaciones de subsunción, lo que significa que los conceptos subordinados heredan las características de sus superordinados. Por ello, en principio, si un predicado selecciona un concepto, selecciona también a todos los subordinados del concepto.

Como comentamos arriba, las 18 “entradas genéricas” de PRÁCTICO –que asemejan un pequeño índice conceptual– participan en las mismas combinaciones por el hecho de pertenecer a un mismo campo semántico. Si el factor determinante para una combinación dada es que el sustantivo designe un color, por ejemplo, es lógico centrar el análisis en determinar las combinaciones que ocurren a partir del concepto ‘color’, no de las unidades léxicas individuales que designen ‘color’ (*rojo, verde, azul*), aunque esas sean las lexicalizaciones correspondientes. A la luz de esto, para poder reducir las clases léxicas de REDES a un inventario conceptual habría que comprobar empíricamente las propiedades sintácticas de los miembros de esas clases y de las clases subordinadas. Bosque (2006) señala que la reducción conceptual de las clases léxicas puede provocar predicciones demasiado fuertes, creadas por simplificaciones excesivas. Lo comprueba en casos sencillos: por ejemplo, un ARMA se puede *disparar* (porque *se disparan las pistolas*, etc.), pero *cuchillo* es también un ARMA y no se dispara. El simplificar las clases léxicas requiere, por lo tanto, ponerlas a prueba para evitar que generen predicciones incorrectas.

Una de nuestras tareas de investigación ha sido reducir el inventario de clases léxicas vinculándolas a la ontología conceptual de FunGramKB, sin perder información sobre los argumentos particulares, que, como hemos visto, participan en las clases léxicas por múltiples razones. Queremos intentar dar cuenta de todos los argumentos de las clases léxicas –o de la mayor cantidad posible– al traducir la información de REDES a FunGramKB, tanto los que puedan vincularse con una entidad conceptual como los que no. Para ello nos tendremos que valer tanto del módulo conceptual como del módulo léxico en FunGramKB, que tiene un apartado para colocaciones léxicas, como veremos en el capítulo III. Sin embargo, nuestro trabajo solo constituye un primer paso hacia la reducción definitiva, pues, como hemos visto, esa meta requeriría el cotejo del comportamiento de la clase léxica en otros contextos, y del comportamiento de sus clases léxicas subordinadas.

6. Las restricciones de selección léxica

REDES es un diccionario de restricciones léxicas –o, mejor, restricciones léxico semánticas–, que intenta acotar o delimitar las restricciones significativas que los predicados imponen a sus argumentos a través de la selección léxica (Bosque, 2004b: XCVI), ya sean de naturaleza aspectual, nocional o de otro tipo. La información sobre combinatoria léxica de REDES no se obtiene analizando el mundo y considerando las propiedades objetivas de los objetos, sino examinando la estructura del idioma y las clases lingüísticas que surgen de él. Los rasgos que más interesan en REDES son, pues, los datos lingüísticos, que están sujetos a variación histórica y dialectal y que no necesariamente coinciden con los de otras lenguas. De hecho, los componentes de la definición que la sintaxis no realiza de forma manifiesta quedan fuera del concepto de restricción léxica que maneja REDES (Bosque, 2004). Las clases léxicas de REDES responden, por lo tanto, a un acercamiento inductivo, que se mueve de abajo (uso lingüístico) a arriba (organización por criterios semánticos y conceptuales).

Las restricciones de selección léxica de REDES se ubican en dos tradiciones teórico prácticas diferentes. En el ámbito teórico, las restricciones de selección de REDES son una respuesta a las restricciones selectivas generativistas y las restricciones aspectuales del *Aktionsart*. En el ámbito práctico, REDES responde a los diccionarios de colocaciones. Sin embargo, como veremos, las pautas que establece REDES lo distinguen bastante de esos precedentes.

6.1 Modelos teóricos

Las restricciones selectivas, también llamadas “selección semántica” o “selección-s” (frente a la “selección categorial” o “selección-c”), fueron introducidas en la gramática generativista por Grimshaw (1979), Pesetsky (1982) y Chomsky (1986), entre otros. Estas restricciones de selección incluían rasgos como ‘humano’, ‘animal’, ‘concreto’, ‘abstracto’, ‘objeto material’, ‘objeto líquido’, ‘artefacto’, ‘instrumento’, etc. Según Bosque (2004b), las restricciones selectivas no fueron capaces de sentar las bases semánticas de la combinatoria gramatical, porque en muchos casos solo aportaban información extralingüística. Por ejemplo, el hecho de que *cantar* se predique de los humanos, los pájaros y tal vez las ballenas, no es parte del significado de *cantar*, sino de nuestro conocimiento de las entidades que pueden cantar en el mundo. Por ello, las restricciones selectivas tradicionales no aportaron lo suficiente a formular propiedades lingüísticas.

Aunque las restricciones léxicas que verdaderamente interesan en REDES son las que no se derivan de nuestro conocimiento del mundo, podemos encontrar en el diccionario algunas restricciones léxicas de naturaleza extralingüística, como hemos señalado antes. Estas figuran frecuentemente en la primera parte del artículo o entradilla, entre las clases léxicas abiertas, que contienen paradigmas léxicos fácilmente ampliables. Por ejemplo, *alimentarse (de)* se combina con la clase léxica abierta de ‘comida, bebida y –en general– todo lo que pueda ser ingerido’, que incluye *pan, agua, leche, carne, chocolate*. Estas combinaciones responden a nuestro conocimiento sobre lo que se puede comer, no a nuestro conocimiento del idioma. Por otro lado, cabe señalar que no todas las clases léxicas abiertas reflejan conocimiento extralingüístico: en la entradilla de *alimentarse (de)*, figuran también clases léxicas que recogen ‘expresiones artísticas de muy diversa naturaleza’ (*versos, pintura, cine*) y ‘unidades de información’ (*datos, noticias*), que reflejan nuestro conocimiento del idioma.

A diferencia de las restricciones selectivas, la información aspectual que presenta el *Aksionsart* sí es exclusivamente lingüística. Las restricciones aspectuales del *Aksionsart*, específicamente las clases léxicas de eventos de Vendler (1967) –estados, actividades o procesos, realizaciones y logros–, han contribuido indudablemente a un mejor entendimiento de la relación léxico sintaxis. Hay selección léxica en verbos como *avisar* y *narrar*, que admiten como únicos complementos nombres de eventos, por ejemplo, *accidente* o *partido*: *Me avisaron del accidente de María* o *Narró muy bien el partido del equipo*. No podríamos decir **Me avisaron de María* o **Narró muy bien el equipo* (ejemplos de Bosque, 2004b: CIII).

El aspecto léxico, sin embargo, puede dar cuenta de la selección léxica “solo en parte”, según Bosque (2004b: CIII). Aunque aporta algunos rasgos relevantes, no necesariamente *explica* la selección. Bosque señala que el adverbio *brevemente* se combina con verbos de actividad o estado (*conversar brevemente*), lo que ilustra un caso de selección léxica de naturaleza aspectual; sin embargo, no es lo mismo decir que un adverbio dado se combina con verbos de actividad, que afirmar que cualquier verbo de actividad admitirá ese adverbio, o que este adverbio no se podrá combinar con verbos de otros tipos (2004b: CVI). Según el autor de REDES, al igual que en el caso de las selecciones restrictivas, tal vez el desdoblamiento de los rasgos aspectuales podría lograr un mayor alcance descriptivo²¹.

²¹ El cotejo de Croft (2012) quedó fuera del alcance de esta investigación, pero su identificación de nuevos tipos y subtipos de aspectos léxicos podría arrojar luz sobre las motivaciones aspectuales de la combinatoria sistemática.

Las limitaciones de las restricciones selectivas y del *Aksionsart* no implican que la relación entre predicado y argumento no se pueda restringir mediante informaciones semánticas que constituyen parte de nuestro conocimiento del idioma (Bosque, 2004b: C). Describir estas relaciones es justamente la motivación de REDES, especialmente cuando la información combinatoria se vale de procesos que van más allá de los sentidos rectos de los predicados y argumentos. Por ejemplo, el hecho de que un *pájaro planee* es predecible por nuestro conocimiento extralingüístico. Eso está recogido en la entradilla de *planear*, que especifica que “en su sentido físico se combina con sustantivos que designan seres voladores (*aeroplano, avión, gaviota, águila*)”. Lo que interesa es lo que no es predecible, por ejemplo, que en español una *duda planee*. De hecho, *duda* pertenece en REDES a la clase de sustantivos que denotan ‘duda, falta de certeza y otras nociones caracterizadas por la ausencia de información, de seguridad o de confianza’ (*incertidumbre, duda, interrogante, misterio, miedo, temor, incógnita, sospecha, pregunta, confusión*).

6.2 Diccionarios de colocaciones

El hecho de que REDES sea un diccionario combinatorio lo coloca de inmediato en la tradición de los diccionarios de colocaciones, que se han encargado de tratar lexicográficamente aspectos de la combinatoria de las palabras²². Bosque (2001c) plantea diversos problemas con la noción misma de ‘colocación’, que considera mal delimitada, en gran medida porque considera que el fenómeno en sí está mal concebido. Este autor defiende que las combinaciones de palabras tradicionalmente llamadas “colocaciones” no son unidades aisladas, sino manifestaciones del fenómeno más amplio de restricción de selección léxica. La relación entre las dos palabras en una colocación raramente es biunívoca, señala Bosque (2001c). Por ejemplo, la colocación *enemigo acérrimo* es realmente parte de la restricción léxica de *acérrimo*, que selecciona no solo a *enemigo*, sino también a *adversario, detractor, rival, crítico, opositor, etc*²³. Las restricciones sistemáticas de selección léxica pertenecen, pues, a la interfaz léxico sintaxis, y no a la fraseología, como ha sido alegado tradicionalmente.

Bosque (2004b) establece varios parámetros para caracterizar los diccionarios de colocaciones existentes. En primer lugar, distingue dos grandes corrientes: una basada en

²² Para la bibliografía sobre colocaciones de 2000 al 2009, ver Koike (2009): <http://www.ner.takushoku-u.ac.jp/~kkoike/Colocaciones2000.pdf>. Para un repaso de los trabajos sobre colocaciones en español, ver el volumen 23 (2001) de *Lingüística Española Actual*. Para un panorama general hasta el 2000, ver Koike (2000).

²³ Esta es solo una de las clases léxicas con que se combina *acérrimo* en REDES.

frecuencia de coaparición de dos unidades léxicas (BBI, LTP, OCD, CCEC²⁴), y otra en que las colocaciones se entienden como la manifestación léxica de relaciones semánticas llamadas “funciones léxicas” (DEC, DICE, LAF). Ante el primer acercamiento, basado en una alta frecuencia de coaparición, Bosque (2001c) responde que el constatar que ciertas palabras se realizan con frecuencia junto a otras palabras particulares no necesariamente representa información lingüísticamente relevante ni ayuda a comprender ninguna característica fundamental del idioma. Para sustentar esto, cita algunos ejemplos del CCEC, que usa un programa informático de extracción de coapariciones frecuentes. Un caso es el del adverbio *rarely* (‘infrecuentemente’), que se combina con los verbos *see, use, find, work, get*; estos datos no dicen mucho sobre *rarely*, un adverbio de frecuencia que podría combinarse con un sinnúmero de verbos, sin mucha restricción, y tampoco sobre el grupo de verbos que selecciona: el hecho de que este grupo de verbos designe acciones que se pueden realizar ‘infrecuentemente’ no lo restringe lo suficiente para señalar alguna propiedad lingüística particular.

Los diccionarios de colocaciones se pueden clasificar además en términos de la teoría que los subyace: algunos se basan en el “contextualismo británico” impulsado por Firth²⁵ (BBI, LTP, OCD, CCEC), mientras que otros están vinculados a la Teoría Sentido Texto (TST) desarrollada por I. Mel’čuk (DEC, DICE, LAF). En cuanto a las lenguas que tratan, los diccionarios de colocaciones en inglés incluyen el CCEC, BBI, OCD, LTP, los del francés el DEC, LAF, y los del español el DiCE. Un criterio final es el de la direccionalidad: todos estos diccionarios abordan las combinaciones de argumento a predicado (A>P).

A la luz de estos criterios, el diccionario REDES se presenta como un diccionario interesado en las relaciones semánticas de la combinación, enmarcado en varias teorías, que trata el idioma español y que organiza los datos de predicado a argumento (P>A). En este último aspecto, REDES se distingue de todos los otros diccionarios.

Conclusión

La riqueza de información que ofrece cada entrada analítica de REDES es evidente, al igual que la riqueza de las interrelaciones que se establecen a través de las referencias cruzadas y los vínculos conceptuales. Es difícil explotar cabalmente estas informaciones e interrelaciones en el formato lineal del libro, aunque estén codificadas a través de las

²⁴ Ver Siglario, al principio de la tesis.

²⁵ A quien se atribuye acuñar el término “colocación” en los años 1930.

múltiples referencias internas que mencionamos. REDES opera a partir de la noción de que, en términos lexicográficos, la intensión de los predicados se puede reducir, en la medida de que la extensión se amplíe. Esto coincide con las pautas fundamentales de las bases de conocimiento como FunGramKB. Se trata pues de un contenido ideal para ser traducido a una plataforma electrónica, que permitiría saltar con el toque de un botón de un predicado léxico a los conceptos que codifican sus sentidos, o de un concepto (evento) a todos los predicados que lo lexicalizan. Se podría también mostrar la combinatoria de los predicados en términos de conceptos (entidades), e ir del concepto a sus lexicalizaciones o bien a los demás predicados o conceptos que lo seleccionan. Todo ello requiere, sin embargo, una base de conocimiento de cierta complejidad que permita vincular la información léxico semántica de REDES a un eje conceptual común que facilite visualizar y explotar todas las relaciones.

Bosque reconoce el potencial de la reducción conceptual de las clases léxicas, al sugerir que REDES se conecte con una ontología de subdominios léxicos, un diccionario ideológico o una base de datos léxicos como WordNet o FrameNet u otros proyectos de procesamiento del lenguaje natural (2004b: CXXXV). Por ello, una base de conocimiento léxico conceptual como FunGramKB, que busca representar e incluso entender nuestro conocimiento del mundo y del idioma, y que incluye entre sus lenguas el español, se presenta como un marco ideal para ensayar la traducción conceptual de REDES. Para FunGramKB, será valioso contar con la información de REDES, que refleja conocimiento lingüístico humano que no es predecible de las definiciones de diccionarios y que supone un gran trabajo de recogida y análisis. Para efectos de la WSD en el PLN, la vinculación conceptual de la relación entre significado y combinatoria podría ser en extremo útil.

El gran reto radica en poder dar cuenta de los fenómenos de selección léxica aparentemente idiosincrásicos que presenta REDES en una base de conocimiento como FunGramKB, con un inventario de conceptos con definiciones fijas; para ello, recurriremos no solo a la parte conceptual de FunGramKB, sino a su módulo léxico.

CAPÍTULO II. LA LEXICOGRAFÍA INFORMATIZADA DEL SIGLO XXI

Introducción

El elegir la base de conocimiento FunGramKB como marco teórico y práctico de esta tesis enmarca este trabajo en el ámbito amplio de la lexicografía informatizada del siglo XXI o del tercer milenio. Esta nueva lexicografía –que puede estar al servicio tanto de la consulta humana como del PLN– incluye recursos lexicográficos informatizados de una complejidad y profundidad crecientes, mucho más ambiciosos que las conocidas versiones digitales de los diccionarios tradicionales. Se trata de bases de datos léxicas o bases de conocimiento léxico, que se conciben y construyen en plataformas electrónicas, y cuyo diseño particular –junto con el tipo de datos que lo pueble– permitirá la explotación y aplicación de la información de maneras determinadas.

Estamos ante un dinámico campo donde la informática emplea al máximo sus métodos y herramientas para respaldar una multiplicidad de tareas lexicográficas, que incluyen la recopilación de información lexicográfica proveniente de diferentes fuentes, el uso de las bases de datos para almacenar la información lingüística relevante, la creación y edición de entradas lexicográficas y la disseminación de productos lexicográficos (Boas, 2009c). Cabe señalar que una de las principales consecuencias del progreso de los recursos lexicográficos electrónicos ha sido la mejoría de los sistemas prácticos de PLN.

Las diversas manifestaciones de la lexicografía informatizada comparten el rasgo de querer vincular múltiples tipos de información en múltiples dimensiones. Estos nuevos recursos van más allá de ofrecer una definición de la palabra, y pueden incluir relaciones sintácticas y semánticas del lema con otras palabras, información sobre valencias sintácticas y semánticas, vínculos con ejemplos del uso real en corpus, equivalencias en otras lenguas y vínculos con ejes conceptuales. Este tipo de estructura multidimensional es justamente lo que se necesita para dar cabida y explotar el entramado de informaciones diferentes que establece REDES.

Nos interesa el hecho de que al enfrentar la construcción de los lexicones de gran escala, especialmente los lexicones multilingües, muchos investigadores llegan a la conclusión de que los conceptos –y no las palabras– deben ser representados como unidades prioritarias (Zaenen, 2002). Como veremos en este capítulo, varias bases de datos léxicas importantes utilizan índices conceptuales universales en función de interlinguas. Estos índices tienen usualmente la forma de taxonomías u ontologías semánticas o conceptuales.

Vossen y Fellbaum (2009) señalan, en referencia a la construcción del Global Word Net Grid (GWG), que estos índices permitirán la investigación empírica de gran escala sobre preguntas teóricas fundamentales, como, por ejemplo, cuáles unidades léxicas reflejan conceptos universales y cuáles son idiosincrásicas de una lengua. Este es el caso de FunGramKB, cuyos creadores afirman que el acercamiento conceptual tiene mayor capacidad explicativa que los modelos lingüísticos para determinar el complejo proceso de la construcción del significado (Mairal *et al.*, 2011).

El tratamiento que se dé al significado en las taxonomías u ontologías, por otro lado, puede responder a una semántica superficial o profunda (Velardi *et al.*, 1991). Mairal y Perrián (2010) señalan como ejemplo de semántica superficial el acercamiento “relacional” al significado que se utiliza en WordNet y EuroWordnet, en donde se representa el significado de una palabra a través de su relación (sinonimia, meronimia, etc.) con el significado de otra palabra. Por su parte, los acercamientos de semántica profunda, como el de FunGramKB, definen formalmente cada concepto. Como veremos, los acercamientos superficiales son más fáciles de construir y de constatar objetivamente que los acercamientos profundos, pero los acercamientos profundos establecen modelos más robustos y de mayor alcance semántico, por ejemplo, al ser capaces de hacer inferencias.

Tanto los acercamientos de semántica superficial como los de semántica profunda se pueden considerar enfoques “simbólicos”, porque explícitamente desarrollan reglas, principios y axiomas para dar cuenta del conocimiento lingüístico. Frente a estos enfoques simbólicos desarrollados por lingüistas, se presentan los enfoques estadísticos, que utilizan algoritmos o métodos probabilísticos para representar el conocimiento lingüístico sin necesidad de postular teorías lingüísticas. Curiosamente, los acercamientos estadísticos son los más exitosos actualmente. Hanks señala:

In recent years, ‘knowledge-poor’ statistical methods in computational linguistics have achieved remarkably—some would say astonishingly—good results, at a coarse-grained level, in applications such as machine translation, message understanding, information retrieval, and idiomatic text generation. At the same time, refined methods based on syntactic and valency theory have yielded largely disappointing results. The same is true of methods based on using machine-readable versions of dictionaries that were designed for human beings. (Hanks, 2013, §13.3.1²⁶)

Sin embargo, según este mismo autor (Hanks, 2013), los métodos estadísticos tienen un límite. A la vez, señala que los modelos determinísticos necesitan reevaluar la relación que

²⁶ Esta cita es tomada del manuscrito final de este libro, no de la versión publicada, por lo que no contamos con la página exacta.

postulan entre léxico y sintaxis, para mejorar su eficacia. Mairal (2013) añade que, independientemente del acercamiento (ya sea a través de algoritmos o de conocimiento), el objetivo central de la investigación es crear sistemas de etiquetado semántico que sean comprensibles por la máquina; por ello, “en principio, la apuesta por incorporar un modelo como el MLC [Modelo Léxico Construccional] en entornos de PLN parece una iniciativa metodológicamente plausible” (Mairal, 2013: 3).

En este capítulo presentaremos un resumen del panorama actual de la lexicografía informatizada, lo que nos proveerá el contexto necesario para valorar a FunGramKB en el próximo capítulo. En la sección 1 comentaremos la relación entre informática y lexicografía, destacando los tres tipos de entidades que se distinguen en ella. En la sección 2 presentaremos ejemplos de diccionarios para consulta humana en versión electrónica, en inglés y español, con particular énfasis en los que rediseñan sus estructuras para organizar su contenido en torno a un eje conceptual o semántico. En la sección 3, discutiremos algunos de las principales bases de datos léxicas existentes, que sirven tanto la consulta humana como el PLN: WordNet, EuroWordNet y FrameNet. En la sección 4, nos detendremos en las ontologías, por el importante rol que asumen en muchas bases de conocimiento léxico. A lo largo del capítulo, según comentamos los recursos lexicográficos, pondremos atención al manejo que da cada uno a las restricciones o preferencias de selección (ya sea en forma de colocaciones o estructura argumental). Terminaremos el capítulo con una conclusión.

1. Relación entre lexicografía e informática

La lexicografía está vinculada a la informática desde fines de la década de los 1940, cuando la traducción automática se convierte en una meta computacional. El gran problema de la traducción automática era la WSD, lo que dirigió a los investigadores a mirar los diccionarios existentes como referencias que proveyeran los sentidos de las palabras. Ello supuso transferir el contenido de los diccionarios a la plataforma electrónica por primera vez. Aunque el interés por la traducción automática estableció un vínculo temprano entre computadoras y diccionarios, la meta de la WSD para el PLN todavía no ha sido alcanzada cabalmente, como veremos en el capítulo IV (en la discusión sobre el manejo de la polisemia en la lexicografía informatizada).

Sintetizando lo que ha sido la relación entre computadoras y diccionarios hasta el momento actual, Pajzs (2009: 327) identifica tres tipos de entidades²⁷:

- (1) los diccionarios impresos, que son hoy preparados en bases de datos electrónicas pero que retienen la mayoría de sus características tradicionales;
- (2) las versiones electrónicas de dichos diccionarios, que tienen información y posibilidades de búsqueda adicionales; y
- (3) las bases de datos lexicográficas multifuncionales, que se diseñan en y para la plataforma electrónica, usando el más avanzado conocimiento sobre tecnologías del lenguaje y lexicografía computacional.

El grupo (1) se empieza a constituir con la introducción de las computadoras en el mundo profesional. Desde que se comercializan, las computadoras se van haciendo imprescindibles en todo trabajo que requiere el procesamiento de datos. El trabajo actual de las editoriales e instituciones encargadas de diseñar, redactar y publicar diccionarios es impensable sin los recursos electrónicos disponibles hoy en día, incluidas las bases de datos y los grandes corpus lingüísticos. El resultado son diccionarios impresos que tienden a mantener las características lexicográficas tradicionales, aunque hayan sido preparados electrónicamente.

El grupo (2) comienza con las transferencias del contenido de los diccionarios impresos a formatos electrónicos, inicialmente como herramienta para la traducción automática. Esta fase inicia a finales de la década de los 1960 con la transliteración manual del *Webster's Seventh New Collegiate Dictionary* a cintas magnéticas legibles por máquinas, lo que resultó en el primer diccionario procesable computacionalmente (*Machine Readable Dictionary*, en adelante MRD). Luego se les fue dando un formato estructurado a estos contenidos: el *Longman Dictionary of Contemporary English* o LDOCE (Procter, 1978), por ejemplo, se convierte en la fuente principal de investigación y desarrollo en este campo en la década de 1970. Poco tiempo después, se crean las primeras bases de datos léxicas en propiedad, que permiten almacenar y manipular la información. Los investigadores comienzan a usar esas bases de datos para poblar los componentes léxicos de sus sistemas de PLN. De esas bases de datos, surgen, en la década de 1980, los primeros diccionarios electrónicos para consumo del público en general, presentados en discos o CD-ROM,

²⁷ Estos tres grupos reflejan las tres fases en la relación entre diccionarios y computadoras identificadas por Cerquiglini (citado por Pruvost, 2000: 188): (1) la lexicografía tradicional asistida por computadoras; (2) la transferencia de los diccionarios tradicionales al medio electrónico; y (3) los diccionarios informatizados, concebidos desde un inicio en el medio electrónico.

dispositivos manuales o Internet. Aquí se incluyen la mayoría de los diccionarios digitalizados que consultamos en Internet.

El grupo (3) inicia con la construcción de los primeros diccionarios concebidos y diseñados para el formato electrónico. Se trata usualmente de bases de datos léxicas o bases de conocimiento léxico de cierta complejidad, la mayoría de las cuales están pensadas para el PLN. Aquí se conjugan la lingüística y la informática para maximizar las posibilidades de interacción entre el lenguaje humano y las computadoras. Esta tercera etapa de la lexicografía incluye los recursos más modernos existentes hoy –en toda su diversidad– y también constituye la puerta al futuro de la lexicografía. En el grupo (3) se incluyen los lexicones electrónicos como WordNet, EuroWordNet, FrameNet, que discutiremos en este capítulo, y FunGramKB, en el próximo.

Los diccionarios electrónicos pensados principalmente para consumo humano (grupo 2) y los lexicones electrónicos pensados principalmente para el PLN (grupo 3) presentan diferencias significativas. Las estructuras en que se almacenan y los formalismos con que se codifican los datos en los lexicones para el PLN los hacen muy diferentes a los diccionarios electrónicos para consulta humana. No obstante esto, de Schryver (2003) señala que existe una tendencia de incluir más componentes del PLN en los diccionarios electrónicos, y más posibilidades de consulta humana en algunos lexicones del PLN –como WordNet o FrameNet–. FunGramKB, por ejemplo, cumple con el doble objetivo de ser usado por computadoras y personas:

our resource (i) has been primarily developed to be reused in various NLP tasks—e.g. information retrieval and extraction, machine translation, or dialogue-based systems, and (ii) can also be used by human users (such as linguists, translators or language learners) in order to retrieve relevant lexical information through a web-based interface. (Periñán y Arcas, 2006: 69)

Las metas para el futuro de los diccionarios son ambiciosas. Para efectos de la consulta humana, incluyen, por ejemplo, que el diccionario se acomode a las necesidades específicas del usuario en cada momento de consulta (Tarp, 2009). La visión del diccionario del futuro ha sido tal vez articulada con mayor entusiasmo por Varantola (2002: 31):

I will be shamelessly selfish and ask for the impossible. I will advocate for a dictionary that will always adapt to my needs and always be ready to provide me with exactly the answer that I need and will also agree with. I also expect the dictionary to be able to give satisfactory answers to those questions that I forget to ask. Varantola (2002: 31)

Para efectos del PLN, las metas incluyen todo las aplicaciones posibles relativas a descodificación y codificación lingüística automática, y tocan el campo de la IA cuando proponen el razonamiento automático. Estas metas implican trabajar no solo con el procesamiento sino también con la comprensión del lenguaje natural, esto es, el que las computadoras sean capaces de derivar significado a partir del lenguaje natural humano.

Para alcanzar estos anhelos lexicográficos, de Schryver (2003) afirma que se requerirán estructuras masivas de datos, que tal vez no conformen un diccionario como tal, sino que permitan que el diccionario se constituya al momento de consulta. La lexicografía tendrá que evolucionar, según ese autor, en aspectos que –en gran medida– están vinculados al desarrollo mismo de las tecnologías: espacio y velocidad; uso de multimedia; corpus y programas adicionales; acceso (comunicación con la base de datos; búsqueda en la totalidad de la base de datos; acceso a la macroestructura; acceso a la microestructura; acceso a la medioestructura); y facilidad del manejo y entallado a las necesidades del usuario o de la aplicación de PLN.

2. Diccionarios electrónicos

2.1 Inglés

En la actualidad, la mayor parte de los diccionarios impresos importantes en inglés cuenta con una versión digital²⁸, mientras que en español existe un buen número, pero no es la mayoría. En general, aunque representan el cambio de la consulta en papel a la consulta electrónica, los contenidos de estos diccionarios son relativamente cercanos, si no idénticos, a los de las versiones impresas (Atkins, 1996; Nesi, 2000; de Schryver, 2003; Bergenholtz *et al.*, 2009). En ese sentido, se podría decir que la mayoría aún pertenecen al grupo (2) antes descrito. Tarp (2009: 28) recoge esta idea en la siguiente cita:

In 99.9 per cent of the cases, these new electronic dictionaries and lexicographic works are more or less tailored to the same pattern as their printed relatives with the only exception that data access is generally much quicker due to the use of search machines and links.

Las limitaciones de los diccionarios tradicionales pueden causar ciertas molestias al usuario que los consulta, pero en la lexicografía electrónica pueden tener repercusiones más serias. Recordemos que los diccionarios electrónicos necesitan ser absolutamente autosuficientes en información. Velardi *et al.* (1991: 157-158, traducción mía) enumera estas limitaciones:

²⁸ Nesi (2009) ofrece un exhaustivo estado de la cuestión sobre los diccionarios electrónicos, específicamente los diccionarios monolingües o bilingües para hablantes del inglés.

(a) Las definiciones son frecuentemente circulares, por ejemplo, la definición de un término A puede referirse a un término B que, a su vez, se define con A;

(b) las definiciones no son homogéneas en cuanto a la calidad y cantidad de información provista: pueden ser muy difusas, o dar información estructural detallada o listar ejemplos de tipos de usos o describir rasgos internos;

(c) un diccionario es el resultado de un esfuerzo de conceptualización llevado a cabo por algún especialista humano; este esfuerzo puede ser consistente con, o apropiado para, los objetivos de una aplicación para la cual se construye un procesador de lenguaje.

Hay varios diccionarios electrónicos que comienzan a moverse del grupo (2) al (3), según de Schryver (2003), porque empiezan a rediseñar su organización para la plataforma electrónica. Entre los principales diccionarios electrónicos del inglés que están disponibles en Internet, se incluyen:

Tabla 7. Principales diccionarios en inglés disponibles en Internet

| Título | Dirección electrónica |
|---|---|
| <i>Oxford English Dictionary</i> | http://www.oed.com |
| <i>Longman Dictionary of Contemporary English</i> | http://www.ldoceonline.com/ |
| <i>Merriam-Webster Dictionary</i> | http://www.merriam-webster.com |
| <i>Collins Cobuild English Dictionary & Thesaurus</i> | http://dictionary.reverso.net/english-cobuild/ |
| <i>Cambridge Dictionaries Online</i> | http://dictionary.cambridge.org |
| <i>The American Heritage of the English Language</i> | http://dictionary1.classic.reference.com/help/ahd4.html |
| <i>McMillan Dictionary</i> | http://www.macmillandictionary.com |

De los anteriores, nos interesa destacar el *Oxford English Dictionary* y los *Cambridge Dictionaries*, como ejemplos de diccionarios que comparten rasgos del grupo (2) y el (3), porque comienzan a explotar la plataforma electrónica y a organizarse en torno a un eje semántico o conceptual. El primero está vinculado a un tesoro histórico (*Historical Thesaurus of English*), que contiene una clasificación taxonómica de la mayoría de los sentidos y lemas del diccionario. Los editores lo describen como un tipo de índice semántico al diccionario. Contiene un total de 800,000 palabras y sentidos, vinculados a un total de 235,000 categorías semánticas. La palabra *lose*²⁹, por ejemplo, se vincula en este tesoro con 31 categorías. Presentamos una muestra en la imagen siguiente:

²⁹ El verbo *perder* está entre las palabras que estudiamos en este trabajo, por lo que, al comentar diferentes recursos lexicográficos a lo largo de esta tesis, aprovecharemos para mirar el tratamiento que le dan a *perder*, al igual que a su equivalente en inglés, *lose*.

Imagen 1. Entrada de *lose* en el *Historical Thesaurus of English* (vista parcial)

Query has found 31 result(s) for the word 'lose' in the following categories.

- Disordered hearing 01.02.01.01.04.13.01 [Browse this category](#)
- Practice of healing art 01.02.01.03.02.02. [Browse this category](#)
- Killing 01.02.02.04. . . [Browse this category](#)
- Slim shape/physique 01.02.05.02.02. . [Browse this category](#)
- Hunting 01.02.08.04. . . [Browse this category](#)
- Not seeing/preventing from seeing 01.03.07.04. . . [Browse this category](#)
- Action/process of leaking 01.04.05.07.03. . [Browse this category](#)

En cada categoría semántica, el tesoro expone las otras unidades léxicas que se vinculan a ella diacrónicamente. Por ejemplo para la categoría de ‘Slim shape/physique’, nos muestra cientos de unidades léxicas que, a través de la historia de la lengua inglesa, se han vinculado a esa categoría. Presentamos en la imagen siguiente algunas unidades léxicas que se asocian con la categoría ‘Slim shape/physique’, incluida *lose*:

Imagen 2. Categoría de ‘Slim shape/physique’ en el *Historical Thesaurus of English* (vista parcial)

| | |
|----------------------------------|----------------------|
| (.shrink) | lank |
| (.shrink) | pitch |
| (.shrink) | fine down |
| (.shrink) | fall in |
| Slim | slenderize |
| Slim | slighten |
| (.shed weight, fatty tissue etc) | lose |
| (.thin) | mægerian |
| (.thin) | (ge)þynnian |
| (.thin) | lense < (ge)hlænsian |
| (.thin) | lean < hlænian |

El tesoro y el diccionario están conectados en las publicaciones de *Oxford*, así que las búsquedas pueden iniciar desde cualquiera de los dos: en el diccionario, la búsqueda comienza desde la palabra, mientras que en el tesoro, la búsqueda puede comenzar desde el sentido. Al organizar el contenido por significado, en vez de solo alfabéticamente, el tesoro

ofrece una nueva perspectiva sobre el material del diccionario y destaca información que de otro modo no sería accesible. Se añade una nueva dimensión a la experiencia lexicográfica al mostrar las palabras dentro de una categoría semántica, en relación con sus sinónimos y con palabras afines, organizadas de lo más general a lo más específico. Tener acceso a esas posibilidades, sin embargo, requiere pago.

Los *Cambridge Dictionaries*, por su parte, vinculan los sentidos de las palabras primero con nociones conceptuales, dentro de las cuales ubican las definiciones lexicográficas. Por ejemplo, en los diccionarios *Cambridge*, *lose* está vinculado con seis nociones conceptuales:

Imagen 3. Entrada de *lose* en <http://dictionary.cambridge.org>

Results for lose

lose was found in the American English Dictionary at the entries listed below.

- **lose** v NOT BE ABLE TO FIND
- **lose** v NO LONGER POSSESS
- **lose** v BE DEFEATED
- **lose** v NOT MAINTAIN
- **lose** v HAVE LESS OF
- **lose** v CONFUSE

Cada uno de estas nociones conceptuales se precisa con una o varias acepciones lexicográficas. Por ejemplo, bajo el concepto NO LONGER POSSESS, se presentan tres definiciones. En algunas definiciones se ilustra la combinatoria léxica del verbo con roles temáticos en la definición y luego ejemplos de uso: “if you lose someone, that person dies: *George lost his wife in 1990*”. El problema es que esta información no parece ser parte de una estructura formal que permita su explotación.

Imagen 4. Definiciones de *lose* (NO LONGER POSSESS) en <http://dictionary.cambridge.org>

lose

US  /luz/ V [T] (PAST TENSE AND PAST PARTICIPLE lost /lost/)
(NO LONGER POSSESS)

Definition



➤ **to no longer have something, because it has been taken away from you, either by accident or purposely:**

Workers will lose their jobs if the plant closes.

He lost his leg in a car accident.

➤ **If you lose someone, that person dies:**

George lost his wife in 1990.

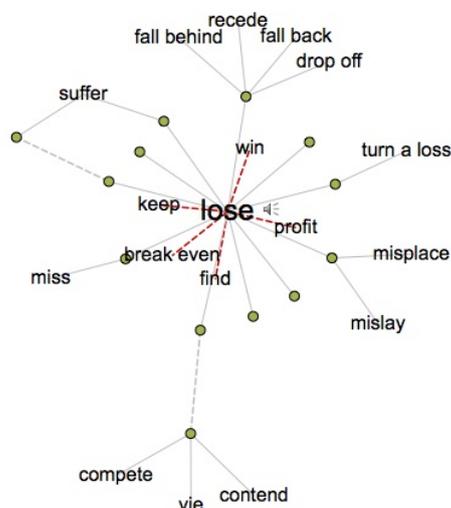
➤ **If you lose money you have risked, you do not make a profit and do not get your money back.**

➤ **A business that is losing money is spending more money than it is receiving.**

(Definition of lose v ((NO LONGER POSSESS)) from the Cambridge Academic Content Dictionary © Cambridge University Press)

Los diccionarios *Cambridge* también se relacionan con un tesoro visual (*Visual Thesaurus*), una herramienta interactiva que presenta mapas de relaciones de palabras y conceptos. Este tesoro permite que las relaciones se visualicen partiendo de diferentes “núcleos”. La imagen siguiente replica el tesoro visual de *lose*: cada rama que parte del núcleo con puntos verdes indica una rama de sentidos diferentes para *lose*; las rayas entrecortadas rojas señalan antónimos:

Imagen 5. Tesoro visual de la palabra *lose* (<http://www.visualthesaurus.com>)



Los diccionarios *Oxford* y *Cambridge* muestran algunas de las ventajas de la plataforma electrónica para relacionar la información contenida en los diccionarios, especialmente cuando hay un eje conceptual central. Estos diccionarios nos dan una idea de lo que se podría lograr al vincular los datos lingüísticos de REDES con una base de conocimiento léxico conceptual como FunGramKB, que cuenta con una ontología conceptual como eje central. Ese vínculo permitiría también explotar las búsquedas no solo a partir de la palabra, sino también a partir de la noción conceptual.

Nuestro trabajo con REDES y FunGramKB busca además añadir la representación conceptual de la combinatoria léxica sistemática de los predicados, de manera que las restricciones de selección participen también de un vínculo con la ontología. El codificar la combinatoria léxica conceptualmente nos permitiría hacer búsquedas no solo de la palabra o noción seleccionadora, sino de la palabra o noción seleccionada también. En ese respecto, estos dos diccionarios parecen limitarse a la representación de la estructura argumental del verbo en sus diferentes sentidos.

2.2 Español

En español, la utilización de computadoras para la preparación de diccionarios (grupo 1) se tardó más que en el resto de Europa. Reporta Dodd que el *Diccionario de la Lengua Española* (1984) es “probably the last large European dictionary to be completed using exclusively the traditional methods of hadwriten slips and letterpress composition and printing” (1989:85). Hoy, aunque los principales diccionarios del español tienen una versión electrónica, esta es prácticamente idéntica a sus predecesores en papel; los títulos más conocidos se muestran en la tabla siguiente. Cabe señalar que el DRAE (2001) también tiene una versión en CD-ROM, que permite ciertas búsquedas más avanzadas, y que el *Diccionario de Uso del Español* de María Moliner (1998) tiene una versión en CD-ROM, pero no en Internet (aparte de una versión no oficial).

Tabla 8. Diccionarios en español disponibles en Internet

| Título | Dirección electrónica |
|--|---|
| <i>Diccionario de la Lengua Española</i> (DRAE) | http://buscon.rae.es/draeI/ |
| <i>Diccionario CLAVE</i> | http://clave.librosvivos.net |
| <i>Diccionario Salamanca de la Lengua Española</i> | http://fenix.cnice.mec.es/diccionario |
| <i>VOX y Larousse</i> | http://www.diccionarios.com |
| <i>Diccionario de Sinónimos y Antónimos</i> (Espasa Calpe) | http://www.wordreference.com/sinonimos/ |
| <i>Diccionario de Sentidos Verbales</i> | http://grial.uab.es/sensem/verbos/login |

No tienen versión electrónica diccionarios importantes del español como el *Gran Diccionario de Uso del Español Actual* de Aquilino Sánchez (2001) y el *Diccionario del Español Actual* de Manuel Seco (1999), ni tampoco diccionarios con enfoques particulares, como el mismo *REDES. Diccionario combinatorio del español contemporáneo* y *Diccionario práctico del español contemporáneo* de Ignacio Bosque (2004b y 2006, respectivamente), a pesar de que todos se fundamentan en grandes corpus del español. Igualmente, no tienen versión electrónica diccionarios dialectológicos de gran alcance, como el *Diccionario de Americanismos*, de la Asociación de Academias de la Lengua Española (2005).

Están surgiendo nuevos diccionarios en español con enfoques especializados, que pertenecen a la transición del grupo (2) al (3), o al grupo (3) en propiedad, pues son creados en plataformas electrónicas y explotan las posibilidades que esta ofrece. En español, por ejemplo, está el *Diccionario de colocaciones del Español* o DiCE (<http://www.dicesp.com/paginas>), dirigido por Margarita Alonso Ramos, un diccionario de colocaciones del español fundamentado en la Teoría Sentido Texto (Mel'čuk, 1995). Existe además el *Diccionario de partículas discursivas del español* (<http://textodigital.com/P/DDPD/>), dirigido por Antonio Briz. Por otro lado, varias de las bases de conocimiento multilingües existentes incluyen el español, incluidas EuroWordNet, FrameNet –que incluye Spanish FrameNet, dirigido por Carlos Subirats (Subirats, 2009; Subirats y Petruck, 2003). – y FunGramKB.

3. Lexicones electrónicos

Los términos “lexicón electrónico” y “base de datos léxica” son equivalentes: apuntan a nociones amplias que abarcan prácticamente todas las bases de datos lexicográficas creadas

en plataformas electrónicas y procesables computacionalmente. Por su parte, los términos “tesauro electrónico” y “base de conocimiento léxico” se refieren a entidades con características más específicas. En todos los casos se trata de recursos que han sido creados directamente en Internet, con el propósito de servir el PLN, aunque –como hemos dicho antes– algunos de ellos tienen una interfaz para consultas humanas.

Los lexicones pueden estar diseñados específicamente para manipular información léxica o pueden ser bases de datos multipropósito a las que se ha añadido información léxica. Entre los principales lexicones electrónicos existentes, se destacan WordNet, EuroWordNet y FrameNet, que discutiremos en los siguientes apartados. Estos lexicones electrónicos coinciden al considerar el lexicon de un idioma un repositorio de miles de conceptos y palabras vinculados unos a otros en una gran red (Fontenelle, 2000a; Zaenen, 2002).

Los “tesauros” son bases de datos léxicas con una característica particular: organizan las palabras principalmente mediante la relación de sinonimia. En el tesauro, cada categoría corresponde aproximadamente a un sentido, por lo que una palabra polisémica aparece bajo diferentes categorías. Cuando varias palabras diferentes están ubicadas en una misma categoría, se trata de palabras relacionadas semánticamente. Los tesauros pueden o no ofrecer definiciones explícitas. El tesauro más usado para la WSD es probablemente el *Roget’s International Thesaurus*. El *Historical Thesaurus of English* de la editorial Oxford, que vimos antes, fundamenta el nivel más alto de su taxonomía en categorías semánticas como las del *Roget’s*. WordNet tiene rasgos de tesauro también, como veremos más adelante.

Las “bases de conocimiento léxicas” pertenecen al nivel más complejo de desarrollo en este campo: son estructuras robustas, tanto en cuanto a información como a procesos. Se trata de repositorios de información sobre las unidades léxicas y sus relaciones, que proveen mecanismos de recuperar, organizar, investigar, utilizar o compartir la información. Las bases de conocimiento léxicas pueden servir muchas aplicaciones de la lingüística computacional, la IA y las ciencias de la información. Algunas son capaces de sugerir soluciones a problemas esporádicos o aprender de la experiencia. Pueden contener información derivada de los MRD, textos completos de libros de referencia, resultados de análisis estadísticos de uso real o datos sobre el conocimiento humano del mundo. En este grupo de lexicones se incluye FunGramKB.

3.1 WordNet

WordNet (Miller *et al.*, 1990) (<http://wordnet.princeton.edu>) es una base de datos léxica de gran tamaño del inglés, elaborada por la Universidad de Princeton. Motivada en sus inicios por la psicolingüística, WordNet fue diseñada principalmente como un recurso computacional para dar cuenta de la capacidad humana de categorización lingüística (Fellbaum, 1998). Por ello, cada categoría gramatical –verbos, nombres, adjetivos y adverbios– se analiza por separado de las otras. WordNet también quería poner a prueba las teorías sobre la memoria semántica humana que alegaban que el conocimiento sobre un concepto incluía también conocimiento del concepto superordinado y de sus partes.

WordNet contiene una extensa serie de categorías semánticas llamadas *synsets*, que en su versión actual (3.0) alcanza los 117,000. Los *synsets* son series de sinónimos cognitivos, cada uno de los cuales expresa un concepto distinto. La decisión de organizarse a través del significado léxico, y no de las unidades léxicas, le otorga a WordNet rasgos tanto de diccionario tradicional como de tesoro (Boas, 2005). Los *synsets* agrupan unidades léxicas con significados similares; por ejemplo, *chair* y *bench* son unidades léxicas similares y se agrupan en el *synset* {*chair*, *bench*}. Otro sentido de *chair* ('cátedra') aparece en el *synset* {*professorship*, *chair*}. Los *synsets*, por lo tanto, no están compuestos de unidades léxicas, sino de significados léxicos, que se definen de manera indirecta por el *synset*. Cada *synset* tiene una glosa explicativa; por ejemplo, para {*chair*, *bench*} es: 'a long seat for more than one person' y para {*professorship*, *chair*} es 'the position of professor'. La polisemia se maneja en WordNet, por lo tanto, mediante la asociación de la unidad léxica con varios *synsets*.

Los *synsets* se organizan de forma jerárquica, de modo que los más generales incluyen a los más específicos. Se trata de una relación tipo "IS-A", de superordinados a subordinados, o de hiperónimos a hipónimos. Estas relaciones varían según cada una de las cuatro sub-bases en las que se divide WordNet, que responden a la categoría gramatical de sus miembros: verbos, nombres, adjetivos y adverbios. Las jerarquías nominales empiezan en el nodo raíz {*entity*}. Un *synset* general como {*motor vehicle*, *automotive vehicle*} incluye otros *synsets* progresivamente más específicos, como {*car*, *auto*, *automobile*, *machine*} y luego {*ambulance*}. La relación jerárquica de hiponimia es transitiva: si una ambulancia es un tipo de automóvil y un automóvil es un vehículo de motor, entonces una ambulancia es un tipo de un vehículo de motor. WordNet codifica también la meronimia, la relación de parte-todo que ocurre entre *synsets* como {*car*, *auto*, *automobile*, *machine*} y {*car mirror*} o

{gasoline engine, petrol engine}. Señala Boas (2005) que una de las ventajas de estructurar una red semántica de nombres de esta manera es que revela posibles lagunas léxicas al comparar idiomas; ello puede apuntar a diferencias en estructuras conceptuales, más que en estructuras léxicas.

Los *synsets* de verbos se organizan en jerarquías también. Los verbos ubicados hacia la parte baja del árbol son “tropónimos” que expresan maneras cada vez más específicas de caracterizar un evento. En la secuencia {communicate}-->{talk}-->{whisper} se ilustran maneras de comunicación que se hacen cada vez más específicas. La dimensión específica de variación dependerá del campo semántico, por ejemplo {move}-->{jog}-->{run} ilustra un cambio en velocidad y {like}-->{love}-->{idolize} ilustra un cambio en intensidad de emoción.

Los sentidos de WordNet son de una granularidad muy fina, más que los de la mayoría de los MRD, lo que para Tomuro (1998) causa que sea un recurso muy ambiguo y que la asignación de sentidos resulte muy difícil. Por ejemplo, en la versión actual de WordNet el verbo *write* tiene ocho sentidos y el nombre *book* tiene once. Esto quiere decir que un proceso automático de desambiguación semántica tendrá que elegir entre ocho posibles sentidos para *write* y once para *book*. Para ilustrar la granularidad semántica de WordNet y compararla con los diccionarios que hemos visto antes, la próxima imagen muestra la entrada de *lose* en WordNet, que tiene 10 sentidos:

Imagen 6. Entrada de *lose* en WordNet (<http://wordnet.princeton.edu/>)

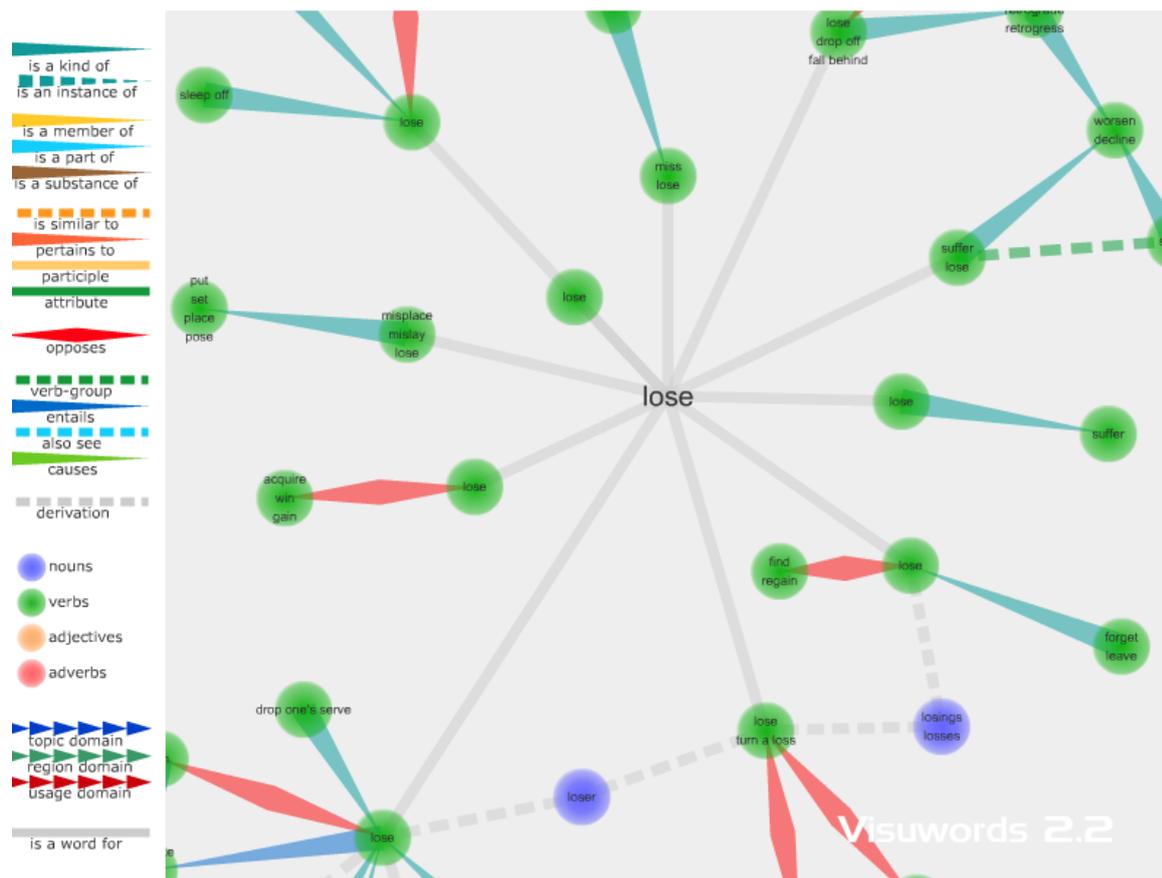
Verb

- **S:** (v) **lose** (fail to keep or to maintain; cease to have, either physically or in an abstract sense) *"She lost her purse when she left it unattended on her seat"*
- **S:** (v) **lose** (fail to win) *"We lost the battle but we won the war"*
- **S:** (v) **lose** (suffer the loss of a person through death or removal) *"She lost her husband in the war"; "The couple that wanted to adopt the child lost her when the biological parents claimed her"*
- **S:** (v) **lose** (miss from one's possessions; lose sight of) *"I've lost my glasses again!"*
- **S:** (v) **lose** (allow to go out of sight or mind) *"The detective lost the man he was shadowing after he had to stop at a red light"; "lose the crowds by climbing a mountain"; "the lost tribe"*
- **S:** (v) **lose**, **turn a loss** (fail to make money in a business; make a loss or fail to profit) *"I lost thousands of dollars on that bad investment!"; "The company turned a loss after the first year"*
- **S:** (v) **lose** (fail to get or obtain) *"I lost the opportunity to spend a year abroad"*
- **S:** (v) **miss**, **lose** (fail to perceive or to catch with the senses or the mind) *"I missed that remark"; "She missed his point"; "We lost part of what he said"*
- **S:** (v) **lose** (withdraw, as from reality) *"he lost himself in his music"*
- **S:** (v) **suffer**, **lose** (be set at a disadvantage) *"This author really suffers in translation"; "The painting loses something in this light"*

La asignación de sentidos en WordNet tiene un vínculo indirecto con las colocaciones de una palabra. Como vemos arriba, en algunas glosas, se mencionan algunas valencias semánticas: “suffer the loss of a person through death or removal” o “miss from one’s possessions” (subrayado nuestro). Asimismo, se ilustran los argumentos del verbo en los ejemplos: “She lost her husband in the war” o “I’ve lost my glasses again” (subrayado nuestro). Sin embargo, no se hace ninguna generalización a partir de esa información.

Los datos de WordNet se han traducido a un tesoro visual llamado VisuWords (<http://www.visuwords.com>). En este diagrama, cada una de las 10 ramas que sale de *lose* responde a un “synset”:

Imagen 7. Tesauro visual de la palabra *lose* en WordNet (<http://www.visuwords.com>)



Aunque WordNet es uno de los lexicones más utilizados en la actualidad, para efectos de este trabajo, presenta dos limitaciones: en primer lugar, no provee un espacio formal para codificar las preferencias de selección de las palabras, y segundo, la granularidad de 117,000 “synsets” resulta demasiado fina. Aunque WordNet contiene información única sobre las relaciones de las palabras entre sí (hiperonimia, sinonimia, hiponimia, etc.), y las organiza en una taxonomía, en este trabajo nos interesan en particular las relaciones que se puedan revelar entre las palabras relacionadas semánticamente en función de una combinatoria similar.

3.2 EuroWordNet

EuroWordNet (<http://www.illc.uva.nl/EuroWordNet/>) (Peters *et al.*, 1998; Vossen, 2004) es una base de datos multilingüe con redes de palabras en varias lenguas europeas: holandés, italiano, alemán, francés, checo, estonio y español. Las redes de palabras están estructuradas de la misma manera que en WordNet en inglés, en términos de *synsets* que mantienen relaciones semánticas básicas entre sí. Cada WordNet representa un sistema único de lexicalizaciones dentro de cada idioma.

Al igual que WordNet, las relaciones de sentido que establece EuroWordNet entre las palabras semánticamente relacionadas –hiponimia, antonimia, meronimia, etc.– representan relaciones entre los *synsets*. Las relaciones de sentido son, pues, internas a la arquitectura conceptual de EuroWordNet. El acercamiento relacional al significado –es decir, el establecer el significado mediante las relaciones de palabras con otras palabras–, característico de una semántica superficial, se ha señalado como una limitación de WordNet y EuroWordNet, frente a los acercamientos de semántica profunda en que el significado se define *a priori*, y luego vinculan las unidades léxicas a estas estructuras semánticas o conceptuales. En ese sentido, Boas (2009a) señala que EuroWordNet depende de conceptos descontextualizados para realizar sus descripciones léxicas, porque cada idioma tiene su serie de conceptos y relaciones léxico semánticas basadas en los patrones de lexicalización de cada lengua.

Por su parte, Perrián y Arcas (2007b) añaden que el acercamiento relacional al significado que utiliza EuroWordNet tiende a causar mayor redundancia en la base de datos y tener menor poder expresivo que el acercamiento de semántica profunda. Los autores señalan como ejemplo de redundancia la relación de meronimia que se identifica entre pares de palabras en cada idioma; en este caso, la relación entre las palabras que designan ‘pájaro’ y ‘pluma’ se tienen que repetir tantas veces como idiomas haya:

Tabla 9. Relación de meronimia en EuroWordNet (tomado de Perrián y Arcas, 2007: 7)

| Palabra | Relación | Palabra |
|---------|---------------|---------|
| bird | HAS_MERO_PART | feather |
| Vogel | HAS_MERO_PART | Feder |
| pájaro | HAS_MERO_PART | pluma |
| ucello | HAS_MERO_PART | piuma |
| oiseau | HAS_MERO_PART | plume |

La redundancia causada por el multilingüismo no ocurre en una base de conocimiento como FunGramKB, que está basada en una ontología donde las unidades léxicas de diferentes idiomas con el mismo significado se vinculan a un mismo concepto. El concepto se define una vez, mediante una representación cognitiva de su significado. Por ejemplo, uno de los enunciados del postulado de significado del concepto +BIRD_00 (‘pájaro’) informa de que sus partes incluyen muchas plumas, dos patas y dos alas. Al concepto +BIRD_00 se vincularían todas las unidades léxicas detalladas en la tabla anterior, por lo cual la especificación de las partes del ‘pájaro’ se haría una sola vez. Las bases de conocimiento que definen formalmente cada concepto tendrán mayor poder expresivo, o sea, mayor capacidad de expresar el contenido cognitivo de las unidades léxicas que las definiciones relacionales de EuroWordNet. Una ontología que define formalmente cada concepto supone, por supuesto,

un esfuerzo mucho mayor de construcción que una base de datos relacional.

No obstante lo anterior, EuroWordNet reconoce la importancia de contar con un metalenguaje verdaderamente independiente en la construcción de las bases de datos. Existe ya un primer paso en el índice interlingüístico (Inter Lingual Index, o ILI) de 1024 conceptos fundamentales que vincula los WordNets de cada lengua. Cada concepto del ILI consiste de un *synset* y una glosa en inglés que especifica su significado y fuente. Los *synsets* de cada lengua pueden estar vinculados directamente al ILI, o establecerse como variantes, por ejemplo, un *synset* en una lengua puede ser un hiperónimo del concepto del ILI. Vossen y Fellbaum (2009: 319) consideran que la conexión de las lenguas no se debe hacer a través de un índice en una lengua en particular, como es el caso del ILI, que conecta los *synsets* de cada lengua a una lista no estructurada de conceptos derivada del inglés. Estos autores defienden la ampliación del modelo de WordNet a un “Global WordNet Grid” (GWG), que incluiría una ontología como base para un índice de conceptos universales. La idea se propuso en 2006 en el Global WordNet Conference en Korea, y plantea que los conceptos universales provengan de la ontología SUMO (Niles y Pease, 2001). En la GWG, las lenguas se vincularán a un índice basado en una ontología formal, independiente de lenguas particulares.

3.3 FrameNet

FrameNet (<https://framenet.icsi.berkeley.edu/fndrupal/>) es un proyecto de lexicografía informatizada sobre la relación entre las propiedades semánticas y las propiedades sintácticas de las palabras del inglés. FrameNet se fundamenta en un gran corpus electrónico, al cual se aplican procedimientos manuales y automáticos, que es capaz de proveer diversos tipos de reportes generados en la web (Fillmore *et al.*, 2003a). Este proyecto comenzó en el International Computer Science Institute de la universidad de Berkeley, California en el 1997, y está disponible gratuitamente para consulta e investigación. Se fundamenta en el trabajo teórico de Fillmore y sus colegas (Fillmore, 1976, 1977, 1982, 1985, entre otros). La idea fundamental de este modelo es que los sentidos de las palabras pueden comprenderse mejor sobre la base de un marco semántico (“semantic frame”), que es una estructura conceptual que describe una situación, objeto o evento particular, y a sus participantes (Ruppenhofer *et al.*, 2010). FrameNet analiza los sentidos de las palabras a través de estos marcos, preguntándose cómo las propiedades semánticas adquieren forma sintáctica.

La unidad básica de estudio de FrameNet es la “unidad léxica” (“lexical unit”, en adelante LU³⁰), que equivale a una palabra en uno de sus sentidos. Por ejemplo, el verbo *bake* es polisémico, así que está vinculado a tres marcos diferentes (Ruppenhofer *et al.*, 2010); esto quiere decir que hay tres unidades léxicas con la forma *bake*:

Tabla 10. Marcos semánticos de *bake* en FrameNet

| Marco semántico | Ejemplo |
|------------------|--|
| Apply-heat | <i>Michelle baked the potatoes for 45 minutes.</i> |
| Cooking_creation | <i>Michelle baked her mother a cake for her birthday.</i> |
| Absorb_heat | <i>The potatoes have to bake for more than 30 minutes.</i> |

Boas (2005) señala que esta categorización de FrameNet se diferencia de la de WordNet porque define los distintos sentidos de una palabra con referencia a los marcos semánticos subyacentes, en vez de a través de otros verbos que pertenecen a distintos *synsets*. Como explican Fillmore y Atkins (1992: 76-77):

Within such an approach, words or word senses are not related to each other directly, word to word, but only by way of their links to common background frames and indications of the manner in which their meanings highlight particular elements of such frames.

Al comparar a EuroWordNet con FrameNet, Boas (2009) señala que los marcos semánticos de FrameNet están motivados por conceptos lingüísticos independientes de las unidades de análisis, a diferencia de los *synsets* de EuroWordNet, donde cada idioma tiene su serie de conceptos y relaciones léxico semánticas particulares. FrameNet codifica las diferencias entre lenguas en el nivel de la realización sintáctica, pero dentro de un marco semántico compartido (Boas, 2009b).

FrameNet, por lo tanto, usa unidades organizacionales independientes más grandes que la palabra: los marcos semánticos. En este sentido, FrameNet se aleja de WordNet y se acerca a FunGramKB, que también se organiza por criterios mayores a la palabra. Sin embargo, FrameNet amplía su interlingua de marcos semánticos según se van añadiendo lenguas, por lo que Boas (2009b) la llama una “representación interlingüística” más que una interlingua como tal; por el contrario, como veremos, FunGramKB crea su interlingua *a priori* del trabajo con lenguas particulares, a partir de conceptos cognitivos de ontologías independientes.

FrameNet expresa la valencia semántica de una palabra en términos de los tipos de entidades que pueden participar en los marcos que “evoca” cada palabra. Estas entidades se

³⁰ “Lexical unit” o LU es un término tomado de Cruse (1986).

llaman “elementos del marco” (“frame elements”, en adelante FE). Por ejemplo, el verbo *to give* puede ser descrita por el marco semántico y FE que se presentan en la tabla siguiente. Esto quiere decir que alguien (DONOR) está en posesión de algo (THEME) y luego causa que otra persona (RECIPIENT) entre en posesión del THEME, lo cual tal vez cause que el THEME se mueva al RECIPIENT. Los FE se definen como roles semánticos específicos a una situación. Difieren de los roles semánticos universales tradicionales (o roles zeta), como Agente y Paciente, en que son específicos a cada marco semántico en que participan (Boas, 2009a).

Tabla 11. El verbo *to give* en FrameNet

| Marco | FE |
|------------------------|-------------------------|
| Transfer ³¹ | DONOR, THEME, RECIPIENT |

Los FE de los marcos semánticos, tanto en FrameNet como en SpanishFrameNet, se definen en relación con el marco semántico específico que se evoca. Esto significa que incluso cuando dos marcos semánticos evocan los mismos FE, estos FE se consideran diferentes por pertenecer a marcos diferentes. Ello responde a la dificultad de establecer una lista fija de etiquetas para estudiar el léxico de las lenguas naturales, incluso a través del análisis de los marcos semánticos (Subirats, 2009: 138, haciendo referencia a varios estudios de Fillmore).

Cada uno de los FE se vincula luego con un tipo de frase (“Phrase type” o PT) y con una función gramatical (“grammatical function” o GF). Toda esa información constituye el “patrón de valencias” del verbo, o la constelación “triple” que describe la realización del verbo en la oración. Una de las ventajas de describir una LU a través de un marco semántico es el reconocimiento de que los FE pueden ocurrir en diferentes posiciones sintácticas y satisfacer diferentes funciones gramaticales. El patrón de valencias de *to give* en la oración *The teacher gave a book to the student* sería:

³¹ En FrameNet, se usan fuentes tipográficas diferentes para distinguir de la lengua general el nombre de los marcos semánticos y de los FE. No debe pensarse, por lo tanto, que los nombres del marco semántico Transfer, o de los elementos DONOR, THEME, RECIPIENT, son lo mismo que las palabras *transfer*, *donor*, *theme* o *recipient* en inglés. Las propiedades específicas de los FE son establecidas por cada marco.

Tabla 12. Patrón de valencias de *to give* en FrameNet

| | | | | |
|----------------|--------------------|-------------|---------------|-----------------------|
| Marco: | Transfer | | | |
| Oración | <i>The teacher</i> | <i>gave</i> | <i>a book</i> | <i>to the student</i> |
| FE: | DONOR | | THEME | RECIPIENT |
| PT: | NP | | NP | PP ³² |
| GF: | Ext | | Obj | Comp |

Los lexicógrafos de FrameNet no empiezan su trabajo por la LU, sino por el marco semántico. Primero proponen una descripción informal del tipo de evento que el marco representa. Después identifican una serie de palabras que pertenecen a este marco. En el caso del marco *Attaching* ('adjuntar'), se identifican verbos como *append*, *attach*, *connect*, *fasten*, *join*, *link*, *secure*, *tie*, etc. (Fillmore *et al.*, 2003b). En tercer lugar, exploran el corpus de manera inductiva para ver cómo se manifiestan esos verbos. En el caso del marco *Attaching*, por ejemplo, sale a relucir un tipo de FE recurrente, que llaman *CONNECTOR*. Se trata del elemento que conecta las dos cosas que se unen; entre las palabras más típicas que lexicalizan al *CONNECTOR* están *string*, *cord*, *wire*, etc. Estas se registran en la base de datos como "colocaciones típicas". Después de explorar todas las LU asociadas al marco, se generan informes finales que detallan las valencias y las realizaciones posibles de los FE para cada LU dentro de cada marco definido.

El tratamiento de FrameNet de la polisemia radica en el establecimiento de los marcos semánticos, no en la identificación de diferentes sentidos o realizaciones sintácticas. Al proveer descripciones de las posibilidades semánticas y de combinación sintáctica de una palabra, se comprueba que muchas variaciones sintácticas responden al mismo marco semántico subyacente, por lo que no requieren múltiples acepciones. En cuanto a su aplicación a la WSD para el PLN, Boas (2005) alega que los marcos semánticos de FrameNet ofrecen información que permite la desambiguación de casos que de otro modo no podrían desambiguarse. Por ejemplo, para distinguir los sentidos de *cure* en *Nancy cured the ham/the patient*, es necesario saber qué marco semántico activa cada uno de los NP posverbiales, en este caso, *Cure* o *Preserving*.

Veamos ahora el registro del verbo *lose* en FrameNet, que distingue 10 marcos semánticos:

³² Las abreviaturas significan: NP-Noun Phrase, PP- Prepositional Phrase, Ext- External Argument, Obj-Grammatical Object, Compl-Complement.

Imagen 8. Marcos semánticos de *lose* en FrameNet

Lexical unit search results: Closest match is lose

| Lexical Unit | Frame |
|--------------|----------------------------------|
| lose it.v | Losing_it |
| lose.v | Finish_competition |
| lose.v | Earnings_and_losses |
| lose.v | Losing_someone |
| lose.v | Losing |
| lose.v | Finish_game |
| lose.v | Losing_track_of_theme |
| lose.v | Losing_track_of_perceiver |
| lose.v | Change_of_quantity_of_possession |
| lose_sleep.v | Emotion_active |

En la próxima imagen, observemos en detalle el segundo marco semántico de *lose*, ‘Finish_competition’.

Imagen 8. Marco semántico de 'Finish_competition' en FrameNet (vista parcial)

Finish_competition [Lexical Unit Index](#)

Definition:

A **Competition** comes to an end, with a **Competitor** tying, winning, or losing against an **Opponent** (or opponents) according to their **Score**. (In the symmetrical situation of a tie, the **Competitor** and **Opponent** may be jointly mentioned as **Competitors**.) According to the **Score** that the **Competitor** and **Opponent** have achieved, they may receive a **Rank** or a **Prize**. The **Margin** may also be mentioned.

Unbelievable--he **LOST** the gold medal by just .02 points!

Looks like it will come to a **DRAW** between Woods and Stephenson.

FEs:

Core:

| | |
|---|---|
| Competition [] | The event in which the Competitors compete with each other. |
| Competitor [] Semantic Type: Sentient Excludes: Competitors | The individual involved in the Competition . |
| Competitors [] | The individuals who compete against each other. |
| Opponent [] Requires: Competitor Excludes: Competitors | The individual that the Competitor has been competing against. |

Según la imagen anterior, el marco de Finish_competition generalmente involucra una competencia (Competition) que termina cuando una persona que compite (Competitor) gana, pierde o empatara con otro (Opponent), según cuánto anote (Score). Las palabras que evocan este marco, además de *lose*, son *draw*, *fold*, *loss*, *show*, *tie*, *victor*, *victorious*, *victory*, *win*, *winner*.

Como diccionario para humanos, FrameNet tiene más de 10,000 sentidos, con ejemplos que ilustran significado y uso. Como recurso para el PLN, tiene más de 170,000 oraciones anotadas manualmente que proveen datos de entrenamiento para etiquetado de marcos semánticos. Los marcos semánticos están enlazados por un sistema de relaciones, que vinculan los marcos más generales a marcos más específicos. Hay varios proyectos de construcción de FrameNets paralelos al inglés para otras lenguas, incluyendo español, alemán, chino, japonés. El Spanish FrameNet, como hemos mencionado antes, está a cargo de Carlos Subirats (Subirats y Petruck, 2003; Subirats, 2009).

4. Ontologías

En esta sección nos detenemos en las ontologías, por el importante rol que asumen en muchas bases de conocimiento léxicas, incluida FunGramKB. Las ontologías son una herramienta clave en los sistemas de manejo de contenido y conocimiento, pues determinan el conjunto de categorías semánticas que refleja la organización conceptual del dominio de información sobre el cual el sistema debe operar; en ese sentido, son un puente formal entre la representación del conocimiento y la léxico semántica computacional (Lenci, 2000).

Según Hanks (2013, §1.7, traducción mía), el término “ontología” tiene al menos tres sentidos actualmente. En esta tesis, al hablar de ontologías, nos referimos a la noción descrita en (ii):

(i) En la filosofía tradicional, “ontología” denota la concepción aristotélica del objeto de la investigación científica: el mundo y todo lo que incluye, físico y metafísico, en otras palabras, todo lo que existe.

(ii) En lingüística computacional, se usa el término para referirse al vocabulario completo de una lengua (o, al menos, a todas sus palabras con contenido). El mejor ejemplo es WordNet, una red de relaciones semánticas entre palabras del inglés y conceptos [...]. Las ontologías como WordNet representan relaciones conceptuales entre unidades léxicas en una red semántica, en particular las relaciones jerárquicas, las llamadas jerarquías IS-A. Las jerarquías IS-A funcionan bien para términos que denotan tipos naturales y artefactos, por ejemplo:

-a *canary* is a *finch* is a *bird* is a *living being* is a *physical object* is an *entity*
-a *gun* is a *weapon* is an *artifact* is a *physical object* is an *entity*³³

No funcionan tan bien para los términos abstractos, o las palabras que denotan partes, propiedades, atributos y ciertas clases de verbos. Se necesita algo más.

(iii) En la terminología de la web semántica, “ontología” se usa actualmente para denotar arreglos organizados de entidades prácticas, como nombres de personas, direcciones, fechas, eventos, citas, productos comerciales, descripciones de productos, en vez de palabras y significados.

La rama de estudio teórico que enmarca la creación de ontologías se llama “semántica ontológica”. Nirenburg y Raskin (2001) la definen como una teoría de significado del lenguaje natural y un acercamiento al PLN que usa las ontologías como el recurso central para extraer y representar significados de textos de lenguaje natural, razonar sobre el conocimiento derivado de textos y generar textos de lenguaje natural basados en representaciones de su significado. La semántica ontológica respalda a la traducción automática, la extracción de información, el resumen de textos, las respuestas a preguntas, los

³³ Mantenemos los ejemplos de Hanks en el inglés original tomado de WordNet. Cabe señalar que Hanks aclara que las jerarquías IS-A de hiperónimos de WordNet para *canary* y *gun* son más complicadas de lo que él señala en el ejemplo, e incluyen varios pasos intermedios sutiles.

trabajos colaborativos entre humanos y programas computacionales, etc. Entre las ontologías más conocidas se encuentran:

- SUMO (*Suggested Upper Merged Ontology* - [http:// www.ontologyportal.org/](http://www.ontologyportal.org/)), que ha sido vinculada a todas las unidades léxicas de WordNet;
- DOLCE (*Descriptive Ontology for Linguistic and Cognitive Engineering* - <http://www.loa.istc.cnr.it/DOLCE.html>),
- CyC (<http://cyc.com/>);
- Microkosmos (Mahesh, 1996; Mahesh y Nirenburg, 1995)
- WEBode (<http://mayor2.dia.fi.upm.es/oeg/>).

Lenci (2000) identifica las varias decisiones que se tienen que tomar a la hora de diseñar una ontología. En primer lugar, hay que elegir el tipo de conocimiento que se va a representar, pues de esto dependerá la estructura que se construya. Existe una distinción básica entre el “conocimiento terminológico” o conocimiento específico a un dominio y el “conocimiento general”. Usualmente el conocimiento terminológico es más homogéneo y está explícitamente estructurado, por ejemplo, el ordenamiento de las plantas o animales, o de las funciones particulares de una computadora. Por otro lado, el conocimiento general –que puede incluir todo el conocimiento humano– es muy heterogéneo y su estructura es más laxa. Ello hace que el primero sea más fácil de representar en una ontología que el segundo, pero naturalmente limita su alcance al ámbito que se trata y el alcance de sus aplicaciones.

Un segundo asunto importante a la hora de diseñar una ontología es el propósito que cumplirá: hay ontologías que tienen usos específicos y otras que son multipropósito. Las ontologías que sirven una tarea específica normalmente restringen su contenido a ese ámbito particular, y destacan las conexiones entre los tipos semánticos que mejor sirvan el propósito establecido. Lenci (2000) ofrece el ejemplo de una ontología que maneje las correlaciones entre los accidentes automovilísticos y el tipo de auto y la edad de los conductores; esta ontología contendrá una clasificación de granularidad fina de esos tres tipos de elementos. Estas ontologías tienden a ser muy eficientes, pero no son flexibles ni portables a otras tareas. Por su parte, una ontología multipropósito no tiene las restricciones que impone una tarea específica, pero ello le requiere construir una arquitectura flexible, lo cual es más complicado. Las ontologías multipropósito o “generales” permiten compartir recursos y portar aplicaciones a múltiples dominios. Usualmente son construidas de arriba hacia abajo, e intentan lograr una cobertura universal de las categorías humanas. Una de sus ventajas es que constituyen una suerte de interlingua para sistemas que manejan conocimiento de

diferentes ámbitos, porque estandarizan y uniforman el conocimiento, de manera que se pueda compartir y reusar.

Finalmente figura el problema de determinar los tipos semánticos que serán incluidos en la ontología, y cómo se definirán. Lenci (2000) explica que una ontología es un sistema de categorías que son seleccionadas por su utilidad para capturar correlaciones y similitudes entre pedazos de la realidad. Los “tipos” son las herramientas de clasificación. Esto requiere que estén asociados con definiciones que establezcan las condiciones que una entidad debe satisfacer para estar subsumida o clasificada dentro de un concepto. Sowa (2000) establece dos tipos de definiciones: las definiciones axiomáticas del sistema de tipos y las definiciones basadas en prototipos. Estos dos tipos de definiciones son más eficaces en los casos de las ontologías de dominios específicos, en las que es más fácil definir los conceptos en términos de condiciones suficientes y necesarias. Por el contrario, las ontologías generales tienden a proveer definiciones implícitas e informales, lo que genera un alto nivel de vaguedad y ambigüedad. Las ontologías Mykrokosmos y Cyc, citadas arriba, son ontologías generales y grandes; por ejemplo, Cyc tiene más de 100,00 tipos conceptuales.

Las ontologías conceptuales hacen una distinción entre las palabras particulares del idioma y los conceptos independientes del lenguaje. Según Vossen y Fellbaum (2009) hay tres preguntas fundamentales que provocan las ontologías:

- (1) si realmente existe un “lexicón universal”, es decir, un inventario de conceptos que están codificados léxicamente (o que podrían estarlo) en todas las lenguas;
- (2) qué tipos de conceptos incluiría un lexicón como ese y cuán grande sería el núcleo común de conceptos lexicalizados para la mayoría de las lenguas; y
- (3) cómo irradian desde ahí las lexicalizaciones específicas a cada idioma.

Sobre estas preguntas, conviene tomar en cuenta la postura de Velardi *et al.* (1991) y recordar que el rol de estos constructos no es necesariamente replicar el conocimiento lingüístico humano, sino producir alguna descripción formal del fenómeno que permita alguna aplicación del PLN:

No language model can prove to be objective, or even plausible. In principle, language rules and primitives do not exist. But even though symbols are arbitrary, their role is not to mimic human comprehension, but rather to produce some formal description of raw textual input, in a form that is ultimately useful for some relevant NLP application (Velardi *et al.*, 1991: 169).

El poder de las ontologías para representar el conocimiento léxico es también su mayor debilidad. Por un lado, constituyen aparatos formales que organizan el significado léxico como tipos ontológicos: un sentido de una palabra se vincula a un tipo particular; los

tipos pertenecen a una estructura; y la estructura da cuenta de las relaciones entre los sentidos de las palabras en términos de las relaciones entre los tipos. Se trata de una organización muy comprensible, en donde las herencias o implicaciones de sentidos están claramente trazadas. Sin embargo, esta clara organización no se corresponde con la realidad del conocimiento léxico, que es complejo e inherentemente heterogéneo, y está implícitamente estructurado. Ello hace que sea extremadamente difícil capturarlo a través de asignaciones únicas a puntos en la ontología.

La diferencia entre ontología y lenguaje natural es un asunto importante, teórica y prácticamente, arguyen Nirenburg y Raskin (2001). Según estos autores, esta distinción no radica en la definitiva ambigüedad de las lenguas naturales y supuesta no-ambigüedad de las ontologías. La diferencia se encuentra en que los conceptos y etiquetas ontológicas son entes estáticos, que han sido deliberadamente construidos y explícitamente definidos, sobre los cuales no puede operar nada involuntariamente para introducir ambigüedad; mientras que el lenguaje natural está sujeto a eventos que no se pueden controlar y que introducen ambigüedad. Lenci (2000) señala además que las palabras tienen sentidos de diferente naturaleza, por lo que usar una ontología unidimensional presenta un desafío, por ejemplo, los términos que nombran clases naturales pueden organizarse en jerarquías taxonómicas, mientras que los términos que se refieren a artefactos se organizan de acuerdo con su función; los términos abstractos se diferencian de los términos concretos; y las categorías gramaticales difieren entre sí.

A pesar de las dificultades que supone su construcción, las ontologías son herramientas poderosas y versátiles para la representación del conocimiento léxico. Además, los sistemas computacionales del PLN necesitan de estos repositorios léxicos (en muchos casos, multilingües) de amplio alcance y gran escala para llevar a cabo sus tareas, por lo que los trabajos en el ámbito de las ontologías están destinados a continuar.

Conclusión

En este capítulo hemos repasado los principales participantes del campo de la lexicografía informatizada actual. Se trata de un ámbito amplio donde conviven productos lexicográficos muy diversos: diccionarios tradicionales digitalizados, diccionarios tradicionales que se comienzan a reorganizar para la plataforma electrónica, lexicones diseñados en y para la plataforma de Internet y ontologías léxico conceptuales. En el grupo de los lexicones –donde se ubica FunGramKB–, destacamos a WordNet, EuroWordNet y FrameNet. Todos hacen referencia a informaciones más allá de la palabra –*synsets* en los dos

primeros, y marcos semánticos en el tercero–, pero no están vinculados a una ontología estable. En esto se distingue FunGramKB que, como veremos, responde a una visión conceptualista y se construye alrededor de una ontología preestablecida.

CAPÍTULO III. FUNCTIONAL GRAMMAR KNOWLEDGE BASE (FUNGRAMKB)

Introducción

FunGramKB, el modelo teórico y práctico que utilizamos en esta tesis, es una base de conocimiento léxico conceptual *multipropósito* y *multilingüe*³⁴. Es multipropósito, pues está diseñada para tareas de PLN como la traducción automática, los sistemas basados en diálogos (contestación y respuesta), la recuperación y la extracción de información, la categorización de textos y la minería de datos–, y multilingüe, pues al día de hoy incluye el inglés, español, francés e italiano (Mairal, 2013). Además del PLN, FunGramKB también puede servir a usuarios humanos –lingüistas, traductores o estudiantes de idiomas–, al proveer información léxica relevante a través de una interfaz basada en la web (Periñán y Arcas, 2006). El proyecto FunGramKB emblematiza el acercamiento entre dos disciplinas –la lingüística y la ingeniería del conocimiento–, que van estrechando progresivamente su colaboración y que deben continuar haciéndolo (Mairal, 2013).

FunGramKB se distingue de otros lexicones como WordNet, EuroWordNet y FrameNet principalmente porque describe formalmente el contenido cognitivo de las unidades léxicas, y ubica este contenido dentro de una taxonomía ontológica previamente diseñada. Ello responde a una metodología basada en un enfoque de semántica profunda, que justamente implica, entre otras cosas, que se definan formalmente las unidades conceptuales. De hecho, para formalizar las definiciones, FunGramKB desarrolló un lenguaje de representación conceptual (*Conceptual Representation Language*, en adelante COREL), que discutiremos más adelante en este capítulo. El desarrollo a gran escala de recursos basados en una semántica profunda, como FunGramKB, requiere más tiempo y esfuerzo que el desarrollo de recursos basados en una semántica superficial o relacional, como WordNet y EuroWordNet. Sin embargo, el poder expresivo de los primeros es más robusto y menos redundante que el de los segundos, y permite que se explote la información para más tareas de PLN.

Para darle al lector una referencia preliminar de la interfaz gráfica de esta base de conocimiento, proveemos una imagen de la pantalla del Editor de FunGramKB (todos los componentes y nociones que adelantamos aquí se discutirán más a fondo a lo largo del

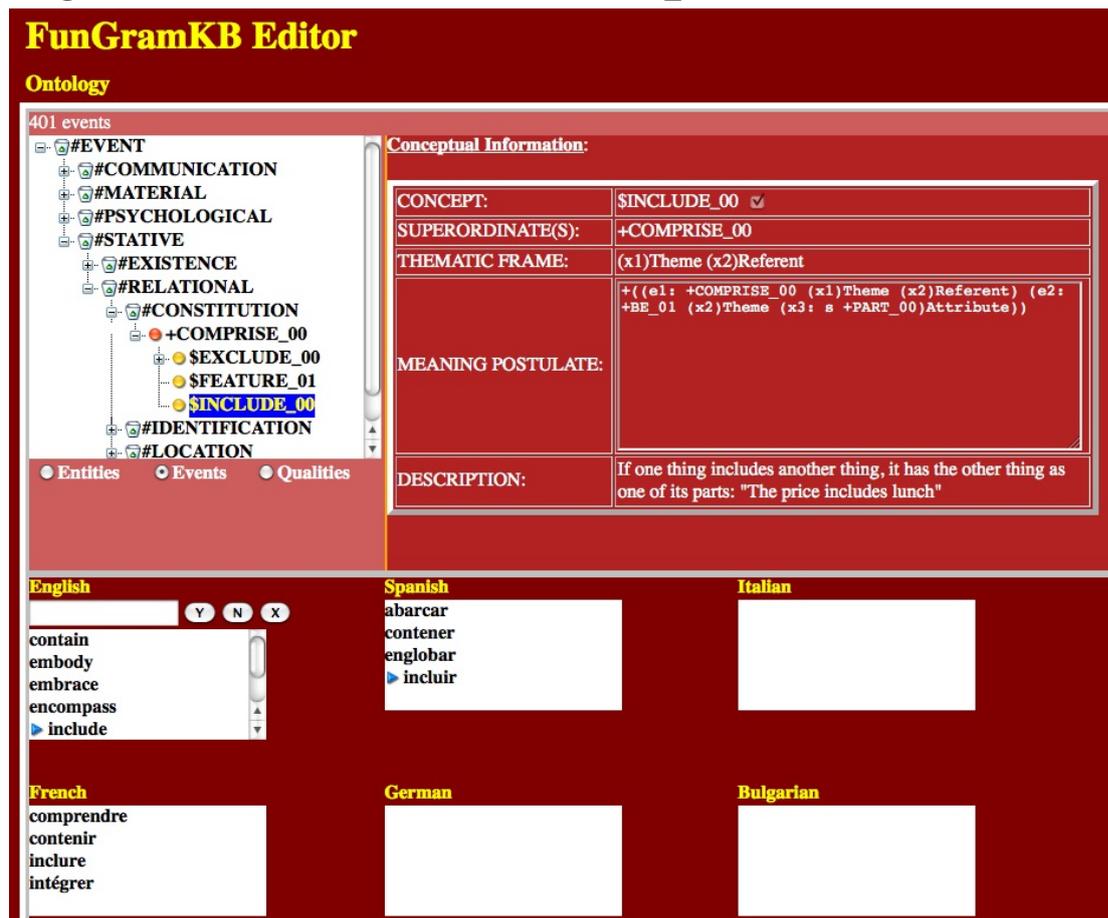
³⁴ La bibliografía de este proyecto está disponible en la página web: <http://www.fungramkb.com>, e incluye, entre otros, a Periñán y Arcas (2004, 2005, 2006, 2007ab, 2010ab), Periñán y Mairal (2009, 2010, 2011, 2012), Mairal y Periñán (2009ab, 2010), Mairal *et al.* (2012) y Mairal (2013).

capítulo). La imagen refleja el resultado de la búsqueda del concepto terminal \$INCLUDE_00 ('incluir'). En la izquierda, se muestra la ruta que conduce a \$INCLUDE_00 en la taxonomía de eventos: #EVENT → #STATIVE_00 → #RELATIONAL_00 → +COMPRISE_00 → \$INCLUDE_00. Como se trata de un evento del nivel terminal, para llegar a \$INCLUDE_00, pasamos por el nivel de metaconceptos (#EVENT, #STATIVE_00, #RELATIONAL_00) y luego por el de los conceptos básicos (+COMPRISE_00). El concepto terminal \$INCLUDE_00 constituye una restricción semántica de +COMPRISE_00, el concepto que lo precede.

A la derecha, bajo el título "Conceptual information" ('Información conceptual'), se detalla el concepto superordinado de \$INCLUDE, que es +COMPRISE_00, así como su Marco temático (MT), Postulado de significado (PS) y una descripción del significado en inglés. El MT señala que los participantes del evento \$INCLUDE_00 son un Tema y un Referente que no están especificados semánticamente. El PS expresa, en el primer enunciado (e1), que el Tema (x1) comprende al Referente (x2), y en el segundo enunciado (e2), que el Referente (x2), ahora en rol de Tema, tiene el Atributo de ser parte (x3) del Tema (x1).

Finalmente, los recuadros inferiores de la imagen muestran el vínculo del concepto \$INCLUDE_00 con los módulos léxicos de FunGramKB para el español, el inglés y el francés, y presenta sus lexicalizaciones en cada una de esas lenguas.

Imagen 9. Vista en FunGramKB de \$INCLUDE_00



En este capítulo repasaremos los aspectos fundamentales del modelo de FunGramKB. En la sección 1, repasaremos el trasfondo teórico de FunGramKB. En la sección 2, discutimos los distintos niveles de FunGramKB, y los módulos que los componen. En la sección 3, damos atención especial a la ontología de FunGramKB y a los tipos de conceptos que la pueblan. En la sección 4, discutimos las propiedades de los conceptos, específicamente, los MT y PS, además de la construcción del significado a través del lenguaje COREL. Dedicamos la sección 5 al manejo de las preferencias de selección en el nivel conceptual de FunGramKB. La sección 6 discute el manejo de las colocaciones en el nivel léxico de FunGramKB. Terminamos el capítulo con una conclusión.

1. Trásfondo teórico

FunGramKB se basa en el Modelo Léxico Construccional (MLC), un modelo de construcción de significado elaborado por el grupo de investigación Lexicom

(www.lexicom.es), que combina perspectivas funcionales y construccionistas del lenguaje (Ruiz de Mendoza y Mairal, 2008; Mairal y Ruiz de Mendoza, 2008ab, 2009³⁵).

El objetivo principal del MLC es dar cuenta de los procesos y niveles lingüísticos que participan en la construcción de significado, incluidos aquellos que van más allá de lo gramatical. Las entradas léxicas del MLC se representan como “plantillas léxicas”, una evolución de las estructuras lógicas de la Gramática del Papel y Referencia (*Role and Reference Grammar* o RRG) de Van Valin y LaPolla (1997) y Van Valin (2005). Sin embargo, el MLC ofrece sistemas de descomposición semántica más elaborados que las estructuras lógicas de la RRG –que recogen solo los aspectos que tienen visibilidad sintáctica–, dado que incluye el conocimiento semántico, pragmático y cultural que también forma parte del conocimiento que tiene un hablante de una palabra (Mairal y Perriñán, 2010)³⁶. El sistema de clases verbales se basa en las distinciones del *Aktionsart* de Vendler (1967), y el sistema descomposicional es una variante del propuesto por Dowty (1979).

Se distinguen en el MLC los siguientes niveles de construcción de significado (Ruiz de Mendoza y Mairal, 2008), que en la arquitectura de FunGramKB se traducirán en diferentes componentes:

- el Nivel 1, o estrato argumental: se ocupa de propiedades gramaticales nucleares de las unidades léxicas;
- el Nivel 2, o nivel pragmático: es el estrato de implicatura conversacional, que trabaja con el significado inferido de situaciones de nivel bajo (o escenarios específicos), tradicionalmente manejado en la pragmática mediante la teoría de implicatura;
- el Nivel 3, o estrato ilocutivo: se ocupa del significado ilocutivo convencionalizado, en situaciones de alto nivel (o escenarios genéricos);
- el Nivel 4, o estrato discursivo: captura los mecanismos lingüísticos basados en estructuras discursivas, con particular énfasis en la cohesión y coherencia.

Según Butler (2009), el MLC combina exitosamente los acercamientos funcionalistas, especialmente los de orientación proyeccionista, con los acercamientos cognitivos,

³⁵ Además de estas obras fundacionales, la página electrónica www.lexicom.es presenta el resto de la bibliografía sobre el MLC –incluidos los más recientes avances–, organizada en los siguientes temas: Semántica ontológica y construcciones, Ingeniería del conocimiento: FunGramKB, Aplicaciones: manejo y traducción terminológica, y Modelaje cognitivo, además de otros escritos sobre el modelo en general.

³⁶ Entre los trabajos más recientes que abordan la relación entre la RRG, el MLC y FunGramKB, figuran las publicaciones que mencionamos a continuación. Van Valin y Mairal (en prensa) abordan la interfaz léxico ontológica mediante un algoritmo de enlace. Mairal *et al.* (2012) tratan el funcionalismo en la teoría lingüística, con énfasis particular en la RRG; en ese volumen, Mairal *et al.* (2012) hablan de la representación léxica con enfoque ontológico, mientras que Perriñán y Mairal (2012) discuten la dimensión computacional de la RRG. Finalmente, Ruiz de Mendoza y Mairal (2011) discuten las restricciones en la alternancia sintáctica dentro del MLC.

principalmente construccionistas. Recordemos que los enfoques proyeccionistas mantienen que las propiedades sintácticas de un predicado se pueden predecir a partir de su significado, mientras que los construccionistas defienden que lo que determina el significado oracional no es solo el significado de un predicado, sino también la aportación de las construcciones en que participa. Según este autor, el MLC logra conciliar lo mejor de los dos mundos: el compromiso de los modelos funcionales proyeccionistas con una semántica léxica de granularidad fina y el compromiso de los acercamientos construccionistas más “macroscópicos” con la adecuación cognitiva³⁷. Butler propone que a esta fusión se le llame “lingüística funcional-cognitiva” (2009:119).

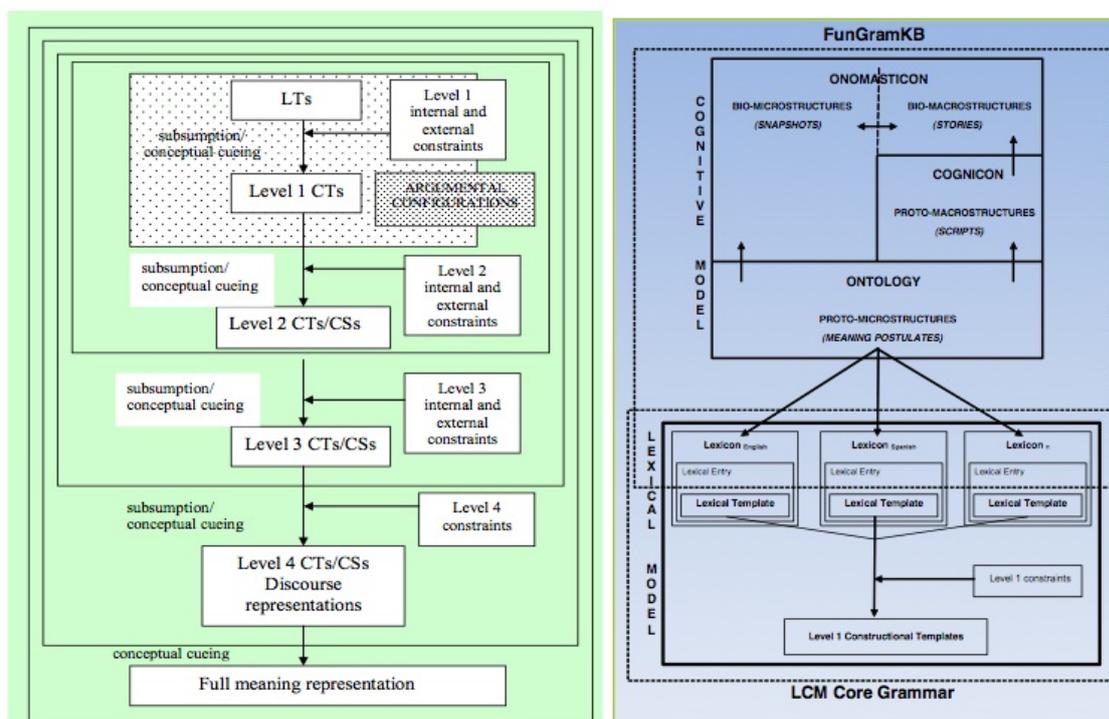
El MLC también tiene el propósito de crear descripciones que se puedan implementar computacionalmente en bases de conocimiento orientadas al PLN, en particular, en proyectos de representación del conocimiento. Según Mairal (2013: 2):

En esencia, el MLC nos ofrece las herramientas analíticas necesarias para un análisis detallado de las dimensiones proposicional y pragmático-discursiva del significado, lo que permite formular una interpretación semántica plenamente especificada de un texto, que bien podría ser el texto de entrada de un ordenador, o un robot o un sistema inteligente.

Con esos fines, el MLC expandió su arquitectura para acomodar a FunGramKB, y tradujo los cuatro niveles del MLC descritos arriba en componentes de la base de conocimiento. Se integró el MLC como parte del nivel lingüístico de esta base de conocimiento. El resultado es un modelo conceptual que consta de dos niveles de representación bien diferenciados –un nivel cognitivo (*Cognitive Model*) y un nivel lingüístico (*Lexical Model*)–, pero bien relacionados, como se muestra en la imagen siguiente:

³⁷ Véase Luzondo (2011) para una discusión a fondo sobre la convergencia exitosa en el MLC de aspectos relevantes de las teorías de Goldberg (1995, 2006) y de Boas (2003, 2008ab), típicamente vistas como teorías divergentes en el marco de las Gramáticas de Construcción.

Imagen 10. Representación de la arquitectura del MLC y su inserción en FunGramKB (tomado de Mairal y Periñán, 2010)



La inclusión de una base de conocimiento como FunGramKB en el MLC provoca un cambio profundo: el modelo ya no comienza en el componente léxico sino que comienza en el nivel conceptual; por consiguiente, “el componente léxico deja de ser el motor de arranque de la maquinaria lingüística para ser recipiendario de todo el caudal de información que le aporta el nivel conceptual y, más en particular, la ontología” (Mairal y Periñán, 2010:5). La imagen anterior ilustra este giro cognitivo: a la derecha vemos cómo el nivel conceptual alimenta a los lexicones de las diversas lenguas. Esto quiere decir que el peso de la descripción semántica recae en la ontología, mientras que las entradas léxicas quedan más simplificadas, aunque –como veremos– también contarán con una cantidad sustancial de información lingüística. La práctica le ha demostrado al MLC que el acercamiento conceptual puede tener mayor capacidad explicativa que los modelos lexicistas para determinar el complejo proceso de la construcción del significado (Mairal *et al.*, 2011; Mairal, 2013).

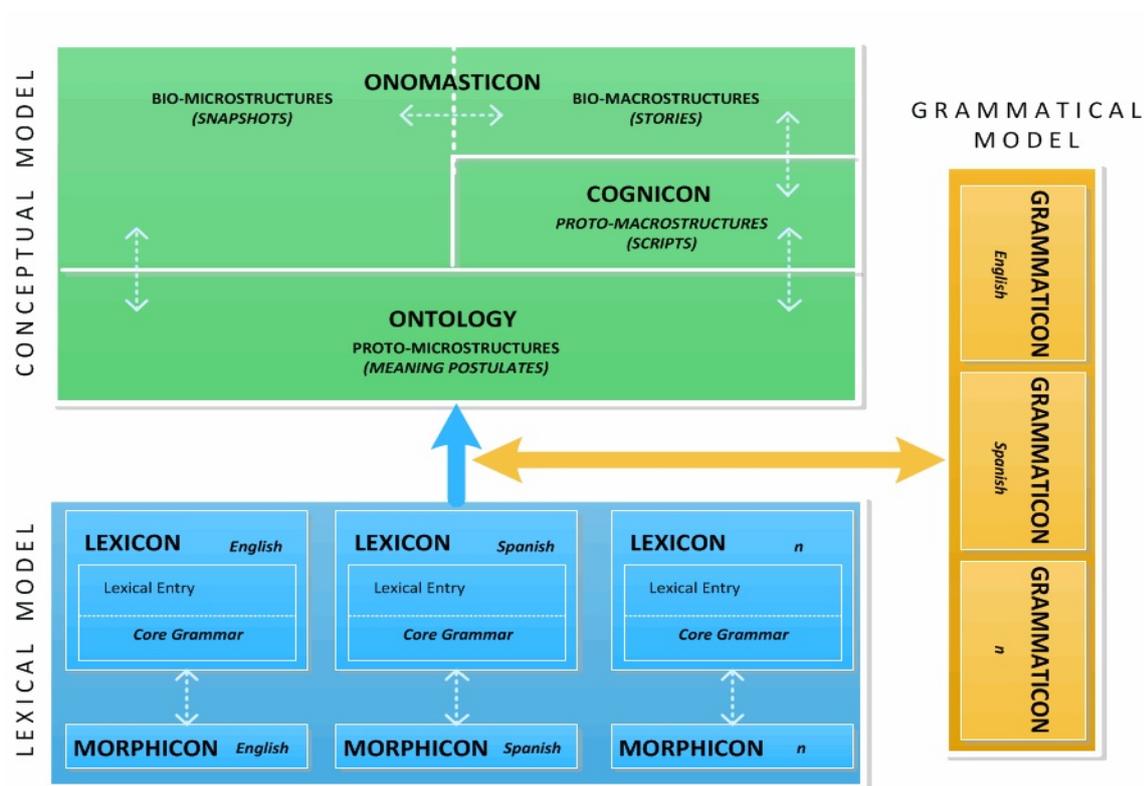
2. Niveles de FunGramKB

Como hemos visto antes, FunGramKB separa el conocimiento cognitivo o universal, que es común a todas las lenguas, del conocimiento lingüístico, que es dependiente de cada lengua. En su versión actual, estos dos tipos de conocimiento se reparten en tres niveles

principales: conceptual, léxico y gramatical, llamados “modelos” en inglés: *Conceptual Model*, *Lexical Model* y *Grammatical Model*. El conocimiento cognitivo o universal está contenido en el nivel conceptual, mientras que el conocimiento lingüístico está repartido en los niveles léxico y gramatical. En el nivel conceptual se encuentra la ontología, que constituye el pivote para toda la arquitectura. Esto quiere decir que los ingenieros del conocimiento solo tienen que construir un nivel conceptual para procesar cognitivamente el insumo de cualquier idioma, mientras que los lexicógrafos computacionales tienen que desarrollar niveles léxicos y gramaticales para el inglés, el español y cualquier otra lengua que se incluya (Periñán y Arcas, 2006). Cada uno de los tres niveles se compone de módulos independientes pero interrelacionados.

La relación entre los tres niveles y los módulos que los componen se ilustra en la siguiente imagen, que también nos permite observar cómo el MLC y, por lo tanto, FunGramKB, intentan reconciliar los enfoques proyeccionistas y los construccionistas sobre el idioma. Así, el nivel léxico (lexicón y morfocón) recoge las propiedades de las entradas léxicas, y el nivel gramatical (gramaticón) se encarga de los esquemas construccionales. Ambos niveles se vinculan al nivel conceptual, que incluye a su vez el onomasticón, el cognocón y la ontología. Cada uno de los niveles del modelo se discute en los apartados que siguen.

Imagen 11. Arquitectura completa de FunGramKB (fungramkb.com)³⁸



2.1 Nivel léxico³⁹

Como se aprecia en la imagen 10, el nivel léxico de FunGramKB –que es particular de cada lengua– incluye los módulos lexicón y morficon. En el lexicón, FunGramKB almacena la información morfosintáctica, pragmática y colocacional de las unidades léxicas, mientras que el morficon maneja los casos de morfología flexiva.

Para lograr un modelo léxico robusto para FunGramKB, se tomaron como punto de partida los estándares lexicográficos de los modelos de EAGLES⁴⁰ y OLIF⁴¹. Sin embargo,

³⁸ De los componentes de FunGramKB, esta tesis utiliza principalmente la ontología y ciertas partes del nivel léxico.

³⁹ El nivel léxico está descrito exhaustivamente en Mairal y Perrián (2009a).

⁴⁰ EAGLES (*Expert Advisory Group on Language Engineering Standards*) es una iniciativa auspiciada por la Comisión Europea que busca proveer recomendaciones para la estandarización de las tecnologías del lenguaje. Más específicamente, el *Computational Lexicons Interest Group* está a cargo de analizar las principales prácticas de la codificación lexicográfica mediante comparar los recursos léxicos disponibles en las lenguas europeas como el catalán, danés, holandés, inglés francés, alemán, griego, irlandés, italiano, portugués, español y sueco.

la plantilla léxica de FunGramKB fue evolucionando para incluir información de tres tipos:

(i) información morfosintáctica, por ejemplo, las variantes gráficas, las abreviaturas, el tipo de sintagmas e información categorial;

(ii) información gramatical (*core grammar*): tipo de estructura *Aktionsart*, número de variables, asignación de marca de macro-papel o macro-rol⁴², etiquetado de cada una de las variables con las funciones temáticas, colocaciones, tipo de construcciones sintácticas;

(iii) información miscelánea: tipo de dialecto, estilo, rasgos de un dominio discursivo, etc⁴³.

Tomamos de Mairal (2013) el siguiente esquema, que ilustra la información gramatical que incluirá la entrada léxica del predicado *romper*:

Tabla 13. Representación simplificada de romper (tomada de Mairal, 2013)

| LEXICÓN: Español | |
|------------------------|---|
| Unidad léxica: | <i>romper</i> |
| Concepto: | +BREAK_00 |
| Aktionsart: | Realización causativa |
| Variables: | x, y |
| Macro-papel: | x = Actor y = Undergoer |
| Enlace Marco temático: | x=Theme y= Referent |
| Construcción: | Resultativa Incoativa Media, etc. |

La imagen siguiente, por otro lado, ofrece una representación parcial del componente léxico de FunGramKB en español, tal cual se ve en la pantalla (cabe recordar aquí que toda unidad léxica está vinculada a un concepto en la ontología). En esta tesis, nos va a interesar en particular la sección de la plantilla léxica dedicada a las Colocaciones, en específico en la posición Y:

⁴¹ OLIF (*Open Lexicon Interchange Format*) es un formato de intercambio léxico y terminológico que está en cumplimiento con XML y fue diseñado específicamente para usuarios de herramientas de tecnologías del lenguaje.

⁴² La tipología de *Aktionsart* y los macropapeles como sistema intermedio de enlace entre la sintaxis y la semántica son nociones teóricas tomadas de la RRG que se adoptan en el formato de las entradas. La RRG postula dos macropapeles: Actor y Padecedor (*Actor* y *Undergoer* respectivamente), que corresponden a las nociones tradicionales de ‘sujeto lógico’ y ‘objeto lógico’ (Van Valin, 2005).

⁴³ Véase Mairal y Perrián (2009) para una descripción detallada de cada uno de estos parámetros.

Imagen 12. Representación parcial del nivel léxico en español

The image shows a software interface titled "LCM CORE GRAMMAR". It is divided into two main vertical sections. The top section is labeled "AktionsArt:" and contains a large empty rectangular box. Below this box is the text: "You determine the canonical lexical class(es) of the verb." The bottom section is labeled "Lexical Template:" and contains several sub-sections: "Variables:" with a text input field; "Idiosyncratic features:" with a field "[MR :], [U :]"; "Thematic frame mapping:" with fields "X = : , Y = : , Z = :"; "A REMINDER OF FUNGRAMKB PARTICIPANTS:" with a small icon; "Prepositions:" with fields "X = : Y", "Y = : Y", and "Z = : Y", each with a small circle containing "N" below it; and "Collocations:" with fields "X = Y", "Y = Y", and "Z = Y", each with a small circle containing "N" below it. An orange arrow points from the left towards the "Prepositions:" section.

2.2 Nivel gramatical

En el nivel gramatical, o Gramaticón, se almacenan esquemas de construcción basados en la RRG y en las Gramáticas de Construcción (Goldberg, 1995, 2006 y Boas, 2003, 2008ab son las fuentes que guardan una relación más estrecha con el MLC⁴⁴). El Gramaticón se compone de cuatro módulos inspirados en los cuatro niveles de construcción de significado formulados en el MLC y descritos arriba: argumental, implicacional, ilocutivo y discursivo. La siguiente imagen presenta la puerta de entrada a estos cuatro niveles en FunGramKB; no nos detendremos en ellos aquí, pues no los utilizamos directamente en esta tesis.

⁴⁴ Véase Luzondo (2011) para un estado de la cuestión exhaustivo sobre la Gramática de Construcción, en particular de la aportación de esos dos autores.

Imagen 13. Representación de los cuatro niveles de construcción del gramaticón



Nuevamente, cabe recordar que el componente gramatical también está vinculado a la ontología, y hereda información de ella. Este vínculo entre el nivel lingüístico (léxico y gramatical) y el nivel cognitivo (conceptual) representa el centro de la interfaz lingüístico-conceptual⁴⁵.

2.3 Nivel conceptual

El nivel conceptual almacena el conocimiento de sentido común que forma parte del razonamiento humano. En este nivel se distinguen tres grandes módulos de representación del conocimiento: el cognición, el onomasticón y la ontología. Estos tres módulos reflejan tres tipos de conocimiento que se distinguen en la psicología cognitiva, que están, a su vez, inspirados por las distinciones establecidas por el modelo de memoria a largo plazo de Tulving (1985): el cognición contiene el conocimiento procedimental, el onomasticón se ocupa del conocimiento episódico y la ontología incluye el conocimiento semántico (Mairal y Perrián, 2010):

(a) El cognición (o conocimiento procedimental) almacena el conocimiento sobre procedimientos, por ejemplo, cómo freír un huevo o cómo comprar un producto. Este

⁴⁵ Véase Van Valin y Mairal (en prensa) para una discusión detallada sobre la interfaz léxico ontológica y su algoritmo de enlace.

conocimiento se codifica por medio de guiones, que son esquemas conceptuales que organizan una secuencia de acciones estereotípicas con continuidad temporal, por ejemplo, ‘comer en un restaurante’ o ‘ir de compras’, etc. Para el diseño de los guiones, se utiliza el modelo temporal de Allen (1983)⁴⁶.

(b) El onomasticón (o conocimiento episódico) guarda información sobre instancias de eventos y entidades, tales como Borges, Bill Gates, el Taj Majal o el 9/11. Este módulo almacena dos tipos diferentes de esquemas (“fotos” y “cuentos”), dado que las instancias pueden ser representadas sincrónica o diacrónicamente. En este componente, los datos se introducen por medio de un sistema semiautomático que usa reglas de “mapeo” (*mapping*) que importan datos de recursos como DBPEDIA (www.dbpedia.org)⁴⁷.

(c) La ontología (o conocimiento semántico) es un catálogo jerárquico de los conceptos que describen el conocimiento semántico de una persona cuando habla de situaciones cotidianas.

La ontología ocupa el rol central no solo en el nivel conceptual, sino en toda la arquitectura de FunGramKB. Esto es consistente con señalamientos hechos desde temprano por Mairal y Faber (2002) y Faber y Mairal (2003) sobre la importancia de enlazar los elementos de las plantillas léxicas a una ontología que permita un acercamiento menos influido por nociones culturales. Dedicamos el apartado siguiente a discutir con más detalle la ontología de FunGramKB.

3. La ontología

La ontología es posiblemente el aspecto más novedoso de FunGramKB. FunGramKB busca proveer representaciones conceptuales que reflejen la estructura del sistema cognitivo de los seres humanos. Por lo tanto, está compuesta de conceptos universales, lo que significa que cada concepto que imaginamos tiene, o puede tener, un lugar en ella. La ontología de FunGramKB está fundamentada en una metodología sólida, basada en principios implementados en otros proyectos de PLN, como Bouad *et al.* (1995), Mahesh (1996) y Noy y McGuinness (2001).

La ontología se organiza como una jerarquía conceptual IS-A que permite la herencia múltiple no monotónica. En la herencia monotónica, el concepto subordinado hereda todos los atributos de su concepto superordinado, sin ninguna excepción. Por el contrario, la

⁴⁶ El trabajo de Perrián y Carrión (2011) ofrece una descripción más detallada del onomasticón.

⁴⁷ Los trabajos de Garrido y Ruiz de Mendoza (2011) y Perrián (2012) ofrecen descripciones detalladas del cognición.

herencia múltiple no monotónica permite que un concepto tenga más de un superordinado y que la información genérica de los superordinados pueda ser rebatida por la información más específica de los conceptos subordinados. Mairal (2012) cita un problema típico que se estudia en la IA: los pájaros vuelan, y el pingüino es un pájaro, pero el pingüino no vuela. La herencia no monotónica es muy propia de las lenguas naturales, por lo cual es preciso que una ontología pueda tratar estos casos sin que exista redundancia de información ni tampoco una explosión desmedida de excepciones⁴⁸.

Aunque la ontología de FunGramKB no es dependiente de ninguna lengua, sí está lingüísticamente motivada. Esto quiere decir que solo se puede introducir un nuevo concepto cuando haya por lo menos una unidad léxica cuyo significado no corresponda a ninguno de los postulados de significado existentes. Es necesario, sin embargo, que las propiedades ontológicas de ese nuevo concepto sean compartidas por todas las unidades léxicas vinculadas con él (Mairal y Perrián, 2009a).

El significado de las palabras se refleja en el vínculo con una unidad conceptual. Por lo tanto, en FunGramKB la polisemia se maneja vinculando unidades léxicas a diferentes conceptos en la ontología. Así, la palabra *country*, que puede significar ‘país’ y ‘campo’, tiene dos entradas léxicas diferentes en FunGramKB, cada una vinculada a un concepto diferente: *country*[+COUNTRY_00] y *country*[+COUNTRYSIDE_00], respectivamente. Esta identificación permite conectar palabras con sentidos a través de las lenguas y en la misma lengua sin ninguna ambigüedad. Por ejemplo, las unidades léxicas *country* y *state* se relacionan a través del concepto +COUNTRY_00, y *country* y *campiña* a través del concepto +COUNTRYSIDE_00.

El proceso para crear la ontología ha integrado acercamientos de “arriba hacia abajo” y de “abajo hacia arriba”, en tres fases. Esas tres fases se traducen en los tres niveles conceptuales de la ontología, que a su vez se componen de tres tipos de conceptos diferentes: en el nivel más alto de la taxonomía o nivel *metaconceptual* se ubican los *metaconceptos*; en el nivel medio o nivel *básico* se ubican los conceptos *básicos*; y en el nivel más bajo o nivel *terminal* se ubican los conceptos *terminales*. Finalmente, también existen los *subconceptos*, que no son conceptos como tal, como veremos. Cada uno de los niveles representa además un nivel diferente de granularidad semántica: los metaconceptos tienen la granularidad más gruesa o genérica, y los conceptos terminales, la más fina o particular. Esto responde al propósito de construir una base de conocimiento que simule el conocimiento humano. A

⁴⁸ Perrián y Arcas (2010a) describen con más detalle el tratamiento de la herencia no monotónica en la ontología de FunGramKB.

continuación describimos con más detalle los conceptos que pueblan cada nivel de la ontología.

3.1 Los metaconceptos

El nivel *metaconceptual* se crea en la primera fase de trabajo, con el propósito de tener una visión compartida del mundo que le dé forma al nivel superior de la ontología. Los *metaconceptos*, por lo tanto, constituyen el nivel más alto de la taxonomía. Representan dimensiones cognitivas, por lo cual constituyen “categorías escondidas”, que son conceptos que no están vinculados a ninguna unidad léxica y no tienen postulado de significado (Periñán y Arcas, 2007a), es decir, una definición formal. Los *metaconceptos* están precedidos por un signo de #, y se dividen en tres subontologías: #ENTITIES o ‘entidades’, #EVENTS o ‘eventos’ y #QUALITIES o ‘cualidades’, que se corresponden en el nivel léxico con tres categorías gramaticales: nombres, verbos y adjetivos, respectivamente.

Los 42 metaconceptos de la ontología se establecieron tras analizar el nivel superior de las principales ontologías lingüísticas, con la meta de tener un diseño que contribuyera a la integración e intercambio con otras ontologías: DOLCE (Gangemi *et al.*, 2002), *Generalized Upper Model* (Bateman, Henschel y Rinaldi, 1995), Mikrokosmos (Mahesh, 1996; Mahesh y Nirenburg, 1995), SIMPLE (Lenci *et al.*, 2000) y SUMO (Niles y Pease, 2001). También se consideró el *Natural Semantic Metalanguage* (NSM) de Wierzbicka, pero las explicaciones de sentido del NSM no resultaron conducentes a una representación concisa.

La próxima imagen, tomada de Mairal (2013), muestra una parte de la subontología de #EVENTS y las algunas dimensiones cognitivas del nivel metaconceptual.

Imagen 14. Vista parcial de las unidades metaconceptuales de la subontología #EVENT

2. #EVENT

- 2.1 #COMMUNICATION
- 2.2 #MATERIAL
 - 2.2.1 #CREATION
 - 2.2.2 #MOTION
 - 2.2.2.1 #TRANSFER
 - 2.2.3 #TRANSFORMATION
- 2.3 #PSYCHOLOGICAL
 - 2.3.1 #COGNITION
 - 2.3.2 #EMOTION
 - 2.3.3 #INTENTION
 - 2.3.4 #PERCEPTION
- 2.4 #STATIVE
 - 2.4.1 #EXISTENCE
 - 2.4.2 #RELATIONAL
 - 2.4.2.1 #CONSTITUTION
 - 2.4.2.2 #IDENTIFICATION
 - 2.4.2.3 #LOCATION
 - 2.4.2.4 #POSSESSION

Tómese en cuenta que cada metaconcepto define de manera particular los participantes que componen su MT (veáse Apéndice 1). Los participantes particulares serán heredados por todos los conceptos subordinados a esa dimensión, por lo que cada metaconcepto determina rasgos importantes que estarán presentes en su “decendencia” conceptual. Por ejemplo, en el metaconcepto #COMMUNICATION, el participante Tema se define como ‘la entidad que transmite un mensaje’, mientras que el Referente es ‘el mensaje, o conjunto de proposiciones, que es transmitido’ y la Meta es ‘la entidad que recibe el mensaje’. Los tres participantes son obligatorios. Así, el concepto básico +SAY_00, ubicado en la dimensión de #COMMUNICATION, tiene un MT que lee (x1:+HUMAN_00)Theme (x2)Referent (x3:+HUMAN_00)Goal. Esto quiere decir que el Tema, que es la entidad que transmite el mensaje, está restringido semánticamente a ‘humano’; el Referente, que es el mensaje en sí, no está especificado semánticamente; y la Meta, que es la entidad que recibe el mensaje, también está restringido a un ‘humano’. Por lo tanto, en el concepto +SAY_00, sabemos que participan dos ‘humanos’ y un mensaje.

Un segundo ejemplo lo provee el metaconcepto “EMOTION”. Aquí el participante Agente se define como ‘la entidad que hace que otra entidad sienta una emoción’, mientras que el Tema es ‘la entidad que siente la emoción’; esta dimensión puede incluir también un

participante Atributo, que sería ‘la entidad o cualidad que describe un atributo de una entidad cuando está sintiendo una emoción’. El concepto +FORGIVE_00, subordinado de #EMOTION, tiene el siguiente MT: (x1:+HUMAN_00)Agent (x2:+HUMAN_00)Theme. Esto quiere decir que el Agente, que hace que otra entidad sienta emoción, está restringido a ‘humano’, mientras que el Tema, que es la entidad que siente la emoción, es también ‘humano’. El Atributo opcional no está presente en este concepto. Los participantes de +FORGIVE_00 son dos ‘humanos’.

Cabe señalar, además, que para cada dimensión metaconceptual, FunGramKB ha desarrollado un inventario de descriptores inspirado en la Estructura de *Qualia* Extendida, que es utilizada por el proyecto SIMPLE (Lenci, 2000; Lenci *et al.*, 2000, SIMPLE Specification Group, 2000). SIMPLE busca extender o ampliar los valores de *Qualia* establecidos por Pustejovsky (1995). Siguiendo el ejemplo de Mairal (2012), el concepto terminal \$TOASTER_00, que pertenece a la dimensión conceptual #SELF_CONNECTED_OBJECT_00, contiene 17 descriptores que ofrecen información sobre el tipo de rasgos ejemplares que son relevantes en esta dimensión, y que pueden o no formar parte del PS del concepto en cuestión. Específicamente, los descriptores abordan preguntas sobre las propiedades formales y constitutivas de la entidad (su olor, su sabor, si tiene partes, etc.), las propiedades télicas (su propósito o función), las propiedades agentivas (de quién se obtiene y cómo). Se trata, pues, de precisiones a los *Qualia* de Pustejovsky.

3.2 Los conceptos básicos

En la segunda fase, los conceptos más descriptivos se organizan jerárquicamente en un nivel medio, o nivel *básico*, usando la relación de subsunción IS-A. Los *conceptos básicos* están precedidos por un signo de +, por ejemplo: +BOOK_00, +HAND_00, +MOVE_00. El punto de partida para identificar los conceptos básicos fue el vocabulario definitorio del *Longman Dictionary of Contemporary English*, aunque se sometió a revisión para reducirlo a 1,300 conceptos⁴⁹ por medio de una metodología que consta de cuatro pasos: conceptualización, jerarquización, remodelación y refinamiento⁵⁰.

Los conceptos básicos sirven en FunGramKB como unidades definitorias: se usan en la construcción de postulados de significado para los mismos conceptos básicos y los

⁴⁹ Este nivel ha sido validado con el vocabulario definitorio de diccionarios de otras lenguas, como el *Diccionario para la Enseñanza de la Lengua Española* (VOX-Universidad de Alcalá de Henares, 1995), para el español.

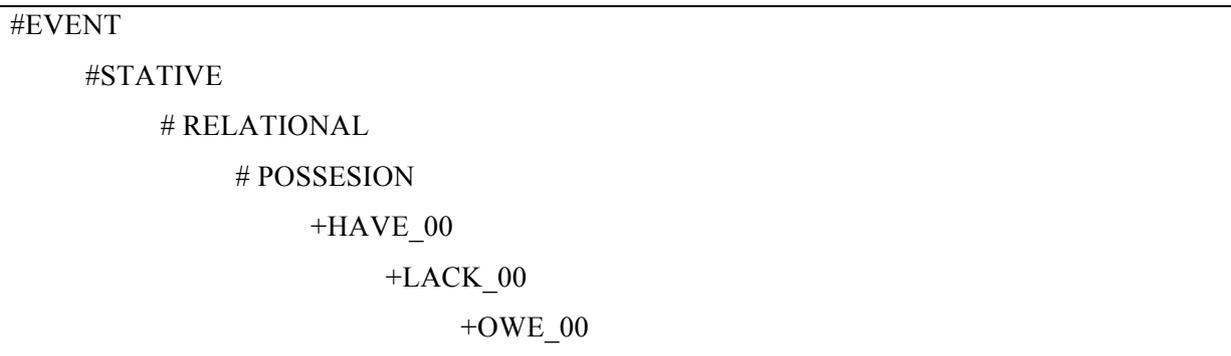
⁵⁰ Perrián y Mairal (2011) presentan una descripción más detallada de esta metodología, que en inglés recibe el nombre de COHERENT: Conceptualization + HiErarchization + REmodelling + RefinemeNT.

conceptos terminales. Son también las unidades que especifican las preferencias de selección de un concepto.

Cabe señalar que el nivel *básico* es el nivel central del conocimiento en la arquitectura de esta ontología: responde a la necesidad de tener un pivote entre las categorías universales que facilitan la interoperabilidad ontológica (los *metaconceptos*) y aquellos conceptos particulares que otorgan aplicabilidad inmediata (los *terminales*), que se discuten a continuación.

Partiendo de la imagen anterior, la imagen que sigue presenta la ruta conceptual del metaconcepto #EVENT hasta el concepto básico +OWE_00, cuyas lexicalizaciones son *adeudar* y *deber*, en español, y *be in debt* y *owe*, en inglés:

Imagen 15. Ruta conceptual del metaconcepto #EVENT al concepto básico +OWE_00



Para describir el concepto básico +OWE_00, FunGramKB establece su Marco temático, que dice que los participantes de este concepto básico son un Tema que es ‘humano’ (x1), y un Referente que es ‘dinero’ (x2). Nótese que estos participantes son conceptos básicos también⁵¹:

Tabla 14. Marco temático de +OWE_00

| |
|-------------------------|
| (x1: +HUMAN_00)Theme |
| (x2: +MONEY_00)Referent |

La definición del concepto básico +OWE_00 está contenida en su Postulado de significado, que en términos simplificados expresa ‘estar en deuda’. Nótese aquí también que las unidades que configuran la definición son todos conceptos básicos:

⁵¹ Aunque ocasionalmente usemos el lenguaje natural para nombrar estos participantes (las palabras “humano” y “dinero”), es imprescindible recordar que se trata de conceptos básicos, que están formalmente definidos en FunGramKB.

Tabla 15. Postulado de significado de +OWE_00

| |
|---|
| <p>+(e1: +LACK_00 (x3: +HUMAN_00)Theme (x2)Referent (f1: (e2: past +LEND_00 (x3)Agent (x2)Theme (x3)Origin (x1)Goal))Reason ^ (f2: (e3: past +SELL_00 (x3)Agent (x4)Theme (x3)Origin (x1)Goal (f3: x2)Instrument))Reason)</p> |
|---|

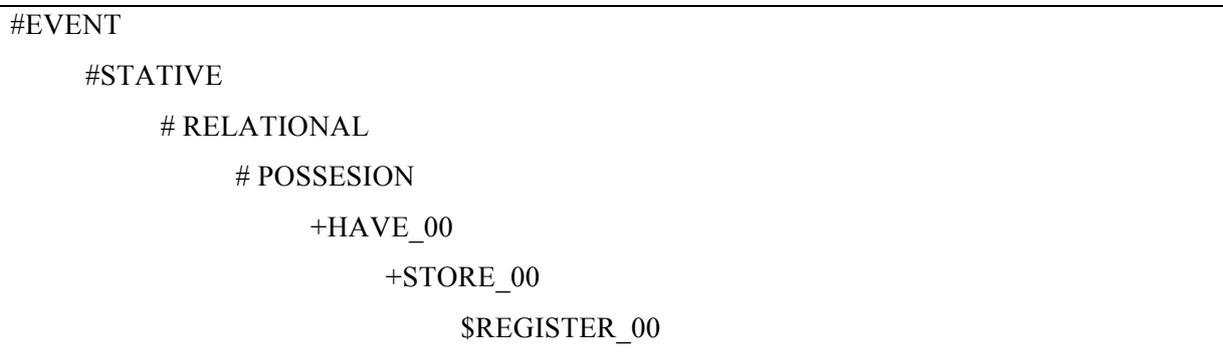
Las nociones de Marco temático y Postulado de significado se discutirán con más detalle más adelante.

3.3 Los conceptos terminales

En la tercera fase, los conceptos más específicos se incorporan en la ontología, en el nivel *terminal*. Los *conceptos terminales* están precedidos por el signo de \$, por ejemplo: \$AUCTION_00, \$CADAVEROUS_00, \$VARNISH_00. La frontera entre los conceptos básicos y los terminales está basada en el potencial de los primeros para tomar parte como elementos definitorios en los postulados de significado. Según continúa el trabajo, se siguen añadiendo conceptos terminales, pero también pueden ascender conceptos terminales al nivel básico o descender conceptos básicos al nivel terminal.

La imagen siguiente presenta la ruta conceptual del metaconcepto #EVENT hasta el concepto terminal \$REGISTER_00, cuyas lexicalizaciones son *grabar* y *registrar*, en español, y *record* y *register*, en inglés:

Imagen 16. Ruta conceptual del metaconcepto #EVENT al concepto terminal \$REGISTER_00



Para el concepto terminal \$REGISTER_00, FunGramKB establece un Marco temático compuesto por un participante Tema que es ‘humano’ (x1), y un Referente que es un ‘objeto de información’ (x2).

Tabla 16. Marco temático de \$REGISTER_00

| |
|--------------------------------------|
| (x1: +HUMAN_00)Theme |
| (x2: +INFORMATION_OBJECT_00)Referent |

La definición del concepto terminal \$OWE_00 está se expresa en su Postulado de significado, que en términos simplificados expresa ‘grabar por escrito’, ‘integrar en un registro’:

Tabla 17. Postulado de significado de \$REGISTER_00

| |
|--|
| +(e1: +STORE_00 (x1)Theme (x2)Referent (f1: +ARTEFACT_00)Instrument (f2:(e2: fut +PERCEIVE_00 (x1)Theme (x2)Referent))Purpose) |
|--|

Nótese que las unidades que se utilizan para definir los conceptos terminales son los mismos conceptos básicos.

3.4 Los subconceptos

Hay un cuarto tipo de unidad en la ontología de FunGramKB, llamado *subconcepto*. Los subconceptos no son visibles en la ontología, porque no se consideran conceptos como tal. Se trata de casos en que la restricción o especificación conceptual ocurre exclusivamente dentro del MT de un concepto básico o terminal, pero no hay cambio en el PS. Esto quiere decir que una unidad conceptual mantiene su significado, codificado conceptualmente como PS, pero cambia alguna preferencias de selección en los participantes del MT. Los subconceptos resultan importantes para cualquier trabajo enfocado en restricciones de selección. Los subconceptos van precedidos por un guion, por ejemplo: -WIELD, -MISPLACE_00 y -SAVE_00.

Para ilustrar lo anterior, veamos los ejemplos provistos por Jiménez y Pérez (2011). El subconcepto -MISPLACE_00, lexicaliza en español como *traspapelar*, es subordinado de +LOSE_00. El concepto +LOSE_00 tiene un MT con una Tema ‘humano’ o ‘animal’ y un Referente ‘corpúscular’, y un PS que expresa que el Tema carece del Referente porque lo colocó en un sitio por una razón y no lo recuerda:

Tabla 18. MT y PS de +LOSE_00

| | |
|-----|---|
| MT: | (x1: +HUMAN_00^+ANIMAL_00)Theme (x2: +CORPUSCULAR_00)Referent |
| PS: | +(e1: +LACK_00 (x1)Theme (x2)Referent (f1: (e2: past +PUT_00 (x1)Agent (x2)Theme (x3)Origin (x4)Goal))Reason (f2: (e3: n +REMEMBER_00 (x1)Agent (x1)Theme (x4)Referent))Reason) |

El subconcepto -MISPLACE_00 se ubica bajo +LOSE_00 y restringe semánticamente el Referente a un documento (+DOCUMENT_00), porque el verbo *traspapelar* requiere combinarse con ‘documentos’.

El subconcepto -WIELD_00, por su parte, se lexicaliza en español como *empuñar*. Es subordinado del concepto +GRASP_00, que tiene un MT con un Tema ‘humano’ y un Referente ‘corpúscular’, y un PS que expresa que el Tema agarra el Referente ‘con fuerza’:

Tabla 19. MT y PS de +GRASP_00

| | |
|-----|--|
| MT: | x1: +HUMAN_00)Theme (x2: +CORPUSCULAR_00)Referent |
| PS: | +(e1: +HOLD_00 (x1)Theme (x2)Referent (f1: +TIGHT_00)Manner) |

El subconcepto -WIELD_00 se ubica bajo +GRASP_00 y restringe semánticamente el Referente a un arma (+WEAPON_00), algo requerido por el verbo *empuñar*, que solo acepta combinarse con ‘armas’.

La importancia del espacio que proveen los subconceptos para ubicar las restricciones de selección se continúa discutiendo en la sección 5 (“Las preferencias de selección”), más adelante.

4. Propiedades de los conceptos

Los conceptos básicos y terminales no son símbolos atómicos, sino que tienen propiedades conceptuales en la forma de MT y PS. Los MT y los PS son esquemas conceptuales, porque no emplean palabras, sino conceptos, como bloques de construcción para la descripción formal del significado.

El acercamiento de definición formal de FunGramKB lo distingue de WordNet y EuroWordNet. Como vimos en el capítulo anterior, estos dos lexicones adoptan un sistema relacional, que representa el significado a través de las asociaciones semánticas (o *synsets*) que establece una palabra con otras. El *synset* captura de manera implícita un sentido compartido por las palabras que lo componen. Aunque los *synsets* van acompañados de glosas definitorias, estas no son tratables por la computadora, por lo que el *synset* constituye la unidad mínima de sentido en WordNet y EuroWordNet. Por su parte, FrameNet vincula las unidades léxicas a marcos semánticos que representan la relación entre las estructuras semánticas y sintácticas de las palabras. Según se estudian las lenguas, se añaden marcos semánticos a FrameNet. Los FE o elementos que componen los marcos no necesariamente pertenecen a un inventario preestablecido, sino que se van creando según lo requiera cada

marco. Todo esto mantiene abierto el universo de marcos semánticos y de FE posibles en FrameNet. Así, aunque tanto los *synsets* de WordNet como los marcos semánticos de FrameNet constituyen nociones semánticas que van más allá de la palabra, FunGramKB se distingue de ellos por estar basado en una estructura conceptual prediseñada y constituida por conceptos definidos formalmente *a priori* del trabajo con idiomas.

Para ilustrar someramente una diferencia entre el acercamiento relacional de EuroWordNet y el acercamiento de definición formal de FunGramKB, usaremos el ejemplo de BIRD, presentado por Perriñán y Arcas (2007b) y Mairal y Perriñán (2010). Partiendo del entendimiento de que un pájaro típico tiene muchas plumas, dos patas y dos alas, el acercamiento de significado relacional de EuroWordNet nos informará de que existen relaciones de meronimia en los siguientes *synsets*. Se necesita una entrada para detallar la relación de cada parte del pájaro al todo: *pájaro, pluma; pájaro, pata; pájaro, ala*:

Tabla 13. Ejemplo de significado relacional en EuroWordNet para el inglés

| |
|-------------------------------|
| HAS_MERO_PART {bird, feather} |
| HAS_MERO_PART {bird, leg} |
| HAS_MERO_PART {bird, wing} |

Asimismo, en una base de datos multilingüe como EuroWordNet estos *synsets* tendrán que repetirse para cada lengua. El hecho de que no haya una interlingua independiente de idiomas genera redundancias notables en la codificación. La próxima tabla, tomada de Perriñán y Arcas (2007b), ilustra las múltiples iteraciones necesarias para codificar la relación entre *pájaro* y *pluma* en varios idiomas:

Tabla 14. Ejemplo de significado relacional en EuroWordNet para varias lenguas

| | | |
|---------|---------------|---------|
| bird | HAS_MERO_PART | feather |
| Vogel | HAS_MERO_PART | Feder |
| pájaro | HAS_MERO_PART | pluma |
| uccello | HAS_MERO_PART | piuma |
| oiseau | HAS_MERO_PART | plume |

Frente a esto, el acercamiento de definición formal de FunGramKB usa un solo postulado de significado que define el concepto de +BIRD_00, al cual se vinculan los módulos léxicos. Ello reduce significativamente la redundancia en la base de datos, pues el hecho de que típicamente un ‘pájaro’ tiene muchas plumas, dos patas y dos alas solo tiene que ser codificado una vez en toda la base de datos:

Tabla 15. PS de +BIRD 00

| |
|---|
| <p>+(e1: +BE_00 (x1: +BIRD_00)Theme (x2: +VERTEBRATE_00)Referent) *(e2: +COMPRISE_00 (x1)Theme (x3: m +FEATHER_00 & 2 +LEG_00 & 2 +WING_00)Referent) *(e3: +FLY_00 (x1)Agent (x1)Theme (x4)Origin (x5)Goal)</p> |
|---|

Antes de proceder a discutir los MT y PS, haremos un repaso de algunos rasgos de COREL, el sistema de notación que se usa en la ontología. Se trata de un sistema de representación conceptual propio y común a los tres componentes de la base de conocimiento (la ontología, el cognición y el onomasticón).

4.1 El sistema de notación COREL⁵²

Para describir formalmente las representaciones mentales en el MT y PS, FunGramKB usa el lenguaje COREL, que consta de su propia semántica y sintaxis. En esto FunGramKB se diferencia de las bases de conocimiento que recurren al desarrollo de algoritmos para la representación de sus unidades conceptuales (Mairal, 2013). La tarea de moldear y gestionar el conocimiento mediante el lenguaje COREL es responsabilidad de los ingenieros del conocimiento, que trabajan mano a mano con lingüistas y lexicógrafos en el proyecto Lexicom. Según Perrián y Mairal (2010), al igual que con otros lenguajes de representación, COREL requiere dos componentes:

- (1) Una colección de términos, es decir, las unidades conceptuales que se organizan jerárquicamente para permitir la herencia y así facilitar la brevedad definitoria y aumentar el poder expresivo.
- (2) Un sistema de notación, es decir, la sintaxis por la que se tienen que regir las representaciones interlingüísticas para poder considerarlas estructuras bien formadas.

Dado que el enfoque de FunGramKB es mentalista, las unidades conceptuales se conciben como representaciones mentales estructuradas. Estas representaciones buscan capturar el conocimiento que tenemos los humanos de las propiedades características de una entidad típica de su clase. En ese sentido, el modelo intenta capturar las regularidades que percibe el ser humano en el mundo. Esto incluye también el conocimiento de sentido común que no necesariamente aparece en los diccionarios porque los humanos lo consideran obvio. Recordemos que para una aplicación computacional no hay nada obvio, por lo que todo debe ser explicitado. Por otro lado, toda predicación tiene regularidades pero también excepciones; la notación de COREL provee para que se marquen unas y otras.

⁵² Perrián y Mairal (2010) describen exhaustivamente la gramática de COREL.

Tomemos como ejemplo la representación del concepto básico +OPEN_00:

```
*(e1: +MOVE_00 (x1)Agent (x2)Theme (x3)Location (x4)Origin (x5)Goal (f1)Instrument (f2:
(e2: +BE_01 (x2)Theme (x6: $OPEN_N_00)Attribute))Condition (f3: (e3: +BE_01 (x2)Theme
(x7: +OPEN_00)Attribute))Result)
```

Según Perrián y Mairal (2010), la lectura de esta representación es como sigue: el enunciado (e1) expresa un Agente (x1) mueve una entidad Tema (x2) que se encuentra en una localización (x3) y la mueve desde un origen (x4) hasta una meta (x5) con un instrumento (f1). En (e2) se añade una condición previa que debe ser satisfecha para que ocurra (e1): la entidad (x2) que se mueve no tiene que estar inicialmente abierta (x6). El tercer enunciado (e3) expresa que el evento (e1) produce un resultado tal que la entidad (x2) que movemos terminará estando abierta (x7).

Cabe recordar aquí que la ontología de FunGramKB tiene una organización taxonómica, en donde los conceptos heredan las propiedades (MT y PS) de sus superordinados. Por ejemplo, el concepto +OPEN_00, que vimos arriba, mantiene el mismo MT de su superordinado +MOVE_00 y utiliza a +MOVE_00 en el primer enunciado de su PS, añadiendo luego las precisiones que son necesarias. FunGramKB también permite la herencia múltiple, en la que una unidad conceptual puede estar subsumida por dos conceptos o más, lo que crea jerarquías complejas en forma de un tejido en vez de un árbol. Por ejemplo, +SOLID_00 es subordinado de +CORPUSCULAR_00 y de +SUBSTANCE_00. La ruta conceptual de un subordinado de +SOLID_00, como +BALL_00, estará especificada en la primera predicación de su PS, en este caso, la ruta de +CORPUSCULAR_00. Finalmente, nótese que en FunGramKB, cada predicación en un PS está precedida por un operador de razonamiento que establece si la predicación es estricta (+) o rebatible (“defeasible”) (*). En el primer caso, la regla se toma como ley, mientras que en el segundo, se admite la posibilidad de que la regla sea supereditada por evidencia contraria. Tomemos como ejemplo nuevamente el concepto +BIRD_00, en específico, su PS:

```
+(e1: +BE_00 (x1: +BIRD_00)Theme (x2: +VERTEBRATE_00)Referent)
*(e2: +COMPRISE_00 (x1)Theme (x3: m +FEATHER_00 & 2 +LEG_00 & 2
+WING_00)Referent)
*(e3: +FLY_00 (x1)Agent (x1)Theme (x4)Origin (x5)Goal)
```

El primer enunciado es estricto: todo pájaro es un vertebrado. Los enunciados segundo y tercero, sin embargo, son rebatibles: aunque la mayoría de los pájaros tienen muchas plumas, dos alas y dos patas, no podemos decir que un pájaro mutilado, que haya perdido alguna de

estas partes, deje de ser pájaro. Asimismo, aunque los pájaros prototípicos pueden volar, existen excepciones como los pingüinos, que no por ello dejan de ser pájaros.

COREL también provee tres conectores lógicos: uno de conjunción, uno de disjunción y uno de exclusión. La conjunción se marca con el símbolo “&” y es el conector por defecto. La frase “a & b” quiere decir “a y b”. Podemos ver un ejemplo de conjunción en el PS de +BIRD_00, en el enunciado (e2), que establece que el pájaro tiene típicamente muchas plumas y dos patas y dos alas: (x3: m +FEATHER_00 & 2 +LEG_00 & 2 +WING_00)Referent). La disjunción se marca con el símbolo “|” y significa ‘una cosa o la otra, o bien ambas cosas’. La frase “a | b” significa “a y/o b”. Por ejemplo, el PS del concepto +HOLD_00 establece que algo se puede sostener con la mano o el brazo, o ambas partes a la vez: MP: +(e1: +HAVE_00 (x1)Theme (x2)Referent (f1: +HAND_00 | +ARM_00)Location). Finalmente, la exclusión se marca con el símbolo “^”. La frase “a ^ b” significa ‘a o b, pero no ambos’. En el MT de +WEAR_00, el Tema se describe mediante una exclusión: (x1: +HUMAN_00 ^ +PET_00)Theme. Esto quiere decir que el Tema de +WEAR_00 es ‘humano’ o ‘mascota’.

4.2 Los marcos temáticos (MT)

Cada evento y cualidad en la ontología tiene un MT, que establece el número y tipo de participantes involucrados típicamente en la situación cognitiva prototípica (Periñán y Arcas, 2007a; Mairal y Periñán, 2009a, entre otros). Siguiendo a Periñán y Mairal (2010), al especificar los participantes de un MT, no pensamos en el comportamiento léxico de un verbo, sino en los participantes que intervienen en una dimensión cognitiva. Por ejemplo, en la dimensión cognitiva #COMMUNICATION –que no es lo mismo que el verbo *comunicar*–, se podrían identificar básicamente tres papeles temáticos (o conceptuales): un *Tema* (“Theme”), o sea, la entidad que transmite un mensaje; un *Referente* (“Referent”), o sea, el mensaje; y una *Meta* (“Goal”), o sea, la entidad que recibe el mensaje⁵³. Dado esto, “un MT es un constructo cognitivo que refleja nuestro conocimiento del mundo, que no tiene por qué coincidir con el conocimiento lingüístico” (Mairal, 2013: 17).

El MT de un concepto se recibe de parte del concepto superordinado a través del mecanismo de herencia o subsunción. El concepto +ANSWER_00, por ejemplo, está subordinado al concepto +SAY_00, que a su vez está subordinado al metaconcepto #COMMUNICATION, que tiene el MT: (x1)Theme (x2)Referent (x3)Goal. Tanto

⁵³ Periñán y Mairal (2010) describen los papeles temáticos que definen cada dimensión conceptual.

+SAY_00 como +ANSWER_00 y sus subordinados heredan el MT de #COMMUNICATION.

Los MT son también el espacio donde se pueden incluir las preferencias de selección típicamente involucradas en una situación cognitiva. Cuando los participantes obligatorios en el MT tienen que ser rellenados por un concepto o predicación particular, estamos ante una preferencia de selección conceptual. Por ejemplo, el MT del concepto básico +LIKE_00 lee así: (x1)Agent (x2: +HUMAN_00^+ANIMAL_00)Theme. Para entender esto cabalmente, debemos tomar en cuenta que #LIKE_00 pertenece a la dimensión del metaconcepto #EMOTION, que define sus participantes así: el participante Agente es ‘la entidad que hace que otra entidad sienta una emoción’, el Tema es ‘la entidad que siente la emoción’ y puede haber Atributo, que es ‘la entidad o cualidad que describe un atributo de una entidad cuando está sintiendo una emoción’. En este caso, pues, el concepto +LIKE_00 tiene un Agente, que es la entidad que hace que otra entidad sienta una emoción, que no está especificado semánticamente. El participante Tema, que es la entidad que siente la emoción, está restringido semánticamente para un ‘humano’ o ‘animal’. Esto quiere que cualquier entidad puede provocar una emoción de agrado, pero tiene que experimentarla un humano o un animal. El concepto básico +FEAR_00 también es subordinado de #EMOTION, por lo que comparte los mismos participantes que +LIKE_00. En este caso, el MT de +FEAR_00 lee así: (x1)Agent (x2: +HUMAN_00)Theme. Aquí la entidad que provoca la sensación tampoco está especificada, pero el Tema que la experimenta restringe su selección a ‘humano’. El concepto +FEAR, pues, restringe el evento de sentir miedo a un humano.

4.3 Los postulados de significado (PS)

El PS representa la definición de una unidad ontológica en FunGramKB, esto es, una descripción de nuestro conocimiento sobre esa unidad conceptual. Como la descripción formal del significado en FunGramKB se construye con conceptos y no palabras, el PS se convierte en una representación del conocimiento semántico independiente de las lenguas.

El PS está constituido por una serie de una o más predicaciones lógicamente conectadas ($e^1, e^2 \dots e^n$). Estas son construcciones conceptuales que aportan los rasgos genéricos de los conceptos. Los PS también pueden incluir satélites, marcados como f1, f2, etc., para expresar parámetros adverbiales como Manera, Condición e Instrumento⁵⁴. Los satélites pueden estar formulados por conceptos o predicaciones completas.

⁵⁴ Ver Apéndice II, para la lista completa de satélites y su interpretación semántica.

Los PS organizan la ontología de acuerdo con tres principios que ayudan a mantener la coherencia interna en la taxonomía. El primero es el principio de similitud, según el cual todos los conceptos subordinados deben compartir el postulado de significado de su superordinado. El segundo principio es el de especificidad, que dice que todos los conceptos subordinados deben tener un postulado de significado que establezca un rasgo distintivo (o *differentia*) que no esté presente en el PS de su superordinado. El tercero es el principio de oposición, según el cual las *differentiae* en los PS de los conceptos hermanos deben ser incompatibles entre sí.

Para ilustrar lo anterior, veamos la descripción del concepto +STORE_00 en FunGramKB:

Tabla 16. Descripción de +STORE_00 en FunGramKB

| | |
|-------------------|---|
| Superordinado: | +HAVE_00 |
| MT: | (x1: +HUMAN_00 ^ +ANIMAL_00 ^ +ORGANIZATION_00)Theme (x2: +ARTEFACT_00 ^ +CORPUSCULAR_00 ^ +SUBSTANCE_00)Referent |
| PS: | +(e1: +HAVE_00 (x1)Theme (x2)Referent (f1)Location (f2: +LONG_01)Duration) |
| Descripción: | provide storage for or keep in storage |
| Lexicalizaciones: | <i>acaparar, acopiar, acumular, ahorrar, almacenar, atesorar, conservar, economizar, guardar, monopolizar</i> (español), <i>accumulate, bond, garner, hoard, hold, keep, monopolize, save, stock</i> (inglés) |

El MT de +STORE_00 nos informa de que la situación prototípica de ‘almacenar algo’ tiene dos participantes: un Tema y un Referente. Recordemos que la definición de los roles temáticos de FunGramKB depende del metaconcepto bajo el cual esté ubicado el concepto en cuestión, en este caso, el metaconcepto correspondiente es #POSSESSION_00, por lo cual el Tema es ‘una entidad que posee otra entidad’ y el Referente es ‘una entidad que es poseída’. Para la situación cognitiva de ‘almacenar algo’, o sea, para el concepto +STORE_00, el Tema y el Referente están especificados semánticamente: el Tema es típicamente +HUMAN_00, +ANIMAL_00 u +ORGANIZATION_00, y el Referente es típicamente +ARTEFACT_00, +CORPUSCULAR_00 o +SUBSTANCE_00. La especificación semántica de los papeles temáticos constituye en FunGramKB las preferencias de selección, como comentamos antes y ampliaremos más adelante.

El PS de +STORE_00, por su parte, expresa los rasgos genéricos asociados con el significado del concepto; en este caso, el PS expresa que un ser humano, animal u organización (x1=Tema) típicamente almacena artefactos, sustancias u objetos corpusculares

(x2=Referente) en un lugar (f1) por una duración larga (f2). Los participantes del MT de un evento quedan pues perfectamente integrados en su PS.

Para ilustrar los principios que organizan la ontología, veamos ahora el concepto +PRESERVE_00, subordinado de +STORE_00:

Tabla 17. Descripción de +PRESERVE_00 en FunGramKB

| | |
|-------------------|--|
| Superordinado: | + STORE_00 |
| MT: | (x1)Theme (x2)Referent |
| PS: | +(e1: +STORE_00 (x1)Theme (x2)Referent (f1: (e2: +DAMAGE_00 (x3)Theme (x2)Referent))Purpose) |
| Descripción: | keep something so that it will not be damaged |
| Lexicalizaciones: | <i>conserva</i> (español), <i>preserve</i> (inglés) |

Según los principios de similitud y de especificidad, todos los conceptos subordinados deben compartir el postulado de significado de su superordinado, pero a la vez establecer un rasgo distintivo (o *differentia*) que no esté presente en el PS de su superordinado. En el PS de +PRESERVE_00, se ha incorporado la noción de +STORE_00 en el primer enunciado: +(e1: +STORE_00 (x1)Theme (x2)Referent, y se ha añadido un satélite que establece que el propósito de esta acción es evitar el daño al Referente: (f1: (e2: +DAMAGE_00 (x3)Theme (x2)Referent))Purpose). Así, cada concepto subordinado encapsula una versión más específica del concepto precedente.

Para ilustrar el principio de oposición, según el cual las *differentiae* en los PS de los conceptos hermanos deben ser incompatibles entre sí, consideremos el concepto +BREED_00, concepto hermano de +STORE_00:

Tabla 18. Descripción de +BREED_00 en FunGramKB

| | |
|-------------------|--|
| Superordinado: | +HAVE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2: +ANIMAL_00 ^ +PLANT_00 ^ +VEGETABLE_00 ^ +LAND_00)Referent |
| PS: | +(e1: +HAVE_00 (x1)Theme (x2)Referent (f1: (e2: +FEED_00 (x1)Agent (x3: +FOOD_00)Theme (x4)Origin (x2)Goal))Purpose & (f2: (e3: +CLEAN_01 (x1)Theme (x2)Referent))Purpose) |
| Descripción: | keep animals for the purpose of producing more animals: "She breeds dogs |
| Lexicalizaciones: | criar (español), breed (inglés) |

Se mantiene el postulado original: +(e1: +HAVE_00 (x1)Theme (x2)Referent, pero sus participantes son diferentes. En este caso el Tema es ‘humano’, y el Referente puede ser ‘animal’, ‘planta’, ‘vegetal’ o ‘tierra’. Además, a este PS se ha añadido un satélite muy específico: (f1: (e2: +FEED_00 (x1)Agent (x3: +FOOD_00)Theme (x4)Origin

(x2)Goal))Purpose & (f2: (e3: +CLEAN_01 (x1)Theme (x2)Referent))Purpose), que establece que el Tema (x1) alimenta (e2) y limpia (e3) al Referente (x2) con un propósito (f1 y f2). Vemos pues cómo del origen común que provee +HAVE_00, se generan dos conceptos hermanos pero muy distinguidos entre sí: +STORE_00 y +BREED_00, y cómo +STORE_00 da pie a otro concepto más especificado aun: +PRESERVE_00.

5. Las preferencias de selección en FunGramKB

Las preferencias de selección en FunGramKB restringen la naturaleza semántica de los participantes de un concepto, cuando la situación cognitiva expresada por el concepto así lo requiere (Jiménez y Pérez, 2011). Se establecen solo cuando ejercen algún poder predictivo sobre el participante. Un participante –Agente, Tema, Referente, etc.– es un argumento de un verbo si satisface tanto el criterio de obligatoriedad semántica como el criterio de especificidad semántica (Periñán y Mairal, 2010). Ahora bien, cuando ese participante obligatorio tiene que ser rellenado por un concepto o predicación particular, estamos ante una preferencia de selección conceptual.

La noción de ‘comer’, por ejemplo, tiene en nuestras mentes –independientemente de nuestro idioma– un primer participante prototípico que realiza la acción de comer. Usualmente ese agente que come es un humano o animal. Ello está dictado por nuestro conocimiento de sentido común y nuestro modelo del mundo, que nos informan de que el acto de comer requiere una boca y que solo los humanos y los animales tienen boca. En el evento +EAT_00 (‘comer’) en FunGramKB, el MT identifica el primer participante como un Agente y, además, restringe ese Agente a (x1: +HUMAN_00 ^ +ANIMAL_00), donde el operador ^ expresa exclusión entre las preferencias de selección. Vemos, pues, que primero se establecen los participantes de un concepto y luego se restringen semánticamente, si es necesario. Como las preferencias de selección en FunGramKB no están enfocadas en las palabras, sino en los conceptos, se ubican en el nivel conceptual, específicamente en la ontología, como parte del MT y PS de cada concepto.

Las preferencias de selección de FunGramKB pertenecen pues a un modelo cognitivo y conceptual, por lo cual la información que capturan es de naturaleza extralingüística. En esto se distinguen de las restricciones de selección léxica de REDES que, como hemos dicho, reflejan nuestro conocimiento lingüístico. Sin embargo, existe un punto de encuentro entre las restricciones de selección de REDES y las preferencias de selección de FunGramKB que permite la comparación entre ambos modelos. Aunque las restricciones de selección de REDES parten del léxico, tienen una interpretación semántica en REDES que podemos

proyectar a las preferencias conceptuales de FunGramKB.

Es importante destacar, por otro lado, que el traducir las restricciones de selección léxica de REDES al plano conceptual de las preferencias de selección de FunGramKB, las validaría como parte de nuestro conocimiento humano, ubicándolas más allá de las lenguas particulares. Queda claro también que la formalización de las preferencias de selección en FunGramKB nos permitirá ver patrones y hacer generalizaciones que son difíciles de percibir en las descripciones *ad hoc* en lenguaje natural de las clases léxicas en REDES.

Las preferencias de selección se pueden manifestar en los conceptos básicos, conceptos terminales y los subconceptos. A continuación discutiremos el rol de las preferencias de selección en cada uno de ellos.

5.1 Preferencias de selección en los conceptos básicos

La próxima tabla muestra la descripción del concepto +WEAR_00 en FunGramKB.

Tabla 19. Descripción de +WEAR_00 en FunGramKB

| | |
|-------------------|---|
| Superordinado: | +HAVE_00 |
| MT: | x1: +HUMAN_00 ^ +PET_00)Theme (x2: +CLOTHING_00 ^+ORNAMENT_00) Referent |
| PS: | +(e1: +HAVE_00 (x1)Theme (x2)Referent (f1: +BODY_AREA_00)Location (f2: +ON_00)Position) |
| Lexicalizaciones: | <i>llevar, llevar puesto, traer</i> (español) <i>wear, dress, have on</i> (inglés) |

Nuestro sentido común nos dice que el evento ‘llevar puesto’ requiere que alguien lleve puesto algo. Según Jiménez y Pérez (2011), sabemos que son los humanos y las mascotas los que prototípicamente llevan ropas, zapatos, lazos, joyería, etc. Como muestra la tabla anterior, el MT de +WEAR_00 en FunGramKB identifica dos participantes: un Tema y un Referente. Esos participantes se restringen conceptualmente como: (x1: +HUMAN_00 ^ +PET_00)Theme (x2: +CLOTHING_00 ^ +ORNAMENT_00)Referent. Esto quiere decir que en el evento de +WEAR_00 hay un Tema, que puede ser ‘humano’ o ‘mascota’, y un Referente, que puede ser una pieza de ropa o un ornamento. Estas especificaciones semánticas de los participantes de un evento constituyen las “preferencias de selección” de +WEAR_00 en FunGramKB.

Las preferencias de selección +HUMAN_00, +PET_00, +CLOTHING_00, +ORNAMENT_00, etc., por lo tanto, pretenden ir más allá del conocimiento lingüístico para tratar de capturar nuestro modelo del mundo. Al confrontar esta apreciación con los datos lingüísticos de REDES, encontramos una validación. REDES incluye, bajo *llevar*, la

siguiente clase léxica abierta: “sustantivos que designan prendas y complementos de vestir (*sombrero, chaqueta, camisa, calzado, pantalón: La muchacha llevaba puesto un pantalón vaquero y una camisa de flores*)” (Bosque, 2004b: 1279, sub voce *llevar*). Existe una correspondencia entre el Referente +CLOTHING_00 u +ORNAMENT_00 de FunGramKB, y la clase léxica de ‘prendas y complementos de vestir’ de REDES.

Esta tesis tiene entre sus propósitos comparar la información que nos da FunGramKB sobre preferencias de selección conceptual con la información que nos da REDES sobre selección léxica, y utilizar la segunda para rellenar la primera, en la medida de lo posible.

5.2 Preferencias de selección en los conceptos terminales

Los conceptos terminales son necesarios cuando hay una restricción conceptual en el significado de un concepto básico, es decir, en su PS. Por ejemplo, el concepto \$SPORT_00 surge de la necesidad de restringir su concepto superordinado +WEAR_00. Según explican Jiménez y Pérez (2011), el concepto terminal \$SPORT_00, lexicalizado en español por *lucir* y *ostentar* y en inglés por *sport*, se genera porque las personas pueden prototípicamente llevar puestas cosas de una manera orgullosa para que todo el mundo los mire.

Para codificar esta restricción semántica conceptualmente, necesitamos especificar tanto el MT como el PS de +WEAR_00. En el MT de \$SPORT_00, el primer participante (Tema) se restringe solo a ‘humanos’ (x1: +HUMAN_00) y el segundo (Referente) se expande para añadir +HAIR_01. En el PS de \$SPORT_00, se añade un parámetro de propósito (f1), que está constituido por una predicación (e2) que expresa que el Tema (x1) lleva el Referente (x2), y los despliega de manera orgullosa.

Tabla 20. Descripción de \$SPORT_00 en FunGramKB

| | |
|-------------------|---|
| Superordinado: | +WEAR_00 |
| MT: | (x1: +HUMAN_00)Theme (x2: +CLOTHING_00 ^ +HAIR_01 ^ +ORNAMENT_00)Referent |
| PS: | +(e1: +WEAR_00 (x1)Theme (x2)Referent (f1: (e2: +SHOW_00 (x1)Theme (x2)Referent (f2: +PROUD_00)Manner))Purpose) |
| Lexicalizaciones: | <i>lucir, ostentar</i> (español) <i>sport</i> (inglés) |

5.3 Preferencias de selección en los subconceptos

Los subconceptos surgen cuando la especificación o restricción conceptual ocurre dentro del MT de un concepto básico o terminal, sin que haya cambio alguno en el PS. Los subconceptos no se consideran unidades conceptuales reales, sino refinamientos léxicamente motivados de las preferencias de selección en el MT de conceptos básicos o terminales

existentes, que vienen a ser sus “superconceptos”. Los subconceptos comparten con sus superconceptos todas las propiedades conceptuales, incluidos el número y rol de los participantes en el MT (Periñán y Mairal, 2011).

Los subconceptos son usados cuando la naturaleza conceptual de evento es extremadamente específica; por ejemplo, dentro del escenario de ‘emisión de sonido’, el evento de ‘ladrar’ se distingue en que aplica exclusivamente a ‘perros’. Por lo tanto, el subconcepto -BARK se introduce con la especificación de +DOG_00 en el Tema de \$SOUND_00, manteniendo el mismo PS (Luzondo, 2011). Asimismo, bajo el concepto básico +CLEAN_01, el subconcepto -PREEN_00 estipula que el Tema siempre es un pájaro y el Referente siempre son las plumas.

Jiménez y Luzondo (2011) explican la creación de varios subconceptos, por ejemplo, -TAKE_SHOES, bajo el concepto básico +WEAR_00. Este subconcepto mantiene la misma información conceptual de +WEAR_00, salvo que especifica el Referente como +SHOE_00, que es la restricción léxica de *calzar*. Su descripción conceptual es la siguiente:

Tabla 21. Descripción de -TAKE_SHOES en FunGramKB

| | |
|-------------------|---|
| Superconcepto: | +WEAR_00 |
| MT: | (x1: +HUMAN_00)Theme (x2: +SHOE_00)Referent |
| PS: | +(e1: +HAVE_00 (x1)Theme (x2)Referent (f1: +BODY_AREA_00)Location (f2: +ON_00)Position) |
| Descripción: | someone has shoes on the body |
| Lexicalizaciones: | <i>calzar</i> (español), <i>take shoes</i> ⁵⁵ (inglés) |

El subconcepto -WIELD, por su parte, es una restricción del concepto terminal \$GRASP_00. Se lexicaliza como *wield*, *carry*, *bear* (inglés) y *empuñar* (español). En el caso de -WIELD, el Referente +CORPUSCULAR_00 de \$GRASP_00 se restringe solo a armas (+WEAPON_00):

Tabla 22. Descripción de -WIELD en FunGramKB

| | |
|-------------------|--|
| Superconcepto: | \$GRASP_00 |
| MT: | (x1: +HUMAN_00)Theme (x2: +WEAPON_00)Referent |
| PS: | +(e1: +HOLD_00 (x1)Theme (x2)Referent (f1: +TIGHT_00)Manner) |
| Lexicalizaciones: | <i>empuñar</i> (español), <i>wield</i> , <i>carry</i> , <i>bear</i> (inglés) |

⁵⁵ La frase *take shoes* como unidad léxica inglesa podría estar incompleta. La construcción habitual es *take size-X shoes*, por ejemplo: *I take size 9 shoes* o *I take a size 9*. La selección parece ser pues entre *take* y *size*, no entre *take* y *shoes*, como se puede comprobar en los diccionarios del inglés.

Como se verá el capítulo V, los subconceptos son un recurso muy útil para manejar la información combinatoria que nos provee REDES.

6. Las colocaciones en FunGramKB

En FunGramKB, las colocaciones se entienden en un sentido amplio para referirse a las combinaciones de lexemas que común y frecuentemente coocurren en la lengua, e incluyen tanto las colocaciones léxicas como las gramaticales. Las colocaciones están ubicadas en el nivel léxico, específicamente en el lexicón, por considerarse dependientes de cada idioma. Por ejemplo, el hecho de que en español *se depende DE algo*, pero en inglés *you depend ON something*, se ubica en los lexicones español e inglés en FunGramKB, respectivamente (Jiménez y Pérez, 2011). Asimismo, los verbos *atesorar* y *hoard* se vinculan al evento +STORE_00, pero en cada lexicón se registra el hecho de que *atesorar* se combina con *victoria*, *éxito*, *información*, *secreto* y *recuerdo*, mientras que *hoard* se combina con *flaw*, *time*, *misery*, *nostalgia* y *information*.

FunGramKB inserta los colocados como parte de la información morfosintáctica y pragmática vinculada a cada palabra, específicamente, en el bloque de gramática nuclear, en el espacio para colocaciones para el segundo argumento (y) de cada predicado, como se ve a continuación.

Imagen 10. Plantilla léxica de *atesorar* en el lexicon

Lexical Template:

Variables: <-- no functions -->

Idiosyncratic features:
[MR <-- no value selected -->], [U = <-- no value selected -->]

Thematic frame mapping:
X = [no function] , Y = [no function] , Z = [no function]

A REMINDER OF FUNGRAMKB PARTICIPANTS:
THEME: Entity that owns another entity.
REFERENT: Entity that is owned.

Prepositions:
X = a Y Y = a Y Z = a Y
N N N

Collocations:
X = Y Y = Y Z = Y
N N N

éxito
información
recuerdo
secreto
victoria

Ahora bien, cuando entramos en contacto con datos como los de REDES, se hace claro que no podemos solamente vaciar los argumentos léxicos en el espacio de colocaciones del módulo léxico. Eso pasaría por alto las generalizaciones semánticas sobre preferencias de selección que nos permite hacer la selección léxica documentada en REDES. Esas preferencias de selección podrían traducirse en nuevos conceptos terminales o subconceptos, que condifiquen conceptualmente la restricción de selección.

Veamos la descripción de +STORE_00, concepto al que está vinculado *atesorar*:

Tabla 23. Descripción de +STORE en FunGramKB

| | |
|-------------------|---|
| Superordinado: | +HAVE_00 |
| MT: | (x1: +HUMAN_00 ^ +ANIMAL_00 ^ +ORGANIZATION_00)Theme (x2: +ARTEFACT_00 ^ +CORPUSCULAR_00 ^ +SUBSTANCE_00)Referent |
| PS: | +(e1: +HAVE_00 (x1)Theme (x2)Referent (f1)Location (f2: +LONG_01)Duration) |
| Lexicalizaciones: | <i>acaparar, acopiar, acumular, ahorrar, almacenar, atesorar, conservar, economizar, guardar, monopolizar</i> (español) |

Los colocados de *atesorar* que se identifican arriba (*victoria, éxito, información, secreto y recuerdo*) son parte de la combinatoria sistemática documentada por REDES para *atesorar*, que incluye tres clases léxicas abiertas y seis clases léxicas en propiedad. Esas nueve clases léxicas se podrían traducir al plano conceptual, y codificar como preferencias de selección. Por ejemplo, podríamos proponer la creación de un subconcepto -TREASURE_00, que restrinja el Referente de +STORE_00 a entidades como +VICTORY_00 e +INFORMATION_00.

Conclusión

En el contexto de la lexicografía del siglo XXI, la base de conocimiento FunGramKB se destaca por su acercamiento de semántica profunda al significado léxico, frente al acercamiento de semántica superficial utilizado por otros lexicones. Por otro lado, en el continuo de granularidad semántica, esta base de conocimiento se ubica en un punto intermedio con respecto a los lexicones enumerativos existentes.

La estructura conceptual de los eventos de la ontología, descrita en el MT y PS, da cabida a las preferencias de selección de un concepto. Este es precisamente el espacio que nos permite codificar conceptualmente las restricciones de selección léxica que se establecen en REDES. La codificación de nuevas preferencias de selección puede significar la creación de un concepto terminal o un subconcepto. El módulo léxico, por su parte, provee el espacio para integrar las colocaciones de cada predicado que no se puedan describir conceptualmente. FunGramKB provee, por lo tanto, todos los elementos necesarios para ubicar las tres unidades de REDES –predicados, clases léxicas y argumentos–, sin pérdida de información y sin redundancia.

CAPÍTULO IV. LA POLISEMIA EN LA LEXICOGRAFÍA INFORMATIZADA

Introducción

En este capítulo discutiremos el fenómeno de la polisemia y el tratamiento que se le ha dado en la lexicografía, especialmente en algunos lexicones electrónicos y bases de conocimiento, tanto monolingües como multilingües. La resolución automática de la polisemia –es decir, la asignación adecuada de sentidos a palabras polisémicas a través de los mecanismos de WSD– es señalada como el problema fundamental del PLN. Se reconoce que los retos que presenta la WSD no tienen una causa única, sino que responden a múltiples factores (Navigli, 2009). A través de este repaso del estado de la cuestión de la WSD, intentamos enmarcar una de nuestras hipótesis de trabajo: la idea de que la información de REDES podría ayudar a desambiguar las palabras polisémicas en la estructura de FunGramKB, mediante la codificación conceptual de los sentidos de una palabra a la luz de sus preferencias de selección.

La polisemia no presenta dificultad alguna para los hablantes, que codifican y descodifican los sentidos adecuados de las palabras sin ningún esfuerzo ni conciencia de ello, según las pistas que ofrezcan los contextos lingüístico y comunicativo. Sin embargo, la polisemia sí constituye un problema para la teoría léxico semántica y para prácticas como la lexicografía, la traducción, la WSD y, por consiguiente, las diversas aplicaciones del PLN. Todo hispanohablante comprenderá perfectamente los sentidos de *perder* en *perder la cartera*, *perder el campeonato*, *perder el tiempo* y *perder el tren*, pero un programa de traducción automática (*Machine Translation* o MMTT) al inglés necesitará distinguirlos para producir los equivalentes apropiados: *lose one's purse* y *lose the championship*, pero *waste time* y *miss the train*.

La WSD implica, a grandes rasgos, el pareo de una palabra en un texto con el sentido adecuado, el cual debe poder distinguirse de todos los otros sentidos atribuibles a esa palabra. Esta tarea requiere la existencia de tres componentes:

- (1) el texto o discurso en donde ocurre la palabra;
- (2) los posibles sentidos de la palabra, al menos para el texto considerado;
- (3) el método para asignarle el sentido apropiado a esa instancia de la palabra.

Todos los mecanismos de WSD, por lo tanto, extraen información de la palabra en el texto, la comparan con información de diferenciación de sentidos almacenada para esa palabra y le asignan el sentido más apropiado. Cada uno de los tres componentes –texto, sentidos y método– conlleva una serie de decisiones fundamentales, que influirán en la formalización de los mecanismos de WSD. Entre estas se incluyen, según Navigli (2009: 2, traducción mía):

- (a) el acercamiento que se usará para representar los sentidos de la palabra: las posibilidades van desde la enumeración de una serie finita de sentidos hasta la generación de sentidos nuevos a partir de reglas;
- (b) el nivel de granularidad de los inventarios de sentidos: se pueden establecer distinciones semánticas sutiles o limitar la distinción a los homónimos;
- (c) la naturaleza de los textos donde se realice la WSD: estos pueden pertenecer a un dominio temático particular o no tener restricciones temáticas;
- (d) el número de palabras meta que tiene la tarea de WSD: puede atenderse desde una palabra meta por oración hasta todas las palabras.

En este capítulo, repasaremos desde diversos ángulos el problema de la polisemia en la lexicografía, tomando en cuenta algunas de las preguntas que postula Navigli (2009), arriba. En la sección 1, nos detendremos brevemente en las nociones de ‘polisemia’, ‘significado’ y ‘sentidos’, y su relevancia para este trabajo. En la sección 2, repasaremos el panorama histórico de la WSD en los lexicones electrónicos y el PLN. En la sección 3, veremos la influencia de la lexicografía tradicional en la lexicografía informatizada. En la sección 4, examinaremos los diferentes niveles de granularidad semántica en los lexicones electrónicos. En la 5, discutiremos los diferentes métodos de WSD que se utilizan en el PLN. En la sección 6, nos enfocaremos en el rol del contexto para la WSD y en la subsección 6.1 pondremos nuestra atención en las preferencias de selección y la semántica argumental como vías para la WSD. Terminamos el capítulo con una conclusión.

1. Polisemia, significado y sentidos

‘Polisemia’ es el fenómeno por el cual una palabra puede tener más de un significado. El manejo de la polisemia supone una serie de preguntas fundamentales, que no parecen tener una sola contestación: ¿qué es el ‘significado’ o qué son los ‘sentidos’ de una palabra? ¿Cómo se detectan estos sentidos? ¿Qué pistas o pruebas apuntan a uno u otro sentido? ¿Cómo se delimitan entre sí? ¿Cuál es su contenido? ¿Cómo se representan? Y, para efectos

de la WSD y el PLN, ¿cómo se adiestra una computadora para que identifique y gestione los sentidos de una palabra?

No existe consenso entre semantistas, lexicógrafos o investigadores del PLN sobre las nociones de ‘significado’ o ‘sentido’. Por consiguiente, tampoco hay coincidencia sobre los sentidos que pueda tener una palabra dada, cómo se delimita uno de otro y cuál es su contenido semántico. Estudiosos como Kilgarriff (1997) llegan a aseverar que la noción misma de ‘sentido’ es cuestionable como unidad básica de trabajo, dada su falta de delimitación tanto en la investigación lexicológica como el ámbito de la desambiguación semántica. Este autor concluye que la noción de ‘sentido’ depende de los intereses particulares de la tarea que se esté realizando: los sentidos de una palabra no serán los mismos para un lexicógrafo que trabaja con diccionarios impresos que para un ingeniero lingüístico que trabaja en una ontología para el PLN.

A pesar de la imprecisión terminológica y de los debates sobre la realidad teórica, lingüística y psicológica de los sentidos, el trabajo práctico de identificación de sentidos nunca se ha detenido. Tanto en la lexicografía tradicional, como en los proyectos de WSD, se ha avanzado haciendo caso omiso hasta cierto punto de estas preguntas (Agirre y Edmonds, 2006). Sin embargo, cabe señalar, siguiendo a Ide y Véronis (1998), que luego de décadas de trabajo de lingüística computacional y de muchos estudios de WSD que no se pueden comparar porque no tratan los sentidos de las palabras de la misma manera, el problema de la delimitación de las nociones de ‘significado’ y ‘sentido’ –y lo que son los sentidos de una palabra dada– se empieza a reconocer y a discutir críticamente.

Este trabajo reconoce el fenómeno de la polisemia, pero no requiere delimitar críticamente las nociones de ‘significado’ y ‘sentido’. Las unidades de significado en este trabajo serán los conceptos ontológicos de FunGramKB; por ejemplo, la unidad léxica *perder* se vincula con los eventos conceptuales +LOSE_00, +LACK_00, +WASTE_00, +FAIL_00 y \$MISS_00⁵⁶. Esto quiere decir que vinculamos a *perder* con cinco sentidos; esto se discutirá a fondo en el capítulo VII. Aparte de esto, en el texto, cuando identifiquemos sentidos provisionalmente, lo haremos en términos generales, sin pretender proponer definiciones lexicográficas; además, tomaremos por buenas las definiciones de los diccionarios tradicionales que usemos de referencia.

⁵⁶ El concepto terminal \$MISS es una propuesta de nuestra tesis.

2. La polisemia en el PLN: panorama histórico⁵⁷

Los primeros MRD se crean para servir a la investigación sobre la traducción automática, que inicia en la década de los 1940. Desde entonces, la ambigüedad semántica de las palabras –y la gestión práctica para resolverla: la WSD– se identifica como un problema fundamental. Para la lexicografía tradicional, la desambiguación de sentidos es un fin en sí mismo: el diccionario tiene el objetivo final de proveer los sentidos posibles de una palabra. Sin embargo, para el PLN la desambiguación semántica es solo un paso intermedio –aunque obligatorio– para realizar tareas particulares (Niremburg y Raskin, 1996; Navigli, 2009), que incluyen no solo la MT, sino también la búsqueda y extracción de información, la navegación de hipertextos, el análisis temático y de contenido, el análisis gramatical, el procesamiento del habla y el procesamiento de textos (Ide y Véronis, 1998). En el PLN, los errores en desambiguación semántica provocan resultados erróneos en la traducción automática o en la recuperación y gestión de datos, lo cual, a su vez, puede generar graves atascos en la autopista de la información (Mairal y Periñán, 2011).

El memorando sobre traducción automática de Weaver (1949) apunta por primera vez a la desambiguación como un problema central en la lingüística computacional, y destaca varios elementos del trabajo de desambiguación que son relevantes al día de hoy: el contexto como pista importante para establecer los sentidos de la palabra, la necesidad de contar con representaciones estandarizadas de la estructura del lenguaje y la importancia de los estudios semánticos estadísticos. En aquel momento, las palabras de la lengua fuente se organizaban alfabéticamente en un lado de la hoja, y las de la lengua meta en el otro lado; sin embargo, este acercamiento resultaba insuficiente: la traducción de las palabras junto con las reglas de combinación de palabras en la lengua meta no eran capaces de resolver la ambigüedad léxica (Boas, 2009c).

En la década de los 1950, los MRD se hacen más estructurados, parcialmente dado el desarrollo de técnicas más avanzadas de análisis oracional (*parsing*) y los nuevos diseños de sistemas de MMTT que hacían distinciones entre las reglas lingüísticas, la gramática y el léxico (Boas, 2009c). Sin embargo, en su tratado sobre la MMTT, Bar-Hillel (1960) concluye que la identificación automática de sentidos es en ese momento imposible, porque requiere caracterizar la totalidad del conocimiento de mundo. Esta declaración reduce el entusiasmo y el respaldo a los trabajos sobre traducción automática en la década de los 1960.

⁵⁷ Ide y Véronis (1998) ofrecen un estado de la cuestión detallado sobre la desambiguación semántica en el PLN hasta ese momento, mientras que Ravin y Leacock (2002) ofrecen una excelente panorámica abreviada. Agirre y Edmonds (2006) actualizan el panorama de la WSD y articulan las conclusiones alcanzadas a ese momento.

En la década de 1970, el foco gira hacia la IA, donde la desambiguación se convierte en una subtarea de los sistemas de interpretación semántica. Sin embargo, la codificación manual de las reglas en los sistemas de desambiguación resulta muy costosa tanto en el plano humano como en el informático.

En la década de 1980, con el avance tecnológico que populariza las computadoras, comienzan a surgir recursos léxicos de gran escala, como los diccionarios electrónicos⁵⁸. Se pretende entonces extraer información léxico semántica de estos diccionarios para crear grandes bases de conocimiento que sirvan a la desambiguación. Entre ellos, Boas (2009c) destaca el *Collins English Dictionary* (1986), el *Longman Dictionary of Contemporary English* o LDOCE (1987), el *Collins Cobuild English Language Dictionary* (1987), el *Webster's New World Dictionary* (1988), el *Oxford Advanced Learner's Dictionary* (1989). Según Ide y Véronis (1998), la meta no se logra, en parte por las inconsistencias entre los diccionarios mismos y en parte porque el diseño de los diccionarios estaba pensado para humanos, no para máquinas; otra limitación que se detecta en los diccionarios era su carencia de información pragmática.

En la década de 1990, ocurre una revolución estadística en la lingüística computacional, y se empiezan a aplicar técnicas supervisadas de aprendizaje electrónico a la desambiguación. Se cuenta ya con grandes córpora textuales, procesables electrónicamente, que posibilitan el análisis estadístico de los patrones de coocurrencia de palabras en el corpus. Esto se vuelve una práctica generalizada en la lingüística computacional: se toman oraciones que representen un sentido de una palabra y, a partir de ellas, se crean modelos de uso para cada sentido de la palabra. Al desambiguar una palabra en un sentido novel, el sistema escoge el modelo más apropiado computando la similitud entre los rasgos del modelo y los del contexto. Uno de los problemas del trabajo con corpus es, sin embargo, que hay que marcar manualmente los contextos en que aparece cada palabra con cada sentido. Ello se hace para cada palabra individualmente; no se puede proyectar automáticamente a otras palabras. A este problema se ha llamado el “embotellamiento de la adquisición del léxico”⁵⁹. En esa década, también aparece WordNet (Miller, 1990), que da un impulso significativo a la investigación, tanto por estar disponible electrónicamente como por organizar sus sentidos taxonómicamente; como hemos comentado antes, WordNet es hoy el inventario de sentidos más usado en los estudios de WSD.

⁵⁸ Wilks *et al.* (1996) proveen una descripción detallada del uso de los diccionarios electrónicos para la WSD.

⁵⁹ Como indicamos antes, el término “lexical acquisition bottleneck” es de Gale *et al.* (1992).

En 1997, Kilgarrieff organiza SensEval (ahora SemEval⁶⁰), una conferencia para comparar los resultados de los sistemas más actualizados de WSD. En la primera edición, 18 participantes se dedicaron a desambiguar 35 palabras, usando los mismos materiales de entrenamiento y prueba. Se confirmó entonces la llamada “falacia del 80%”, término acuñado por Bar-Hillel, que postula que el que 80% del problema esté resuelto, no significa que el 20% restante se resolverá aumentando por 20% la investigación; más bien, este 20% restante requerirá más trabajo que el 80% inicial.

En la década de los 2000, las técnicas supervisadas de desambiguación de sentidos de granularidad fina parecen alcanzar la precisión máxima posible, que resulta aún muy limitada. La atención gira entonces hacia los sentidos de granularidad gruesa, la consideración del dominio temático en la desambiguación, los sistemas basados en corpus semisupervisados y no supervisados, los métodos que combinan diferentes modelos y los modelos que utilizan gráficas basadas en conocimiento.

Se reconoce, sin embargo, que aunque la WSD sigue sugiriendo el potencial de ser útil, no ha demostrado beneficios significativos en aplicaciones reales (Navigli, 2009; Mihalcea y Edmonds, 2004). De hecho, según Agirre y Edmonds (2006), se ha ido creando una separación entre la investigación sobre la WSD y la investigación para crear aplicaciones de PLN. Estos autores señalan que, aunque las aplicaciones de PLN necesitan la WSD de alguna forma, tal vez su utilidad radique en una codificación implícita dentro de otro modelo, en vez de en una codificación explícita, como se ha intentado hasta ahora.

3. Legado de los diccionarios tradicionales

Los diccionarios tradicionales sirvieron desde temprano como la referencia básica de los lexicones computacionales para la gestión de WSD. Como dijimos antes, los primeros MRD fueron primero versiones electrónicas de diccionarios impresos para consumo humano. Vale la pena repasar, pues, algunos de los rasgos del tratamiento que dan los diccionarios tradicionales a la polisemia, especialmente dado que muchos de los problemas típicos de la lexicografía fueron heredados por los MRD.

Conviene recordar que el diccionario tradicional está complementado por el conocimiento lingüístico del lector: valiéndose de este, el usuario selecciona o descarta, amplía o restringe la información que le brinda el texto. En este respecto, el diccionario para

⁶⁰ La cuarta edición de SensEval (“Sense Evaluation”), llevada a cabo en 2007, cambió su nombre a SemEval (“Semantic Evaluation”); ello consigna el cambio de foco de los trabajos, que inicialmente estaba puesto en los sentidos y ahora recae sobre aspectos semánticos más amplios. El sexto ejercicio de SemEval fue en 2012.

consulta humana es fundamentalmente diferente del lexicón computacional que sirve al PLN, que tiene que ser autosuficiente en cuanto a la información necesaria para determinar el sentido de una palabra. Los programas de WSD necesitan procesar información textual no estructurada y transformarla en estructuras de datos que se pueden analizar para determinar el significado subyacente (Navigli, 2009). En una oración como “Juan *perdió* el tren”, el programa de WSD tendrá que concluir que *perder* es igual a ‘no llegó a tiempo’ y no a los otros sentidos posibles de *perder* (‘dejó de tener’, ‘desperdió’, ‘no ganó’, etc.).

Un primer aspecto importante sobre el tratamiento de la polisemia en la lexicografía tradicional es la separación del significado de la palabra en diferentes acepciones. Típicamente los diccionarios han tratado los diferentes sentidos de una palabra como unidades discretas que se presentan en listas enumeradas que tienden a ser lo más exhaustivas posible. En 1964, Weinreich señalaba ya, con respecto a la tercera edición del diccionario Webster’s, que existía una tendencia a la profusión de acepciones sin validez científica: en ese diccionario, el verbo *to turn* presentaba –entre acepciones, subacepciones y subsubacepciones– un total de 115 sentidos. Esta tendencia se mantiene: en la versión actual del DRAE, el verbo *perder* tiene 27 acepciones.

Cuando se usa el acercamiento de enumeración de sentidos discretos en la lexicografía electrónica, se habla de “lexicones enumerativos”. Sin embargo, como veremos en el apartado siguiente, para efectos de la WSD y del PLN, mientras más sentidos posibles para una palabra incluya un programa, más difícil será su desambiguación automática. Mairal y Periñán (2009a) señalan además que la mera enumeración de acepciones esconde las posibles relaciones entre sentidos; esto dificulta la creación de taxonomías, que constituyen el ordenamiento típico de los sentidos en los lexicones electrónicos. De hecho, a partir de los diccionarios tradicionales, se han producido poco más que un puñado de taxonomías limitadas e imperfectas (Ide and Véronis, 1998).

Se ha señalado también la variabilidad –o, en última instancia, la subjetividad del lexicógrafo– en la determinación de acepciones en los diccionarios (Boas, 2009c). Fillmore y Atkins (2000) demuestran que las definiciones entre diccionarios rara vez coinciden, incluso para palabras medianamente ambiguas como *whistle* (*silbar*), y entre diccionarios de tamaño y alcance similar. Asimismo, Bosque (2004b) confirma divergencias notables en la distribución de las acepciones del verbo *practicar* en varios diccionarios del español, y afirma que abundan los casos análogos. Aunque esto podría deberse a las presiones de tiempo y espacio a las que está sujeta la lexicografía tradicional, el resultado son series de entradas inconsistentes, incluso de unidades léxicas con comportamientos morfológicos, sintácticos y

semánticos similares (Mairal y Periñán, 2009a), vistas incluso en un mismo diccionario (Boas, 2009c).

Otra práctica lexicográfica tradicional que presenta dificultades para la lexicografía electrónica y la WSD es el uso de palabras polisémicas en la definición (Boas, 2009c; Ide y Véronis, 1998). Para un humano, esto supone desambiguar el *definiens* o palabra definidora mentalmente o buscar nuevamente en el diccionario, pero en el ámbito computacional, ello genera un nuevo problema de WSD. Boas (2009c) añade también la omisión de aseveraciones explícitas de datos lingüísticos que pueden ser esenciales a la WSD, por ejemplo, información sobre las valencias de palabras predicativas. Todo esto ha llevado a los lexicógrafos a revisar y reestructurar los MRD para facilitar su acceso y uso.

Finalmente, en la lexicografía tradicional el ordenamiento de la información ha sido típicamente semasiológico, es decir, de la palabra al significado, salvo en las contadas excepciones de diccionarios ideológicos. Sin embargo, como hemos comentado ya, en la lexicografía informatizada hay una tendencia a utilizar los sentidos (o conceptos) como eje organizador de la información. En este contexto, cabe destacar el *Proyecto de Definiciones Mínimas* (PDM), de Bosque y Mairal (2012ab) que, aunque todavía tiene forma de propuesta impresa, presenta un modelo de entradas léxicas relacionales que van de la definición a la voz, y no de la voz a la definición. La organización de los sentidos en este modelo de diccionario se asemeja a una taxonomía semántica con correspondencias onomasiológicas. A continuación se muestra parte de la entrada paradigmática del significado de *verdad*; la sangría refleja las nociones subordinadas que heredan información de los que lo preceden inmediatamente. A diferencia de un diccionario tradicional, este modelo nos permite comparar la manera en que entendemos un grupo de voces como *perogrullada*, *veraz*, *cierto*, *falsedad* y *embuste*. Usualmente las consideramos a cada una aparte, pero aquí las tenemos todas juntas en función de una jerarquía semántica y podemos observar los puntos compartidos y diferentes:

verdad.

1. ↓

verdad obvia: *truismo* [culto], *perogrullada* [coloquial]

que dice o parece decir la verdad: *veraz* (*persona veraz*, *documento veraz*)

 cualidad de veraz: *veracidad*

que posee o contiene verdad: *cierto* (*afirmación cierta*), *verdadero* (*hecho verdadero*), *verdad* (*No es verdad lo que dices*)

 no cierto: *falso* (*declaración falsa*)

 información falsa: *mentira* (*Lo que contó es una enorme mentira*), *falsedad* (*una sarta de falsedades*)

 mentira gruesa: *embuste* [...]

El PDM también presenta información sintagmática. Veamos parte de la entrada de la noción de ‘verdad’, en donde se presentan algunas de las colocaciones verbales que ocurren con el lema *verdad*. Como se trata del orden sintagmático, ahora el pivote es el lema en sí, y las colocaciones son las diferentes palabras que se combina *verdad* en función de diversos significados:

verdad.

2.⇒

COLOCACIONES VERBALES

aparecer la verdad: *salir a la luz, aflorar, surgir*

surgir la verdad de modo indirecto o por indicios: *trahucirse*

surgir o sobresalir la verdad entre lo que la oculta: *resplandecer* [culto, literario]

hacer patente la verdad, ponerla de manifiesto: *decirla, manifestarla, exponerla, declararla, expresarla*

manifestar la verdad siendo forzado a ello: *confesarla* [...]

Un diccionario como el PDM describe en lenguaje natural cada precisión semántica, o cada relación de colocación entre dos palabras. Esto permite que un lector humano tenga una comprensión perfecta de cada caso que se describe, pero limita las posibilidades de hacer generalizaciones sobre clases de palabras. Asimismo, dificulta la explotación del contenido para propósitos informáticos. Para ello habría que uniformar estas descripciones, usando un índice común de primitivos semánticos o conceptos.

4. Granularidad en los lexicones electrónicos

Uno de los aspectos en que se diferencian los lexicones entre sí es la “granularidad” semántica de los sentidos, que puede ser más fina o más gruesa. El nivel adecuado de granularidad dependerá de cada tarea de PLN; por ejemplo: una granularidad gruesa que solo distinga homónimos u homógrafos (por ejemplo: *esta*-adjetivo/*está*-verbo) será suficiente para colocar acentos gráficos, pero una distinción mucho más fina será necesaria para la traducción automática.

El modelo del lexicón enumerativo (Taylor, 1992) defiende la máxima representación semántica de una unidad léxica y la mínima representación del contexto. Se trata de un acercamiento que busca “la inclusión de múltiples entradas para cada uno de los sentidos posibles” de una palabra (de Miguel, 2009a: 361). Pustejovsky (1995: 39, traducción mía) identifica tres limitaciones en la representación del significado en los lexicones enumerativos:

- (i) el uso creativo de las unidades léxicas: las palabras pueden adquirir nuevos sentidos en diferentes contextos;

(ii) la permeabilidad de los significados de las unidades léxicas: los sentidos de las palabras tienen límites borrosos que permiten el solapamiento entre palabras; y

(iii) la expresión de las formas sintácticas múltiples: un único sentido de una palabra puede tener diversas realizaciones sintácticas.

Mairal (2011) elabora un poco estos tres puntos: primero, como el lexicón enumerativo presenta una enumeración finita de sentidos, no puede dar cabida a los usos creativos de una palabra; en segundo lugar, los límites muy rígidos en las definiciones de las palabras en el lexicón enumerativo impiden los solapamientos fluidos entre sentidos; y tercero, el lexicón enumerativo representa como sentidos distintos la variación en la complementación de una palabra, por lo que pierde la generalización que subyace tras las diversas formas sintácticas de una palabra.

La granularidad fina que presentan los lexicones enumerativos dificulta además los mecanismos del PLN: genera efectos combinatorios complicados; hace que las decisiones de etiquetado sean muy difíciles, incluso para lexicógrafos, lo que aumenta el número de discrepancias; y aumenta exorbitantemente la cantidad de datos necesarios. Se ha señalado que las distinciones de un diccionario tradicional serían demasiado finas para efectos del PLN (Slator y Wilks, 1990) y que, incluso, las distinciones de WordNet resultan excesivas (Palmer, 1998). Según Mairal (2011: 5): “la granularidad de los significados en los lexicones computacionales debe ser más genérica, es decir, menos significados y más amplios de los que nos encontramos en la mayoría de diccionarios tradicionales”.

Al otro extremo del continuo de granularidad semántica, se encuentra la teoría del lexicón generativo o TLG (Pustejovsky, 1995 y posterior), que aboga por una representación única subespecificada del significado, donde la interpretación específica de la unidad se logra mediante la información contextual. El significado en la TLG es producto de la cocomposición, o sea, no se refleja mediante categorías con límites rígidos. Se ve el léxico como un número finito de unidades léxicas que pueden generar un número infinito de significados; las palabras pueden usarse en un número infinito de contextos noveles, y activar diferentes sentidos o matices de sentidos. En este modelo, las informaciones subespecificadas de las unidades léxicas se superponen e interactúan de diferentes maneras en la sintaxis, con el resultado de especificar y determinar el significado en el contexto. Para el verbo *bake*, por ejemplo, se diferencian dos significados:

(a) ‘prepare with dry heat in an oven’ (el sentido de ‘crear’): *bake a cake*

(b) ‘cook and make edible by putting in a hot oven’ (el sentido de ‘calentar’): *bake a potato*.

Señala Mairal (2011) que, “mientras el modelo enumerativo almacena estos dos sentidos en el lexicón, el modelo generativo postula la creación de una única entrada léxica infraespecificada”. En este caso, según la visión de la TLG, *bake* no es un verbo polisémico, sino que los significados de sus argumentos –*cake* (artefacto) y *potato* (objeto natural)– concretan el significado del verbo.

La TLG cuenta con cuatro componentes que codifican distintas informaciones en el nivel sub-léxico, que repasaremos a continuación: la Estructura Argumental, la Estructura Eventiva, la Estructura de *Qualia* y la Estructura de Tipificación Léxica.

(1) Estructura Argumental (EA) - La Estructura Argumental (EA) presenta el número de argumentos de un predicado, la clase semántica a la que pertenecen (por ejemplo, [evento], [objeto], [individuo]) y la manera en que se realizan en la sintaxis.

(2) Estructura Eventiva (EE) - La Estructura Eventiva (EE) indica el tipo de evento denotado por un predicado. Los predicados no se tratan como entidades atómicas, sino que cuentan con una estructura interna que se puede descomponer en distintos subeventos o fases. Cada uno de estos subeventos se puede destacar en los distintos contextos sintácticos en que aparece el predicado.

(3) Estructura de *Qualia* (EQ) - La Estructura de *Qualia* (EQ) es reconocida como el componente más novel y productivo de la TLG (Batiukova (2009 a, b), de Miguel (2009 a, b), y más pertinente al estudio del léxico (Hanks, 2013). El nivel sub-léxico de la EQ contiene información potencial sobre las características fundamentales de la entidad a la que se refiere, ya sea una entidad (objeto físico o entidad abstracta), un evento o una propiedad. Hay cuatro *qualia* para las entidades:

Quale agentivo - cómo llega a existir;

Quale constitutivo - cuál es su constitución interna;

Quale télico - para qué sirve;

Quale formal - en qué se diferencia formalmente de otros objetos en un dominio más extenso.

No todas las entidades tienen todos los *qualia*. Por ejemplo, las palabras que designan artefactos manufacturados tienen un *quale* télico, pero las que designan animales, usualmente no. De otra parte, algunas palabras podrían tener más de un *quale* télico. La idea general, sin embargo, es que la combinatoria léxica activará uno o más *quales* de la palabra en cuestión. Por ejemplo, en el caso de los adjetivos *fast* y *slow*, la estructura de *Qualia* de los nombres especifica los posibles eventos asociados que ocurrirán rápida o lentamente. Por ejemplo, el rol télico de *plane* (avión) es *to fly* (volar). El rol télico de *book* (libro) es *to read*

(*leer*). Cuando *fast* se combina con *plane* está modificando el evento relacionado con él (*to fly*), lo cual produce el sentido de ‘avión que vuela rápido’, mientras que cuando se combina con *book* modifica el evento de *to read*, lo que produce el sentido de ‘libro que se lee rápido’.

(4) Estructura de tipificación léxica - La Estructura de Tipificación Léxica (ETL: Lexical Typing Structure, o Inheritance Structure, en Pustejovsky, 1995) es el nivel que explica cómo se relaciona una palabra con otras en el lexicón mental. Este nivel hace referencia a la información de la EQ, que determina las relaciones que puede mantener una palabra con otras: por ejemplo, siguiendo el ejemplo de Pustejovsky (1995), *libro* y *diccionario* comparten el rol formal como [libro], pero se distinguen en su rol télico en que el primero está destinado a ser leído y el segundo a ser consultado.

Sobre estas cuatro estructuras, operan cinco mecanismos: selección o selección pura; acomodación; coacción o coerción; explotación; y cocomposición. Estas estructuras y mecanismos intentan explicar por qué en la frase *construir un hospital*, el verbo activa el sentido de ‘institución’ de *hospital*, mientras que en *pintar el hospital* se activa el sentido de ‘edificio’.

Uno de los problemas fundamentales con el acercamiento de la TLG es que la infraespecificación del significado de una unidad léxica supone la postulación de un grupo de reglas que permitan la generación de los significados en el contexto. Nirenburg y Raskin (1996) señalan que el esfuerzo de postular todas las reglas necesarias en el modelo generativo puede superar el esfuerzo de identificar todos los sentidos en el modelo enumerativo. Además, para poner en práctica la TLG, hay que hacer un trabajo previo, que incluye identificar y codificar la información contextual clave, además de los significados resultantes.

Según Hunter y Marten (1999), aunque la TLG logra derivar nuevos sentidos léxicos, hay problemas con este acercamiento. Estos autores arguyen que la TLG carece de la expresividad necesaria para respaldar el razonamiento contextual en la interpretación de enunciados. La TLG implica un trabajo previo de establecer la sensibilidad contextual de las palabras a través de rasgos semánticos distintivos, pero no cuenta con un mecanismo lógico para unificar estos rasgos, sino que el proceso depende de la aplicación directa de reglas. Añade Mairal (2011:6): “en el caso de *bake*, no sólo debemos crear dos reglas léxicas que vinculen el sentido de creación con los artefactos y el de cambio de estado con los objetos naturales, sino que también es preciso que *cake* esté léxicamente marcado como artefacto y *potato* como objeto natural”.

En la práctica, muchos investigadores han buscado “ambiguar” los sentidos de lexicones existentes (Dolan, 1994), o sea, agruparlos en sentidos de granularidad más gruesa. De hecho, diversos estudios sobre la polisemia regular (Buitelaar, 1998a; Tomuro, 1998; Peters y Peters, 2000, etc.) usan lexicones como WordNet para crear sentidos más amplios y menos especificados que subsuman varios sentidos de granularidad más fina. Ide y Wilks (2006) y Agirre y Edmonds (2006) arguyen que los sentidos de granularidad gruesa son los únicos que tanto máquinas como humanos pueden distinguir con regularidad, y sugieren que ese tipo de distinción semántica es la que se debe tener en cuenta para mejorar el desempeño de las aplicaciones de PLN. Nirenburg y Raskin (2004), por su parte, recomiendan fusionar tantos significados como sea posible con el fin de que quede un número mínimo de significados.

En el punto intermedio del continuo de granularidad semántica se ubica FunGramKB, como veremos en el capítulo V. Como muchas bases de conocimiento léxico, el diseño de FunGramKB refleja experimentos psicolingüísticos como el de Jorgensen (1990), que demuestran que un hablante culto promedio puede identificar unos tres sentidos por cada palabra polisémica. Comparado con los diccionarios tradicionales, la granularidad de FunGramKB es gruesa; por ejemplo, los primeros tres sentidos del verbo *to know* del *Oxford Advanced Learner's Dictionary* se funden en un solo postulado de significado en FunGramKB (Mairal y Perrián, 2009a). Sin embargo, comparado con otras ontologías formales, FunGramKB se puede definir como un lexicón enumerativo robusto, dotado de una granularidad semántica fina y de una extensa red conceptual, que relaciona significado y unidades léxicas a través de los postulados de significado en la ontología (Perrián y Mairal, 2011). Por ejemplo, en la ontología SUMO, hay más de 250 unidades léxicas del inglés que se vinculan al concepto RADIATINGLIGHT (‘luz irradiante’), lo que indica una granularidad semántica bastante gruesa. Por su parte, FunGramKB reparte la noción de ‘luz irradiante’ en más conceptos, incluidos +ENERGY_00 (‘energía’), +ELECTRICITY_00 (‘electricidad’), +LIGHT_01 (‘luz’) (Perrián y Mairal, 2011), entre los cuales se reparten las unidades léxicas relativas a esos conceptos.

5. La WSD: metodologías básicas

En el campo de la lingüística computacional, la WSD se define como el problema de determinar computacionalmente qué sentido de una palabra se activa por el uso de esa palabra en un contexto particular. La WSD es esencialmente una tarea de clasificación: los sentidos de las palabras son las clases, los contextos proveen la evidencia y a cada instancia

de la palabra se le asigna una o más de las posibles clases según la evidencia que se tenga (Agirre y Edmonds, 2006: 1-2).

La precisión de los diferentes métodos de la WSD se mide con referencia a una “precisión base” (*baseline accuracy*), esto es, el algoritmo más sencillo posible que asigne automáticamente a la palabra su sentido más frecuente, sin ninguna otra consideración⁶¹. La precisión base de la WSD se encuentra, según los resultados de SemEval 2007, en 51.4%. Frente a este resultado, la precisión de la WSD automática para los homógrafos en inglés (por ejemplo *bass-peza/bass-instrumento* musical) –considerados el nivel de granularidad semántica más grueso y, por lo tanto, la gestión más sencilla de WSD– es alta: se ha mantenido típicamente sobre el 90%. Sin embargo, las distinciones de granularidad más fina todavía no se distancian tanto de la precisión base: los resultados de SemEval 2007 los ubican entre 59.1% y 69%.

Cuando hablamos de “los posibles sentidos de la palabra” y “el método para asignarle el sentido apropiado a esa instancia de la palabra” (puntos 2 y 3, en la Introducción), hacemos referencia a las fuentes de información y mecanismos que están disponibles para la WSD. Dependiendo de las fuentes de información que se utilicen, se distinguen dos acercamientos principales: los métodos basados en conocimiento (“knowledge-based WSD”) y los métodos basados en corpus (“corpus-based” o “data driven WSD”).

La desambiguación basada en conocimiento recurre a fuentes externas de conocimiento que preestablecen los sentidos posibles de una palabra (Ide y Véronis, 1998). Se trata principalmente de MRD –como LDOCE o OED–, de tesauros –como *Roget’s International Thesaurus*– y de bases de conocimiento –como WordNet, FrameNet o FunGramKB. Por su parte, la desambiguación basada en corpus o en datos trabaja directamente con datos de frecuencia y de coocurrencia de elementos lingüísticos en corpus, que no están disponibles en diccionarios o inventarios de sentidos. Estos corpus se pueden clasificar en *corpus anotados* y *corpus crudos* o *no anotados*.

Los *corpus anotados*⁶² contienen palabras que están etiquetadas con sus sentidos correspondientes. Se utilizan como recurso principal para los métodos *supervisados* de WSD, los cuales parten de la premisa de que el contexto provee la información suficiente

⁶¹ Esta medida parte de la premisa de que las palabras polisémicas tienen un sentido que ocurre con mayor frecuencia que los demás; si se cuenta con los datos de frecuencia, un método sencillo de WSD asigna el sentido más frecuente a cada palabra. Este método se usa a menudo como punto de partida para hacer comparaciones; según Gale *et al.* (1992), los sistemas deben superar este punto de partida para considerarse aceptables.

⁶² Aquí se incluyen Semcor y MultiSemCor, el DSO Corpus, Open Mind Word Expert y FrameNet Examples, entre otros.

para desambiguar las palabras. Estos métodos constituyen la única manera de medir la efectividad de los sistemas de WSD en general, y de compararlos entre sí, y son los que obtienen los mejores resultados al día de hoy. Sin embargo, anotar los corpus resulta muy costoso y laborioso, porque requieren la marcación manual de los sentidos de las palabras en textos de gran extensión.

Las palabras en los corpus *crudos* o *no anotados* no tienen etiquetas de sentidos, aunque pueden tenerlas de otros tipos, por ejemplo, de categoría gramatical. En estos corpus, se utilizan métodos de WSD *no supervisados* por humanos, que incluyen el uso de corpus alineados en diferentes lenguas para recoger evidencia multilingüe para desambiguar sentidos, así como la *inducción* o *discriminación* de sentidos, método por el cual se agrupan los contextos en los que ocurre una palabra meta y se equiparan a un sentido de la palabra⁶³.

Las estrategias para asignarle el sentido apropiado a cada instancia de una palabra son, por supuesto, el componente más complicado de la WSD. Entre los métodos existentes para la WSD, figura el uso de las preferencias de selección, ya sean léxicas, semánticas o conceptuales, al que daremos particular énfasis en este capítulo. Estas pueden servir para restringir los posibles significados de las palabras que participan en una relación léxica dada. En el caso del verbo *perder*, podríamos decir que las preferencias léxicas son indicador suficiente para deslindar varios sentidos del verbo: hay claramente una diferencia semántica entre *perder la cartera* ('dejar de tenerla'), *perder el campeonato* ('no ganar'), *perder el tiempo* ('desperdiciarlo') y *perder el tren* ('no llegar a tiempo').

Utilizar las preferencias de selección supone, sin embargo, conocer las combinaciones sistemáticas que se dan entre palabras, lo que requiere un trabajo previo significativo. REDES satisface esa necesidad, al proporcionar la combinatoria léxica sistemática de 4,000 predicados, extraída de corpus y agrupada en clase léxicas por criterios semánticos. Por supuesto, usar las preferencias de selección para la WSD requiere postular también los sentidos de las palabras que las preferencias de selección activen. Cabe recordar que REDES nos informa de 16 clases léxicas con que se combina *perder*, pero no identifica los sentidos que son activados en cada combinación. Como veremos, la traducción de los datos de REDES a FunGramKB implica codificar ambas cosas –sentidos y preferencias– a la vez.

⁶³ Los corpus no anotados más utilizados son el *Brown Corpus*, el *British National Corpus* (BNC), el *Wall Street Journal Corpus* y el *Reuters News Corpus*.

6. El rol del contexto

El contexto de uso de una palabra polisémica ofrece, en la mayoría de los casos, información suficiente para determinar su sentido en esa instancia particular. Por lo tanto, todos los acercamientos tienen que referirse al contexto de alguna manera u otra. Para los métodos basados en corpus, el contexto también provee el conocimiento previo con que se compara el contexto de uso de la palabra en cuestión.

Al usar el contexto como fuente de información para la desambiguación, los investigadores distinguen el contexto *local* o *micro*, constituido por las palabras que están cerca de la palabra polisémica, del contexto *temático* (“topic / global context”), que son las palabras sustantivas que coaparecen con un sentido de una palabra en una discusión de cierto tema. Sobre el contexto micro, los estudios sugieren que la distancia óptima variará según el tipo de desambiguación, pero el promedio parece estar cerca de las +/- 3 palabras (Leacock *et al.*, 1998). Otro acercamiento de contexto micro parte de la hipótesis de “un sentido por colocación”, impulsada por Yarowsky (1993), que dice que una palabra tiende a preservar el mismo sentido cuando aparece con la misma colocación. Yarowsky (1993) concluye que existe un sentido por colocación –en el sentido de una coaparición que es habitual o usual– en casos de ambigüedad binaria.

El contexto temático o global ofrece también información desambiguadora. Aquí se parte de la premisa de que las palabras en un discurso deben estar relacionadas semánticamente para que el discurso sea coherente. Se puede tomar en cuenta un dominio temático general o uno especializado. Por ejemplo, *racket* y *court* están asociadas temáticamente, y se desambiguan mutuamente, en sus sentidos deportivos. Ahora bien, este acercamiento es computacionalmente muy intensivo: la mayoría de las pruebas se han circunscrito a analizar la vecindad inmediata de la palabra en cuestión, y no a la totalidad temática del texto.

Otros, como Schutze (2000), usan el contexto para identificar patrones de uso similares. En vez de vincular la palabra a un sentido discreto, agrupan los contextos de uso similares en conjuntos (“clusters”) y les asignan un sentido. También ha habido esfuerzos de simplificar las concordancias de corpus; por ejemplo, Hanks (2004) parte de la premisa de que el estudio de las palabras en su contexto identifica los patrones normales de uso: un significado –real o potencial– se puede vincular a cada patrón.

Dentro del campo del análisis de patrones de corpus (Corpus Pattern Analysis o CPA), Hanks (2013) propone una teoría llamada *Theory of Norms and Exploitations* (TNE), que servirá de base a su *Pattern Dictionary of English Verbs*, proyecto que está en curso. La

próxima tabla muestra el perfil sintagmático del verbo *to urge*. Los patrones identificados en la tabla son abstraídos de corpus, y constituyen las “normas” de uso a las que hacen referencia el nombre de la teoría. Este autor señala que el corpus provee información directa sobre los patrones de uso de las palabras, pero solo información indirecta sobre sus sentidos. Arguye que para poder asignarle sentidos a una unidad léxica, primero se deben conocer los diferentes patrones sintácticos y colocacionales en los que la palabra participa. Las “explotaciones” a las que se refiere el nombre de la teoría son los casos de usos metafóricos o figurados que no caen en los patrones. En el caso de *urge*, las explotaciones suman un 10%.

Tabla 24. Perfil sintagmático de *to urge* (tomado de Hanks, 2013)

| | Lemma: <i>urge, urges, urging, urged</i> |
|-------------------|--|
| PATTERN 1 (61%) | [[Human ₁]] urge [[Human ₂]] {to-INF [V]} |
| PATTERN 2 (15.4%) | [[Human]] urge [[Action]] |
| 2.1 | [[Human ₁ Speech Act]] urge [[Action Attitude]] ((up)on [[Human ₂]]) |
| 2.2 | [[Human ₁ State of Affairs]] urge [[[Reason (for Action Attitude)]] ((up)on [[Human ₂]])] |
| PATTERN 3 (4.3%) | [[Human]] urge <i>that</i> [CLAUSE] |
| PATTERN 4 (5.4%) | [[Human]] urge [QUOTE] |
| PATTERN 5 (3.5%) | [[Human]] urge [[Steed]] [Adv[Direction]] |

Para efectos del diccionario *Pattern Dictionary of English Verbs* –la aplicación práctica de la TNE–, para determinar los sentidos de las palabras, el autor parte de los patrones y presenta las principales implicaturas asociadas con cada patrón. Estas descripciones son aproximadamente equivalentes a una definición de diccionario, pero están ancladas en los argumentos de cada patrón.

El programa WordSketch (Kilgarriff, 2006), por su parte, explota los datos de corpus para crear un boceto de la frecuencia con que ocurren colocados con una palabra dada, en las diferentes relaciones gramaticales. Esa información sirve de base a la desambiguación. Por ejemplo, el “word sketch” del sustantivo *goal* (‘gol’/‘meta’) muestra que ocurre en dos ámbitos: si se habla de deportes, *goal* es objeto de verbos como *score*, *concede*, *disallow*, *net*, mientras que si se habla de la vida, ocurre con *achieve*, *pursue*, *attain* o *reach*.

La información sintáctica se ha utilizado también como indicador para la WSD. Esta puede incluir –además de las colocaciones– información sobre categoría gramatical (“part-of-speech” o POS), morfología y subcategorización, según Agirre y Stevenson (2007). Sin embargo, según Boas (2009c), muchas bases de datos multilingües dependen exclusivamente

de información sintáctica para desambiguar las palabras, lo cual es limitante. Sin tener acceso a la información de los diferentes tipos semánticos de las frases que ocurren con la palabra en cuestión, es difícil establecer de qué sentido se trata:

Any MLLD [Multi Language Lexical Database] aimed at providing useful information for humans and machines will therefore have to include detailed syntactic and semantic valence information showing how to map specific sub-senses of a word from one language into another language (Boas, 2009c: 6).

Este repaso hace claro lo que implica usar el contexto sistemáticamente para los procesos de desambiguación. Por un lado, hay que conocer los contextos en que ocurren las palabras sistemáticamente y, por otro, hay que conocer –o establecer– los sentidos que pueden asumir. Dados estos dos requisitos, entendemos que la combinación entre REDES y FunGramKB puede ser en extremo beneficiosa, pues el primer recurso provee la información de combinatoria y el segundo, una ontología de conceptos que aspira a ser autosuficiente. La conexión entre ambos, sin embargo, supondrá un trabajo manual.

7. Las preferencias de selección

El término “preferencia de selección” captura el hecho intuitivo de que los predicados de una lengua tienen un mejor “pareo” semántico con algunos argumentos que con otros (Ó Séaghdha y Korhonen, 2012). La diferencia entre la noción de ‘preferencia de selección’ y otras, como ‘colocación’, por ejemplo, es que las preferencias hacen referencia –en general– a relaciones entre categorías de clases de palabras, y no a las palabras en sí mismas.

Dentro de los métodos de desambiguación basados en conocimiento, las preferencias de selección como medio para restringir los posibles significados de las palabras que participan en una relación léxica dada presentan un potencial atractivo, aunque todavía no supera la asignación de los sentidos más frecuentes (Milhacea, 2007)⁶⁴. Explica esta autora que las preferencias de selección capturan información sobre las posibles relaciones entre categorías de palabras, y representan conocimiento de sentido común sobre clases de conceptos. Por ejemplo, COMER-COMIDA, BEBER-LÍQUIDO son casos de restricciones semánticas que se pueden usar para descartar sentidos incorrectos y seleccionar solo aquellos que coincidan con las reglas de sentido común. En la oración *Mary drank burgundy*, el sentido de ‘color’ de *burgundy* se descarta porque el verbo *to drink* requiere un líquido como objeto directo. Pero Agirre y Martínez (2001) advierten que las preferencias pueden ser más

⁶⁴ Brockman y Lapata (2003) ofrecen un análisis detallado de los acercamientos basados en conteos de frecuencia, medidas de información-teoría, y relaciones de clase a clase adquiridas de taxonomías creadas manualmente, y los comparan con los juicios humanos.

complejas; en el caso de *eat*, además de combinarse con comestibles (*She was eating an apple*), pueden darse casos como: *The acid ate the metal*, *This car eats a lot of gas*, *We ate our savings*, etc. Estos casos, en el contexto de FunGramKB, se vincularían con conceptos diferentes: el sentido literal de *eat* se vincularía a +EAT_00, pero la combinación de *eat the metal* se vincularía a +DAMAGE_00, y las combinaciones *eat gas* y *eat savings* a +USE_00. Naturalmente, las diferentes restricciones de selección tienen que ser codificadas también, ya sea como preferencias de selección conceptuales o como colocaciones en el lexicón.

Aunque las preferencias de selección son intuitivas y ocurren de manera natural, es difícil ponerlas en práctica para resolver el problema de la WSD, apunta Milhacea (2007). Hay dos dificultades principales, que se relacionan de manera circular: por un lado, poder realizar la desambiguación depende de contar con grandes colecciones de preferencias de selección de una palabra dada; por otro lado, aprender las preferencias requiere conocimiento de los sentidos involucrados en una relación léxica. Como hemos dicho antes, en nuestro trabajo, REDES salva uno de los problemas y FunGramKB el otro: REDES provee el corpus de combinatoria léxica sistemática, clasificado semánticamente, y FunGramKB nos lleva a vincular al verbo a un concepto que no solo codifica su significado, sino que también incluye la preferencia de selección como parte de su descripción conceptual. El trabajo de hacer ese pareo, sin embargo, se realiza caso a caso, manualmente.

Las preferencias de selección se han examinado para la WSD, por medio de combinaciones de *palabra a palabra*, *palabra a clase (semántica)* y *clase a clase*. Las relaciones sintagmáticas de palabra a palabra son similares a las preferencias de selección, pero expresan las restricciones en términos de palabras en vez de tipos semánticos que generalizan sobre la serie de palabras (Resnik, 1997). Agirre y Martínez (2001) evalúan los tres tipos de preferencias y concluyen que, aunque no superan los resultados del sentido más frecuente, los modelos de clase a clase producen mejores resultados para la desambiguación que palabra a palabra y palabra a clase. A continuación discutimos cada modelo.

El trabajo de *palabra a palabra* se ha basado en conteos de frecuencia de coaparición de dos palabras. Se cuantifica el pareo semántico (“semantic fit”) entre dos palabras midiendo cuántas veces las dos palabras aparecen, en una relación determinada, en un corpus de gran tamaño. Otro acercamiento relacionado usa probabilidades condicionales para estimar el pareo semántico de una relación dada; este se puede expresar en ambas direcciones: de la palabra-1 a la palabra-2, o viceversa.

En la desambiguación de *palabra a clase semántica*, Resnik (1992, 1997) introduce la asociación de selección como medida del pareo semántico entre verbos y la clase semántica

de sus argumentos nominales. La clase semántica se asocia con un concepto de la base de conocimiento de WordNet. En las asociaciones de selección, la contribución de una clase semántica en una relación dada es cuantificada usando la contribución de todos los conceptos subsumidos en esa clase.

Agirre y Martínez (2001), por su parte, proponen un método para determinar preferencias de selección de *clase a clase*. Su motivación es, por un lado, que los argumentos nominales de los verbos pueden ser indicadores de distintos sentidos verbales y, por otro, que si los sentidos de un grupo de verbos son similares, pueden tener preferencias de selección similares; se puede entonces generalizar que una clase de verbos tiene una preferencia de selección particular. Se trata de un acercamiento probabilístico que toma datos de frecuencia de instancias de este tipo: [*nombre-sentido*]RELACIÓN[*verbo-sentido*], donde la relación es sujeto-verbo o verbo-objeto. Aunque la desambiguación se puede enfocar en ambas direcciones, los autores eligen desambiguar los nombres; estudian, además de las preferencias de clase a clase, las de palabra a palabra y palabra a clase. Aunque ningún método superó los resultados del sentido más frecuente, los modelos de clase a clase tuvieron mejores resultados que los otros dos.

Stevenson y Wilks (2001) también evaluaron las preferencias de selección, implementándolas como rasgos en un sistema más amplio de WSD. Derivaron las preferencias de selección usando: (a) los códigos semánticos de LDOCE; (b) una jerarquía construida a partir de esos códigos, que indicaba, por ejemplo, que Sólido, Líquido y Gas son tipos de Inanimados; y (c) las relaciones gramaticales, como sujeto-verbo, o verbo-objeto. Las preferencias de selección produjeron un 44% de desambiguación en un corpus anotado con sentidos de LDOCE.

Uno de los retos que presentan las preferencias de selección entre clases semánticas es precisamente la creación de esas clases semánticas. Como vimos en el capítulo I, al discutir la naturaleza de las clases léxicas de REDES, se trata de clases léxicas sensibles a la combinatoria, en donde las palabras adquieren sentidos especiales en función del predicado con que se combinen. No obstante eso, lo importante será el sentido que adquieren en la combinación. Para dar cuenta de eso, será utilísimo contar con una ontología de conceptos previamente establecida.

Conclusión

Los datos de restricción de selección léxica de REDES y el marco ontológico de FunGramKB resuelven las dos dificultades citadas por Milhacea (2007), arriba:

- (a) la necesidad de contar con grandes cantidades de datos de combinatoria, y
- (b) la necesidad de establecer previamente los sentidos que hay que desambiguar.

Los datos de REDES ofrecen información combinatoria exhaustiva⁶⁵: se trata de una gran colección de restricciones de selección tomadas de corpus, que van más allá de lo que Milhacea (2007) llama “intuitivo”, pues incluyen no solo las combinaciones predecibles, sino también las combinaciones usuales que no son predecibles. En segundo lugar, al vincular los datos de combinatoria de predicado-argumento de REDES a los conceptos correspondientes evento-entidad de la ontología, establecemos precisamente qué sentidos verbales ocurren con qué argumentos y, por lo tanto, aportamos la información que puede servir de base para la WSD.

El marco de FunGramKB nos permite aprovechar la información de REDES en dos niveles. La información de preferencias de selección conceptual se ubica en la ontología, precisamente como preferencias de selección, mientras que las combinaciones léxicas que no puedan ser descritas conceptualmente se ubican en el lexicón, como colocaciones. El vínculo que FunGramKB establece entre el lexicón y la ontología permiten que se aproveche tanto la información léxica como la conceptual, de modo que, en principio, se podría realizar la desambiguación semántica en los tres niveles: palabra a palabra, palabra a clase (vinculada a un concepto) y concepto a concepto.

⁶⁵ Una entrada de REDES no representa necesariamente todos los sentidos de la palabra, pero de los sentidos elegidos, incluye todas las combinaciones encontradas (comunicación personal entre I. Bosque y R. Mairal, abril, 2011).

CAPÍTULO V. LA POLISEMIA REGULAR

Introducción

El fenómeno que llamaremos “polisemia regular”, siguiendo la designación utilizada por Apresjan (1974)⁶⁶, presenta un problema interesante para la léxico semántica, tanto en el plano teórico como en sus aplicaciones prácticas, incluida la lexicografía electrónica. A grandes rasgos, la “polisemia regular” o alternancia regular de sentidos se refiere al fenómeno por el cual grupos de sentidos se relacionan de manera sistemática y predecible, y se manifiestan del mismo modo en grupos de unidades léxicas.

Entre los casos más citados de sentidos que mantienen relaciones sistemáticas entre sí, y sus manifestaciones léxicas nominales, figuran los que se presentan en la siguiente tabla. Nótese que se necesitan por lo menos dos sentidos que alternan, y por lo menos dos palabras que posean esos dos sentidos:

Tabla 25. Ejemplos de polisemia regular

| Sentidos que alternan | Manifestaciones léxicas |
|---------------------------------|--|
| (1) ‘contenedor’ ~ ‘contenido’: | <i>vaso, taza, caja, baúl</i> |
| (2) ‘edificio’ ~ ‘institución’: | <i>escuela, universidad, instituto</i> |
| (3) ‘animal’ ~ ‘alimento’: | <i>pollo, codorniz, atún</i> |
| (4) ‘lugar’ ~ ‘sus habitantes’: | <i>ciudad, pueblo, Nueva York</i> |
| (5) ‘animal’ ~ ‘su piel’: | <i>cocodrilo, conejo</i> |
| (6) ‘planta’ ~ ‘alimento’: | <i>plátano, aguacate</i> |
| (7) ‘sustancia’ ~ ‘color’: | <i>jade, ámbar</i> |
| (8) ‘fruta’ ~ ‘color’: | <i>naranja, melocotón</i> |
| (9) ‘producto’ ~ ‘productor’: | <i>Toyota, periódico</i> |
| (10) ‘música’ ~ ‘baile’: | <i>vals, rumba, salsa</i> |

La polisemia regular captura procesos que están en gran medida lexicalizados y que son, por lo tanto, recurrentes. Como se muestra en la lista arriba, sabemos que las palabras *vaso, taza, caja, baúl* pueden significar ‘contenedor’ (*eché el agua en el vaso*) o ‘contenido’ (*dame un vaso de agua*). Estos patrones pueden replicarse en otros miembros de una clase de

⁶⁶ Además de *polisemia regular*, este fenómeno ha recibido otros nombres, entre los que destacan *polisemia productiva* (Apresjan, 1974), *polisemia sistemática*, *polisemia lógica* o *metonimia lógica* (Pustejovsky, 1991, 1995). Nunberg y Zaenen (1992) recogen otros términos usados por los investigadores para referirse a este fenómeno, como: *referencia deferida* (Nunberg, 1979); *reglas de transferencia semántica* (Leech, [1974] 1981); *transferencia de sentido* (Sag, 1981); *conectores* (Fauconnier, 1985); *extensiones de sentido* (Pustejovsky, 1991; Briscoe and Copestake, 1991; Copestake and Briscoe, 1991); *redes léxicas* (Norvig & Lakoff, 1987); *subregularidades* (Wilensky, 1991); *reglas de implicación léxica* (Ostler and Atkins, 1991).

palabras, en ese caso, *maleta*, *bulto*, *mochila*. Otro ejemplo lo constituye el patrón de alternancia entre los sentidos de ‘fruta’ y ‘su color’ (*una deliciosa naranja --> una camisa naranja*), que podría en principio extenderse a otras frutas (*un delicioso melocotón --> una camisa melocotón*). Sin embargo, sabemos también que no siempre es posible (*una deliciosa pera --> *una camisa pera*). Los sentidos que alternan en la polisemia regular son –en general– independientes de la situación comunicativa y de las restricciones pragmáticas del discurso.

La polisemia regular se ha documentado en suficientes lenguas como para concluir que es un fenómeno lingüístico universal, pero su manifestación léxica no siempre es igual. Existen casos en que la variación léxico semántica se manifiesta de forma paralela en varias lenguas. Por ejemplo, Wing (2002) ilustra que las palabras *iglesia* en español, *church* en inglés y *kerk* en holandés comparten la alternancia ‘edificio’~‘institución’ (Ver sección 5.5 de este capítulo). Otros advierten, sin embargo, que hay casos en que las diferentes lenguas no concretan léxicamente la distinción de sentidos en palabras equivalentes, sino a través de otros procesos como la derivación y la composición.

La polisemia regular ocurre tanto en nombres –la categoría gramatical más estudiada– como en verbos, adjetivos, adverbios y hasta preposiciones (Buitelaar, 1998a). Los trabajos hechos con verbos (Levin, 1993; Pustejovsky y Busa, 1995) tienden a enfatizar las diferencias aspectuales del verbo, por ejemplo:

- (a) *The enemies sank the boat* (causativo)
- (b) *The boat sank* (incoativo)

El trabajo de Pustejovsky y Boguraev (1993) estudia la polisemia regular en adjetivos, con el caso de *fast*. Aquí la alternancia de sentidos radica en que *fast* destaca diferentes aspectos del evento que está implicado en el sustantivo, principalmente relativos a su *quale* télico⁶⁷ (siguiendo la TLG):

- (a) *fast typist* (--> teclea rápido)
- (b) *fast game* (--> ocurre rápido)
- (c) *fast book* (--> se lee rápido / --> se escribe rápido)
- (d) *fast driver* (--> conduce rápido)

También se han hecho trabajos con palabras con contenido funcional, como las preposiciones, por ejemplo, el estudio sobre “over” de Brugman (1988).

Tomuro (1998) afirma que la polisemia regular que ocurre en cada categoría es

⁶⁷ Recordemos que el rol télico tiene que ver con la función inherente de algo o el propósito de un evento.

diferente: para los nombres, cada sentido polisémico usualmente se refiere a un objeto diferente. Por ejemplo, *escuela* puede referirse a una ‘institución’ o a un ‘edificio’. Para los verbos, los sentidos polisémicos se refieren a diferentes aspectos de la misma acción. Por ejemplo, el verbo *write* en *John wrote a book* es ambiguo entre el sentido de ‘creación’ (del libro) y ‘comunicación’ (a través del contenido del libro), aunque los dos describen la misma acción de que John escribe el libro. Los dos sentidos están relacionados sistemáticamente al referirse al aspecto de *causación* (‘creación’) o al aspecto de *propósito* (‘comunicación’) de la acción de *write*.

La representación de la polisemia regular en una base de conocimiento es una herramienta de sumo valor para la lexicografía moderna y el PLN. Ello se debe a que una vez se identifica un patrón de polisemia regular, por ejemplo ‘edificio’~‘institución’, este se puede extender a todas las unidades léxicas que tengan uno u otro sentido, por ejemplo, *escuela, universidad, alcaldía*, etc. En otras palabras, si vemos la polisemia regular como una fuente de creatividad léxica, podríamos predecir el uso de sentidos noveles a partir de los sentidos establecidos en patrones identificados. Siguiendo el ejemplo anterior, podríamos postular que “cualquier palabra que tenga el significado de ‘edificio’, también tendrá el significado de ‘institución’”. Por otro lado, si se establece que los dos sentidos pueden ocurrir simultáneamente, o sea, si no es necesario precisar el significado de una unidad léxica, el recurso de desambiguación los podría aceptar como una clase semántica subespecificada. En el caso del ejemplo mencionado, estaríamos ante una clase semántica llamada ‘edificio-institución’. Estas posibilidades representarían un ahorro significativo en el esfuerzo de asignar sentidos a las unidades léxicas y una expansión significativa de la cobertura semántica del recurso de PLN. Para las bases de datos multilingües, el alcance de la predicción se multiplicaría.

Como parte de nuestra tesis, exploramos un grupo de cinco verbos que manifiesta el fenómeno de la polisemia regular: *tramar, tejer, urdir, maquinar y cocinar*, y examinamos cómo se codifica este fenómeno en la base de conocimiento FunGramKB. Algo particular de nuestra investigación es que miraremos la relación entre la polisemia regular de estos verbos con su combinatoria léxica sistemática. Además, no nos enfocaremos en la alternancia aspectual de los verbos, sino en su alternancia semántica, según queda plasmada en los vínculos con eventos conceptuales en FunGramKB.

En este capítulo, repasaremos los acercamientos a algunas de las principales preguntas formuladas en torno a la polisemia regular, y dedicaremos bastante atención a los estudios

hechos sobre polisemia regular en lexicones electrónicos. La sección 1 hace un repaso de los principales acercamientos generales a la polisemia regular. La sección 2 considera la naturaleza de la relación que media entre los sentidos que alternan. La sección 3 mira las clases de palabras como entidades importantes en la polisemia regular. La sección 4 considera las posibles motivaciones que activan los sentidos que alternan. La sección 5, que abarca la segunda mitad del capítulo, se dedica a los estudios que se han llevado a cabo sobre la polisemia regular en lexicones electrónicos para el PLN.

1. Acercamientos generales al estudio de la polisemia regular

Según Peters y Peters (2000), Aristóteles hablaba ya de los sentidos que se relacionan de modos sistemáticos y predecibles al discutir tropos lingüísticos como la metonimia y la sinécdoque. Sin embargo, el estudio moderno de la polisemia regular empieza con los trabajos etimológicos diacrónicos, como el de Stern ([1931] 1964), que puso énfasis en la conexión histórica entre los sentidos que alternan. Los estudios sincrónicos modernos comienzan con Weinreich (1964), que hace una analogía entre la polisemia regular y la neutralización fonológica en ciertos contextos (en la que se acepta que ocurre ambigüedad de rasgos fonológico). Weinreich propone un acercamiento similar al de la fonología en el manejo de los rasgos de significado, aunque reconoce que el problema semántico es más complicado, porque el número de elementos básicos es muy grande y los contextos muy difíciles de clasificar.

Weinreich (1964) establece una distinción importante entre dos tipos de polisemia: la “complementaria” y la “contrastiva”. La “polisemia complementaria” –que llamamos aquí polisemia regular– refleja sentidos que se relacionan de manera sistemática, como los anotados en la lista de arriba. Por su parte, la “polisemia contrastiva” se refiere a los homónimos, sentidos que no tienen relación entre sí, como:

(1) *bank*: ‘institución financiera’

(2) *bank*: ‘bajo cercano a la orilla en cuerpos de agua’⁶⁸.

Entre los estudios sincrónicos sobre la polisemia regular, el trabajo de Apresjan (1974) es un hito fundamental. Sin duda, la definición más citada de este fenómeno es la suya:

⁶⁸ Algunos investigadores atribuyen la homonimia a orígenes etimológicos diferentes, mientras que otros la equiparan a una marcada distancia semántica. Sobre el famoso caso de *bank*, la lexicografía inglesa en general valida la noción de que los dos sentidos de *bank* tienen orígenes etimológicos diferentes, mientras que en español, los dos sentidos de *banco* se presentan como relacionados por derivación en fuentes como Corominas (1954) y DRAE (2001).

Polysemy of the word A with the meaning a_i and a_j is called regular if, in the given language, there exists at least one other word B with the meanings b_i and b_j , which are semantically distinguished from each other in exactly the same way as a_i and a_j and if a_i and b_i and a_j and b_j are non-synonymous (Apresjan, 1974: 16).

Este autor también describe la polisemia irregular: “Polysemy is called irregular if the semantic distinction between a_i and a_j is not exemplified in any other word of the given language” (Apresjan, 1974:16). A partir de sus observaciones del ruso, Apresjan identifica diferentes alternancias de sentido en diferentes categorías gramaticales; por ejemplo, en los sustantivos, señala las siguientes alternancias:

‘acción’ ~ ‘sujeto de la acción’
‘acción’ ~ ‘objeto de la acción’
‘acción’ ~ ‘resultado de la acción’
‘acción’ ~ ‘medios para realizar la acción’
‘estado’ ~ ‘causa del estado’
‘propiedad’ ~ ‘sujeto de la propiedad’
‘acción’ ~ ‘sujeto de la acción’

En los verbos, señala que los sentidos de ‘causa’, ‘extracción’ ~ ‘eliminación’, ‘remoción’ ~ ‘procesamiento’, ‘deformación’ participan en alternancias frecuentemente. Dice que casi cualquier combinación de estos sentidos es posible. Naturalmente, dependiendo de los verbos que participen, estos sentidos generales toman matices más específicos, por ejemplo: ‘procesar algo de un modo’ ~ ‘eliminar algo por procesarlo de ese un modo’. En los adjetivos, los ejemplos que da Apresjan son del tipo ‘relativo a X’ ~ ‘causado por X’, o ‘expresando X’ ~ ‘causando X’.

En tiempos más recientes, ha habido dos acercamientos principales a la polisemia regular: uno semántico y uno pragmático (Buitelaar, 1998a). La diferencia entre ellos tiene que ver con la relación que se postule entre las unidades léxicas polisémicas y sus referentes en el mundo. La perspectiva semántica, donde Buitelaar (1998a) incluye a Bierwisch (1982) y a Pustejovsky (1995), entre otros, ubica la relación en el lenguaje mismo, ya sea entre unidades léxicas o entre aspectos semánticos de las unidades léxicas. Así, se reconoce que una unidad léxica polisémica, como *escuela*, puede adquirir diversos sentidos: ‘un edificio’, ‘una institución’, ‘un proceso educativo’ o ‘un grupo de gente’. La visión pragmática, en la que Buitelaar (1998a) incluye a Nunberg (1979), Ruhl (1989) y otros, ubica la relación entre objetos en el mundo. Desde esta perspectiva, la unidad léxica apunta a un objeto, y ese objeto apunta a otro objeto; la relación existe pues entre los dos objetos y es independiente de la palabra. En el caso de *escuela*, la palabra apunta a un objeto, que puede ser la institución,

y ese objeto apunta a otros, como el edificio, la gente y el proceso educativo. La importancia de la perspectiva que se asuma es que tendrá repercusiones en la representación: un acercamiento semántico almacenará la mayor información posible en la entrada léxica, mientras que uno pragmático tendrá una entrada escueta y dependerá de inferencias hechas a partir de principios o convenciones globales (Nunberg, 1979).

Nuestro acercamiento a la polisemia regular será semántico conceptual, dado que investigaremos su representación en la base de conocimiento FunGramKB, específicamente en la ontología conceptual. La mayoría de los estudios sobre la polisemia regular que se han hecho en el ámbito del PLN –que discutiremos en la sección 5– también son semánticos, aunque el tipo de unidad semántica específica depende de las bases de conocimiento que se utilicen.

2. Relación entre los sentidos

Una de las preguntas básicas sobre el fenómeno de la polisemia regular es la naturaleza de la relación que existe entre los sentidos que alternan. A partir de Apresjan (1974), la mayoría de los investigadores la ven como un tipo de metonimia, mientras que otros hablan de una relación de dependencia. Elaboramos las dos perspectivas a continuación.

Apresjan (1974) argumenta que la polisemia regular está motivada por la metonimia, mientras que la polisemia irregular responde a procesos metafóricos. Explica que la regularidad semántica vista en la polisemia regular es un rasgo distintivo de las transferencias metonímicas, mientras que la transferencia metafórica expresa una relación de similitud semántica entre dos conceptos. Este autor añade que hay otros procesos que pueden dar pie a la polisemia regular, como la analogía semántica y la compresión de frases, y compara estos procesos con la formación de palabras porque, en ambos casos, existe un buen número de tipos productivos. Peters y Peters (2000) coinciden en que la polisemia regular surge de la metonimia, la cual expresa una contigüidad semántica derivada del conocimiento de mundo. Peters (2006), por su parte, explica que la polisemia regular es una figura retórica no literal en la que el nombre de una cosa es sustituido por el nombre de otra cosa relacionada con ella; existe pues una relación semántica entre dos conceptos que están asociados con el mismo nombre.

Otros investigadores describen la relación que existe entre los sentidos como una de dependencia. Copestake & Briscoe (1995), por ejemplo, introducen la noción de “extensión

de sentido”, que establece la existencia de un sentido primario que causa la existencia de los sentidos secundarios dependientes. Tomuro (1998) también aboga por la noción de dependencia entre los sentidos, y opina que en el caso del nombre *escuela*, el sentido de ‘institución’ se debe considerar primario y el de ‘edificio’ secundario, porque aunque las instituciones usualmente tienen espacio de oficina, los edificios pueden albergar otras entidades aparte de instituciones. Asimismo, el verbo *escribir* tiene un sentido de ‘creación’ que él considera primario y uno de ‘comunicación’ que considera secundario, dado que el escribir necesariamente crea un objeto, que se puede o no usar para comunicar.

Otros autores desglosan las relaciones entre los sentidos que alternan. Barque & Chaumartin (2006), por ejemplo, formulan un análisis en donde existen tres relaciones posibles que dan pie a la polisemia regular, la especialización, la metáfora y la metonimia:

(1) Especialización: una unidad léxica L2 es una especialización de una unidad léxica L1 si su significado es más específico que el de L1. La unidad léxica *pressure#7* denota un tipo particular de presión (*pressure*), específicamente la que ejerce la atmósfera:

{*pressure#7*} = the pressure exerted by the atmosphere
 [{*presión#7*} = la presión ejercida por la atmósfera]⁶⁹

(2) Metáfora: dos unidades léxicas L1 y L2 están en una relación metafórica si los referentes de L1 y L2 están en una relación de analogía, esto es, si son similares en al menos un aspecto. Por ejemplo, la risa denotada por *cackle#3* se parece al cacareo (*cackle*) de una gallina, según se consigna en la definición siguiente:

{*cackle#3*} = a loud laugh suggestive of a hen’s cackle
 [{*cacareo#3*} = una risa sonora sugestiva del cacareo de una gallina]

(3) Metonimia: dos unidades léxicas L1 y L2 están en una relación metonímica si los referentes de L1 y L2 están en una relación de contigüidad o de contacto, en términos concretos o abstractos. Por ejemplo, la relación entre los dos sentidos de *chestnut* (color y fruta) es metonímica dado que el color denotado por *chestnut#4* es el color de la fruta denotada por la castaña (*chestnut*) en la definición:

{*chestnut#4*} = the brown color of chestnut
 [{*castaño#7*} = el color marrón de las castañas]

Aparte de la relación general que medie entre los sentidos –especialización metáfora, metonimia–, algunos investigadores han intentado caracterizar la relación en términos

⁶⁹ Proveemos entre corchetes la traducción al español de los ejemplos.

semánticos. Tomuro (1998), por ejemplo, les da nombre a las relaciones entre los sentidos: en la polisemia de la palabra *escuela*, la relación entre los sentidos de ‘edificio’ e ‘institución’ es de ‘ubicación física’. Peters (2006), por su parte, señala que hay tres maneras de caracterizar la relación entre los sentidos. Primero, se puede hacer una asociación general del tipo ‘está relacionado con’, que se limita a establecer que hay una relación no especificada. Segundo, las relaciones se pueden describir mediante la introspección humana, comparando los sentidos participantes y dándole nombre a la relación; sin embargo, esto requiere bastante tiempo. Una tercera opción, que Peters (2006) explora en su trabajo, es automatizar el proceso: este investigador explota la información disponible en las glosas semánticas de WordNet para extraer relaciones semánticas explícitas entre los sentidos involucrados. Por ejemplo, de palabras como *football*, *handball*, *baseball*, que alternan entre los sentidos ‘bola usada en juego X’ y ‘juego X’, se extrae la relación ‘usada en’.

Si postulamos que, entre los dos sentidos, existe un significado primario y un significado que se desprende de este, cabe preguntarse en qué dirección opera el fenómeno. Según Peters & Peters (2000), la relación de sentidos parece funcionar en una dirección principalmente –de un sentido base o prototípico a un sentido derivado, y no necesariamente en ambas. Por ejemplo, las palabras que designan un animal pueden asumir a veces un sentido de alimento (*chicken* puede referirse al animal y al alimento), pero la relación inversa no se sostiene (*mutton* se refiere al cordero cocido, pero no al animal). Por ello, defienden esos autores, merece la pena identificar el sentido base de una pareja de sentidos.

Fillmore y Atkins (2000), por su parte, exploran los principios generales de extensión por metonimia, o de extensión de un sentimiento (*I am extremely sad --> It is a sad day*). Sin embargo, se expresan en contra de formular condiciones bajo las cuales estos mecanismos generales apliquen. Su análisis contrastivo del inglés y francés demuestra que no hay un principio productivo que determine cuándo operan mecanismos generales para extender ciertos sentidos. Señalan además muchas extensiones de sentido que resultan de casos específicos de polisemia, como la que deriva el significado de ‘cercanía al piso’ del inglés *crawl*, o la que deriva el concepto de ‘esparcirse’ del francés *ramper*.

En nuestro acercamiento a la polisemia regular, partimos de cinco verbos relacionados semánticamente entre sí y vinculamos sus sentidos a la ontología conceptual, además de comparar sus restricciones de selección léxica para cada uno de los sentidos. Aunque nuestro objetivo no era caracterizar la relación que mediaba entre los sentidos, la relación resultó ser de naturaleza metafórica. Los verbos (*tramar*, *tejer*, *urdir*, *cocinar* y *maquinar*) tienen un

sentido recto base, que se refiere a realizar un trabajo algo intrincado, y tienen un segundo sentido derivado metafóricamente, que se refiere a construir un plan malicioso, o maquinan.

3. Clases de palabras

Algunos investigadores ven una motivación para la alternancia de sentidos en las clases léxico semánticas que participan en ella, por lo que postulan las clases de palabras como entidades importantes en la polisemia regular. Es natural que se postulen clases, dado que para que la polisemia se considere regular tiene que ocurrir en conjuntos de palabras, por ejemplo, el que las palabras similares a *libro* (*artículo, revista, etc.*) muestran alternancias similares entre los sentidos de ‘objeto’ y ‘texto’, o que los sustantivos deverbales como *construcción, cooperación, separación* describan tanto una acción como un resultado (Ravin & Leacock, 2000)⁷⁰. Por su parte, en su trabajo sobre clases de verbos, Levin (1993) se basa en las propiedades sintácticas y semánticas compartidas para proponer clases de verbos cuyas características pueden usarse para predecir miembros adicionales.

En el ámbito del PLN, también ha habido intentos de formular clases para explicar la polisemia regular. Como veremos más adelante en el capítulo, varios investigadores que trabajan con los sentidos de WordNet formulan clases semánticas subespecificadas (Buitelaar, 1998a; Tomuro, 1998; etc.). Se trata de clases semánticas abstractas que agrupan sentidos que se relacionan de maneras sistemáticas y predecibles en una serie de unidades léxicas. Por ejemplo, en CoreLex (Buitelaar, 1998a), la clase ‘state’ incluye nombres que son a la vez ‘state’ y ‘time_period’, como *adolescence, adulthood, babyhood, bachelorhood, childhood, day, drought, eternity*. Martínez *et al.* (2011) también defienden la existencia de clases o tipos semánticos complejos, identificándolos como los *dot types* de la TLG de Pustejovsky (1995). El tipo complejo o *dot type* recoge palabras que son miembros de más de una clase semántica, por ejemplo, el tipo complejo Location/Organization tiene miembros que pertenecen tanto a Location como a Organization e incluye principalmente nombres de países.

Un problema que se ha señalado con el establecimiento de clases de palabras en la polisemia regular es que frecuentemente no se examina el comportamiento de las clases *a priori*, sino que se postula la existencia de las clases después de percibirse la polisemia

⁷⁰ Esta última distinción la hemos visto tradicionalmente en definiciones de diccionarios codificadas como “Acción y efecto de”.

regular. En el ámbito del PLN, estas clases son construidas a partir de inventarios de sentidos particulares, lo que genera clases léxico semánticas que resultan incompatibles con otras clases propuestas siguiendo otros inventarios (Mairal, 2011). En ambos casos, el resultado son clases dependientes de un fenómeno, en vez de clases establecidas por rasgos inherentes. Esto se refleja en el hecho de que los miembros de estas clases no necesariamente son estables. Hay palabras que podrían en principio pertenecer a ellas, pero en la práctica no lo hacen. Por ejemplo, Fellbaum (2000) estudia la polisemia irregular en verbos que manifiestan autotroponimia⁷¹, un tipo de polisemia lexicalizada que es independiente de contexto, de realizaciones sintácticas o de membresía a clase semántica, y concluye con una advertencia en contra de hacer generalizaciones sobre la polisemia en los verbos, por considerarla impredecible.

El trabajar con FunGramKB salva algunos de estos problemas: por un lado, los patrones de polisemia regular que emerjan reflejarán alternancias entre conceptos razonados *a priori* del trabajo lingüístico, compatibles con otras ontologías. En ese sentido, en nuestro estudio, la polisemia regular se reflejará en una alternancia sistemática entre dos conceptos ontológicos. Además, nuestra tesis no solo estudiará un grupo de verbos que exhibe alternancia de sentidos, sino también las clases de argumentos que seleccionan dichos verbos en cada sentido. Ello nos permitirá observar si la alternancia de sentidos se corresponde con una alternancia en selección de argumentos. Todo esto, por supuesto, estará reflejado en el nivel conceptual de FunGramKB.

4. Activación de sentidos

Una de las preguntas recurrentes sobre la polisemia regular es qué mecanismos activan un sentido u otro. En general se ha aceptado que los detonantes se encuentran en el contexto. Como vimos en el capítulo anterior, hay acercamientos a la desambiguación de las palabras polisémicas que enfocan diferentes componentes del contexto: contexto micro, contexto tópico, colocaciones, etc. Para la polisemia regular, repasaremos tres estrategias distintas de explotación del contexto: la subespecificación y composición, las reglas léxicas y la sintaxis.

⁷¹ En la troponimia, como mencionamos en el Capítulo II al discutir WordNet, un verbo expresa una manera específica de elaborar la acción expresada por otro verbo; X es un tropónimo de Y si la acción de X es la acción de Y hecha en cierta manera: por ejemplo, existe troponimia entre *hablar* y *susurrar*, *balbucear*, *gritar*. La autotroponimia es el mismo fenómeno cuando ocurre en una misma unidad léxica, como en el caso de *comportarse*, que puede significar ‘comportarse’ o ‘comportarse bien’.

(a) Subespecificación y composición

En la TLG, Pustejovsky (1995 y posterior) propone nuevas maneras de organizar y representar la información léxica para explicar la polisemia regular a través de mecanismos generativos. Como vimos antes, este autor establece significados subespecificados que se concretan en composición con los elementos del contexto. Por ejemplo, la estructura del nombre *book* contiene tres argumentos: uno para el objeto físico (x), uno para el contenido (y) y un tercero para la combinación de los dos, llamado un tipo complejo o *dot type*, que se escribe x.y. Ahora bien, los *qualia* de *libro* determinan las relaciones que estos argumentos pueden tener entre sí o con otros componentes semánticos en su contexto. Por ejemplo, el *quale* formal especifica que (x) sostiene a (y). El *quale* télico especifica el propósito y función del libro: ser leído por un agente, que aplica al contenido (y) y al concepto combinado (x.y). A través de los mecanismos de cocomposición en contexto, se activa uno u otro argumento y uno u otro *quale*. Por ejemplo, en *This book is broken* se destaca el argumento (x) y el *quale* formal, mientras que en *This book is funny* se destaca el argumento (y) y el *quale* télico. En *I cherish this book*, se activa el tipo complejo (x.y) y ambos *qualia*. En principio, esa alternancia debería ocurrir en otras palabras que designen objetos concretos de lectura, como *periódico* o *revista*, dando pie a una instancia de polisemia regular.

(b) Reglas léxicas

Copestake & Briscoe (1995) distinguen dos tipos de polisemia regular: la polisemia constructiva, un tipo de modulación de sentido donde la unidad léxica especializa contextualmente un sentido particular, y la extensión de sentidos, que relaciona predeciblemente dos o más sentidos de una misma unidad léxica.

Los autores tratan la polisemia constructiva como una entrada léxica subespecificada. En el caso de la especialización contextual, por ejemplo la palabra *reel* –‘a container artifact with the purpose of (un)winding’–, el material que se enrolla está subespecificado. Frases como *cotton reel*, *film reel* y *fishing reel* especifican el material en cuestión. Pustejovsky (1991, citado por Mairal, 2011) expresa que en estos casos la polisemia es más aparente que real, porque léxicamente existe sólo un significado y el proceso de cocomposición sintagmática es el que causa esa modulación.

En el caso de la extensión de sentido, que puede ser metonímica o metafórica, se formulan reglas léxicas. Las reglas léxicas se estipulan con la idea de que se apliquen a clases específicas de unidades léxicas para cambiar su significado de alguna manera.

Tomando como ejemplo una extensión metonímica, la regla tendría que explicar cómo las unidades léxicas que denotan un lugar (*The village is pretty*) pueden referirse a las personas que habitan ese lugar (*The village voted against the proposal*). En una extensión metafórica, la regla explicaría cómo las unidades léxicas que se refieren a animales pueden utilizarse para denotar seres humanos (*John is a pig*).

Las reglas léxicas se interpretan típicamente como relaciones condicionales entre entradas léxicas. Para describir el hecho de que los verbos como *dress* pueden recibir interpretación causativa/incoativa, la regla tomaría los verbos intransitivos de cuidado corporal como su *input* (correspondiente a la lectura incoativa) y derivaría verbos transitivos como su *output* (correspondientes a la lectura causativa).

Briscoe y Copestake (1999) arguyen que las reglas léxicas que generan la polisemia regular son semiproductivas, dado que pueden estar sujetas a bloqueo. El bloqueo ocurre cuando la existencia de una palabra evita la aplicación de una regla productiva que daría pie a una palabra con la misma semántica de la palabra ya existente, o provoca hiatos léxicos arbitrarios o grados variables de convencionalización. Por ejemplo, en el caso de nombres de animales que se vuelven comida, que también son nombres contables que se vuelven no contables, el proceso es productivo para *lamb*, *chicken*, *rabbit*, pero está bloqueado para *cow* por la existencia de un sinónimo: *beef*. Briscoe y Copestake (1999) arguyen que la productividad de las reglas léxicas debe ser estimada empíricamente mediante estudio en corpora.

(c) La sintaxis

Dowty (2000) se destaca entre los investigadores de este tema propone las estructuras sintácticas como principio explicativo para las alternancias de significado. Por ejemplo, el verbo *swarm* exhibe diferencias semánticas en estas dos estructuras sintácticas: *Bees swarm in the garden* (con sujeto agente) / *The garden swarms with bees* (con sujeto locativo), o *Roaches crawled on the wall* (con sujeto agente) / *The wall crawled with roaches* (con sujeto locativo). Dowty agrupa los verbos que participan en este tipo de alternancia –una estructura con sujeto agente y otra con sujeto locativo–, y establece cinco clases semánticas. Este autor propone que estas estructuras sintácticas están correlacionadas sistemáticamente con diferencias semánticas de cierto tipo de verbos. Las propiedades sintácticas de la estructura locativa transmiten un significado específico al verbo: el sujeto locativo convierte la ubicación en el tema del discurso, mientras que el predicado le asigna una propiedad

abstracta, que está reforzada por el uso de plurales y términos incontables (Ravin y Leacock, 2000).

En nuestro trabajo sobre polisemia regular, compartimos la idea de la subespecificación de Pustejovsky, aunque nuestro análisis no estará basado en la TLG. En nuestros cinco verbos –*cocinar, maquinar, tramar, urdir, tejer*– hay al menos dos sentidos potenciales básicos que se actualizan en contexto. Si el argumento es una palabra concreta (*cocinar arroz, tejer un traje*) los verbos adquirirán un sentido general de ‘confeccionar’, pero si el argumento es abstracto (*cocinar un plan, tejer un complot*), el sentido será cercano a ‘conspirar’. Los datos de REDES, sin embargo, nos permitirán ir mucho más allá de la distinción concreto/abstracto, para lograr especificar semánticamente los argumentos.

5. La polisemia regular en la lexicografía electrónica

Desde la década de los 1990, diversos investigadores han intentado elaborar métodos para identificar o extraer –de la manera más automática posible– patrones de polisemia regular en bases de datos electrónicas. Las definiciones de polisemia regular que utilizan varios de estos investigadores (Buitelaar, 1998a; Tomuro, 1998; Lapata, 2001a; Peters y Peters, 2000) se acercan a la siguiente expresión: ‘las alternancias de sentido regulares y predecibles a las que ciertas clases de palabras están sujetas’. Como veremos, la tendencia marcada en estos estudios ha sido a colapsar sentidos, es decir, reducir la granularidad de los lexicones, antes de llevar a cabo la desambiguación (Martínez *et al.*, 2011).

Muchos de los investigadores trabajan sobre WordNet, explotando de diversas maneras los sentidos (*synsets*), la jerarquía taxonómica que los organiza y las glosas que los definen. WordNet reconoce las relaciones de polisemia regular y agrupa las palabras que la exhiben como “primos” (*cousins*). El criterio para ser “primos”, por ejemplo, *magazine* y *newspaper*, es compartir los mismos *synsets* superordinados, en este caso, *publication, publishing house* y *product* (Boas, 2005). El propósito de los trabajos con WordNet es construir mecanismos de inferencia que activen el potencial semántico de las palabras, con el objetivo de ampliar la comprensión de textos por parte de las computadoras, además de enriquecer las relaciones entre los sentidos de WordNet de una manera explícita.

A este enfoque, Mairal (2011) lo llama “dinámico” y “distribucional”, porque la polisemia regular se detecta según la posición que tengan los sentidos en la jerarquía taxonómica. Según este autor, este es el método más deseable para el PLN, porque aumenta el nivel de integración de los sentidos en la ontología léxica. Lo contrasta con el enfoque que

llama “estático”, que simplemente explicita los otros sentidos posibles para cada sentido de una palabra. El enfoque estático también está presente en las reglas léxicas que conectan categorías conceptuales.

En esta sección, repasaremos varios trabajos hechos sobre lexicones enumerativos con métodos dinámicos o distribucionales. Como veremos, todos tienen la tendencia de colapsar sentidos. Primero repasaremos los trabajos hechos sobre WordNet y EuroWordNet: Buitelaar (1998a, 1998b), Tomuro (1998), Peters y Peters (2000), Peters y Wilks (2001), Wing (2002), Peters (2004), Peters (2006) y Barque y Chaumartin (2006). Finalmente, veremos un estudio sobre el lexicón BDéf (Barque, 2007).

5.1 Buitelaar (1998a, 1998b)

El proyecto CoreLex de Buitelaar (1998a) es pionero entre los esfuerzos de elaborar sistemas automáticos de identificación de casos de polisemia regular en lexicones electrónicos. El propósito de CoreLex es identificar clases polisémicas sistemáticas en WordNet 1.5 de manera semiautomática. Para ello, el autor parte de los nombres polisémicos en WordNet. En vez de asignarles una enumeración de sentidos discretos, los vincula a tipos semánticos subespecificados previamente establecidos. Cuando hay series de nombres que se vinculan a los mismos tipos semánticos, se dice que comparten patrones de polisemia regular. Esos nombres se agrupan en clases polisémicas sistemáticas. Su metodología siguió tres pasos:

- (a) Reduce los sentidos de WordNet a un grupo de 39 “tipos básicos”, como ‘animal’, ‘human’, ‘natural_object’⁷², etc. A cada nombre de WordNet se le asignaron los tipos básicos mínimos necesarios. Por ejemplo, la palabra *book* tiene en WordNet siete sentidos; Buitelaar los reduce a dos: ‘art’ (‘artifact’) y ‘com’ (‘communication’).
- (b) Agrupa las unidades léxicas que comparten la misma distribución de tipos básicos en “clases polisémicas sistemáticas”. Por ejemplo, la clase ‘art atr sub’ incluye nombres que son a la vez ‘artifact’, ‘attribute’ y ‘substance’, como *chalk*, *charcoal*, *daub*, *fiber*, *fibre*, *tincture*. Las clases que tienen un solo tipo básico o una combinación única de tipos se eliminan.

⁷² En general, intentaremos replicar el estilo gráfico que usa cada autor para representar conceptos o sentidos. De no ser esto posible, usaremos las minúsculas con comillas simples, como hemos hecho en el resto de la tesis, por ejemplo: ‘humano’ o ‘animal’.

- (c) Representa las clases polisémicas sistemáticas como “tipos semánticos subespecificados”. Por ejemplo, el tipo ‘anp’ describe nombres que corresponden tanto a la clase ‘anm’ (‘animal’) como a ‘psy’ (‘psychology’), esto es, nombres que describen animales de una naturaleza psicológica o conceptual (por ejemplo, las criaturas mitológicas).

Los 39,937 nombres procesados por Buitelaar (1998a) produjeron un total de 529 clases polisémicas, que luego fueron agrupadas en 126 tipos semánticos subespecificados. Estos tipos subespecificados son naturalmente de granularidad gruesa, lo que Buitelaar considera una ventaja sobre WordNet, que tiene 60,000 etiquetas solo para los nombres. El autor indica que muchas tareas del NLP pueden ser atendidas con componentes subespecificados, dado que no requieren una especificación definitiva del significado. Otra ventaja de CoreLex sobre WordNet, señala el autor, es que los sentidos se asignan de una manera consistente, por ejemplo: los sentidos de palabras como *door*, *gate*, *window* tienen paralelismos entre sí en CoreLex, mientras que no tiene relación explícita en WordNet.

Las aplicaciones de CoreLex están principalmente en el análisis del discurso y el etiquetado semántico. Sin embargo, para el autor, reconocer la naturaleza sistemática de la polisemia y su relación a las representaciones subespecificadas es esencial para diseñar ontologías que sirvan al PLN de manera más eficiente y con capacidad de generar más interpretaciones apropiadas en contexto.

Peters y Wilks (2001) y Peters (2006) señalan que una de las debilidades de CoreLex es que los conceptos de WordNet que usa pertenecen a un nivel muy alto de la jerarquía taxonómica, lo que puede resultar en información tan subespecificada que no tenga mucha utilidad. Estos autores indican que el usar niveles altos de la jerarquía tiene menos probabilidades de producir patrones de relaciones semánticas significativas entre los sentidos subsumidos, y, además, permite que se pierdan patrones significativos más específicos en la masa de palabras participantes. Barque y Chaumartin (2006), por su parte, señalan que los resultados de Buitelaar (1998a) no han sido evaluados, así que no hay certeza de si los casi 40,000 nombres polisémicos son realmente manifestaciones de las clases polisémicas descritas por los 126 tipos semánticos. Otra crítica es que los tipos semánticos son muy generales y no muy intuitivos, por lo que ofrecen muy poca información sobre el tipo de relación polisémica. Otro señalamiento a esta propuesta, desde otra perspectiva, es las clases sistemáticas dependen del inventario de tipos semánticos, que, en este caso, ha sido elaborado

de forma arbitraria (Mairal, 2011). Sería preferible contar una ontología de conceptos que sea parte de una base de conocimientos, construida según sus propios criterios.

5.2 Tomuro (1998)

Siguiendo la línea de Buitelaar (1998a), Tomuro propone un método semiautomático para inducir “clases semánticas subespecificadas” para los verbos y nombres en WordNet 1.6. Para él, la clase semántica subespecificada es un tipo abstracto que generaliza el vínculo de un grupo de sentidos relacionados. Su metodología sigue los siguientes pasos:

- (a) Selecciona manualmente una serie de sentidos abstractos (de granularidad gruesa) de las taxonomías de WordNet como “tipos semánticos básicos”, tanto para los nombres como para los verbos. Establece 31 tipos básicos de las categorías más altas de WordNet para los nombres, por ejemplo: `entity(ENT)`, `life_form(LIF)`, `causal_agent(AGT)`, `human(HUM)`. Para los verbos elige 18 sentidos, entre ellos: `change(CHA)`, `communication(COMM)`, `cognition(COG)`, `competition(COMP)`.
- (b) Crea una “gráfica de dependencias de tipos” para las palabras polisémicas de WordNet. Este paso conlleva dos fases: un análisis automático seguido de un filtro manual. Primero se aplican técnicas estadísticas a las palabras polisémicas, para crear automáticamente la gráfica de dependencias de tipos. En la gráfica quedan plasmados los tipos semánticos, que se encierran en un círculo, y las relaciones sistemáticas que existen entre ellos, que se representan como flechas.

Para construir la gráfica, se calcula el grado de asociación entre cada par de tipos semánticos, para llegar al “índice de información mutua”. La asociación se considera significativa cuando el valor de la información mutua es mayor que un umbral preestablecido. El filtro manual requiere verificar estas asociaciones, para eliminar relaciones homónimas y recuperar asociaciones que no emergieron de la concurrencia estadística. Las flechas que representan la dirección de las relaciones se trazan manualmente.

- (c) Asigna clases semánticas subespecificadas según la distribución de tipos básicos de cada palabra, es decir, según su distribución de sentidos. Este paso es completamente automático. De ahí resultaron 136 clases semánticas subespecificadas de verbos y 325 clases de nombres.

Para ilustrar sus resultados, veamos algunas de las clases de verbos que incluyen el tipo `contact(CONT)`: `CONT-MOT` representa un contacto físico que resulta de un

movimiento e incluye verbos como *beat, chop, fumble* y CONT-POSS representa una transferencia de posesión que conlleva contacto físico e incluye verbos como *pluck, release, seize*. Una clase más polisémica es CONT-MOT-POSS, que tiene miembros como *carry, cover, fling, toss*, que cubren los tres tipos.

Para confirmar la utilidad de las clases semánticas en las interpretaciones semánticas y las inferencias contextuales de textos reales, Tomuro las aplicó a las estructuras predicado-argumento del corpus Brown⁷³. Concluyó que la inferencia facilitada por las clases subespecificadas es más significativa cuando tanto el predicado como el argumento son polisémicos.

Este método de inducción puede considerarse un intento inicial de adquirir automáticamente la polisemia sistemática de un recurso léxico de cobertura amplia. Sin embargo, como reconoce el propio autor, el trabajo es preliminar y requiere más estudio.

5.3 Peters y Peters (2000)

Estos autores también recurren a WordNet 1.5 para extraer patrones de polisemia regular, pero se enfocan en explotar la estructura jerárquica en sí misma. En la primera fase del estudio, siguieron estos criterios:

- (a) al menos dos palabras deben compartir la misma combinación de sentidos que representa el patrón de polisemia regular;
- (b) las palabras pueden o no pertenecer al mismo *synset*; y
- (c) los *synsets* involucrados deben pertenecer a la misma clase sintáctica.

Luego seleccionaron nodos del nivel superior de la taxonomía (“unique beginners”) que, en combinación, pudieran compartir una misma palabra como hipónimo en cualquier nivel de la jerarquía. Un ejemplo de una combinación de “unique beginners” es:

artifact-1: un objeto hecho por el hombre
group-1: cualquier número de entidades (miembros) considerados como una unidad

Se encontraron múltiples pares de palabras que lexicalizaban la pareja de “unique beginners”, o sea, que manifestaban la misma alternancia de sentidos, por ejemplo:

⁷³ El Brown University Standard Corpus of Present-Day American English (o “corpus Brown”) contiene 500 muestras de textos en inglés publicados en los Estados Unidos en 1961, que suman un total de un millón de palabras.

| | |
|----------------|---|
| institution-2: | un edificio o complejo de edificios donde se sitúa una organización para la promoción de alguna causa |
| institution-1: | una organización fundada para un propósito específico |
| guard-3: | un dispositivo diseñado para prevenir las heridas o lesiones |
| guard-5: | un grupo de hombres que escolta y protege a una persona importante |

En general los resultados indicaron que no es provechoso empezar con conceptos de muy alto nivel. Las combinaciones de “unique beginners” pueden generar instancias inapropiadas de un patrón o grupos que contienen palabras que no son similares semánticamente y no presentan una relación metonímica. A la vez, usar conceptos de alto nivel para la caracterización de la polisemia sistemática puede bloquear la identificación de subgrupos semánticamente más coherentes.

En la segunda fase del estudio, los autores identificaron combinaciones de hiperónimos en un nivel más bajo y específico que los “unique beginners”, pero que fueran lo suficientemente generales para abarcar varias palabras y constituir grupos semánticamente homogéneos. Establecieron los siguientes criterios:

(a) los pares de nodos deben ser hiperónimos de al menos tres palabras que tengan al menos un sentido en cada rama;

(b) la distancia entre los sentidos y el hiperónimo en la taxonomía de WordNet nunca debe exceder cuatro (4) niveles. Esta distancia limita el tiempo de procesamiento y el volumen de datos resultantes.

Partiendo de los “unique beginners” ‘artifact’ y ‘action’, y usando los nodos más específicos ‘music’ y ‘dance’, se obtuvieron tres palabras: *rumba*, *waltz* y *bolero*. El patrón polisémico sistemático es ‘music’~‘dance’. En el caso de los nodos específicos ‘passage’ y ‘structure’, obtuvieron palabras como *arcade*, *arch* y *gallery*. Hicieron lo mismo con varios otros pares de nodos.

Los patrones polisémicos obtenidos corresponden a varios de los casos conocidos en la bibliografía, especialmente los que tienen un número grande de miembros, como ‘music’~‘dance’ y ‘container’~‘quantity’. Solo un 10% de los miembros extraídos fueron considerados inválidos. Sin embargo, la metodología presentó varios inconvenientes. Primero, solo permite el manejo de parejas de sentidos, y no de grupos de tres o más. Segundo, es difícil extraer automáticamente parejas conceptuales adecuadas según los criterios establecidos; mientras más específica es la pareja conceptual, más definida es la relación semántica que existe entre los conceptos del par, pero más pequeña es la serie. La

tercera desventaja es el límite de cuatro niveles entre los conceptos hiponímicos y los pares conceptuales; esto presupone que la jerarquía de WordNet es equilibrada, cuando en realidad no es así.

5.4 Peters y Wilks (2001)

En este trabajo, los autores intentan hacer explícitas las relaciones metonímicas presentes en la estructura jerárquica de WordNet 1.6. Los patrones resultantes son evaluados contra las colocaciones de sentidos de Semcor⁷⁴.

Los pares de nodos hiperonímicos cumplen con un criterio ya ensayado en Peters y Peters (2000): deben funcionar como hiperónimo de al menos tres palabras que tengan al menos un sentido bajo cada nodo. Esto resulta en grupos más coherentes semánticamente, como:

| | |
|---|-----------------------------------|
| ‘publication’/‘publisher’: | <i>paper, newspaper, magazine</i> |
| ‘musical composition’/‘group of singers’: | <i>trio, quartet, suite</i> |
| ‘building-institution’/‘association’: | <i>school, chamber, court</i> |
| ‘package-container’/‘collection’: | <i>parcel, bundle, pack</i> |
| ‘music-arrangement’/‘formation’: | <i>line, arrangement, chorus</i> |
| ‘construction’/‘body of people’: | <i>house, body, camp</i> |

Para reducir el volumen de datos, se usó una técnica que los proyectó a las colocaciones de sentido de Semcor. Hubo 184 palabras cuyos sentidos coocurrieron dos veces o más en los documentos de Semcor; la mayoría de estas palabras está asociada con más de un par hiperonímico y puede constituir un caso válido de polisemia regular.

Los autores escogieron el subgrupo de 23 palabras con sentidos que coocurrían más de siete veces en los documentos de Semcor para evaluar este método. Postularon que estas combinaciones de alta frecuencia podían maximizar las posibilidades válidas. Solo cuatro se consideraron casos válidos de polisemia regular:

| | |
|-------------------------------|-----------------|
| ‘group action’/‘organization’ | <i>business</i> |
| ‘structure’/‘way’ | <i>door</i> |
| ‘measure’/‘clock time’ | <i>hour</i> |
| ‘message’/‘condition’ | <i>problem</i> |

De las demás palabras, seis fueron casos de metonimia, que, junto a las cuatro, rinden

⁷⁴ SemCor es una parte del corpus Brown que incluye 127 documentos y que ha sido etiquetado con sentidos de WordNet 1.6.

un 43.5% de palabras relacionadas metonímicamente. Ocho sentidos se consideraron como casos de especificación/generalización donde un sentido es una instancia más específica del otro, o ambos muestran un gran nivel de similitud. Ninguno de los sentidos estuvo relacionado metafóricamente.

Según Peters y Wilks (2001), el resultado de esta evaluación a pequeña escala apoya el que la coocurrencia de sentidos múltiples dentro de un discurso puede constituir evidencia de relaciones léxico-semánticas. Las colocaciones dentro de los documentos de Sencor capturan un número sustancial de patrones metonímicos. Los autores indican que estos patrones de polisemia regular extraídos pueden servir para extender al marco ontológico de WordNet con nuevas relaciones. Por ejemplo, según descritos en Peters y Peters (2000), el caso de ‘music’/‘dance’ que cubre *bolero*, *waltz* y *rumba* puede ser caracterizado semánticamente como ‘accompanies’, y el de ‘passage’/‘structure’ que cubre *door* y puede identificarse con una relación de ‘leads_through’ o ‘functions_as’.

5.5 Wing (2002)

Wing (2002) realiza un interesante estudio en el que busca nombres en tres idiomas diferentes que compartan dos hiperónimos en EuroWordNet. Concluye que se trata de un buen método para localizar casos de polisemia regular y que hay coincidencias entre las tres lenguas. Los pasos que sigue son:

(a) Busca todos los casos en que dos palabras (solo nombres) tengan dos sentidos diferentes cuyos hiperónimos sean los mismos. Agrupa los pares de hiperónimos. Por ejemplo, bajo los sentidos ‘fabric’ (algo hecho mediante tejer fibras naturales o sintéticas) y ‘covering’ (un objeto natural que cubre o arropa), encontró palabras como: *fleece*, *hair*, *tapa*, *wool*. De este paso salieron 8,062 nombres en inglés.

(b) Compara tres lenguas: inglés, holandés y español. Los escoge porque quería representar diferentes familias lingüísticas, pero necesitaba idiomas con redes de palabras relativamente completas en EuroWordNet.

(c) Busca palabras en inglés con dos sentidos en diferentes *synsets* donde los *synsets* correspondientes tanto en holandés como en español tengan una palabra en común. Por ejemplo, *church*, *iglesia* y *kerk* se refieren tanto al edificio como a la institución religiosa. De aquí salieron 920 nombres en inglés.

(d) Intersecta las palabras resultantes con las palabras de la sección previa. Resultaron

404 nombres en inglés (5% de las 8,062 palabras iniciales). Como este número era muy alto, descartó los grupos de exactamente dos palabras, e hizo la intersección otra vez. El resultado fue 394 nombres en inglés, una reducción mínima.

(e) Escoge 177 palabras aleatoriamente para evaluación manual. Verificó que cada grupo representado fuera válido. Examinó los dos hiperónimos que definen el grupo y aseguró (i) que fueran razonablemente específicos y (ii) que tuvieran “homogeneidad semántica”, es decir, que de verdad hubiera una relación semántica, por ejemplo, ‘*es la ubicación de*’ en el caso de ‘edificio’ e ‘institución religiosa’. El resultado fue que 109 palabras (62%) demostraron patrones polisémicos válidos, mientras que 68 (38%) no.

Algunos ejemplos de los pares son:

Tabla 26. Par hiperonímico ‘Person’/ ‘Quality’ en EuroWordNet (Wing, 2002)

| | |
|--------------------------------------|--|
| Par hiperonímico: | ‘Person’ (ser humano) y ‘Quality’ (un atributo de algo o alguien esencial y caracterizador). |
| Inglés (11 en total): | <i>attraction, authority, beauty...</i> ⁷⁵ |
| Holandés (1 en total): | <i>schoonheid</i> |
| Español (4 en total): | <i>belleza, atracción, autoridad, imagen</i> |
| Intersección entre las tres lenguas: | 9% de la serie derivada de WordNet |

Tabla 27. Par hiperonímico ‘Control’/ ‘Trait’ en EuroWordNet (Wing, 2002)

| | |
|--------------------------------------|--|
| Par hiperonímico: | ‘Control’ (la actividad de manejar o ejercer control sobre algo) y ‘Trait’ (un rasgo distinguidor de la naturaleza personal de uno). |
| Inglés (7 en total): | <i>abstinence, sobriety, inhibition, restraint, self-control, self-denial, self-discipline</i> |
| Holandés (2 en total): | <i>zelfcontrole, onthouding</i> |
| Español (3 en total): | <i>abstinencia, abnegación, inhibición</i> |
| Intersección entre las tres lenguas: | 36% de la serie derivada de WordNet |

El autor concluye que (a) este mismo método puede ser aplicado más allá de EuroWordNet, a cualquier recurso multilingüe con relaciones y correspondencias hiperonímicas entre lenguas; (b) algunos patrones de polisemia regular son válidos a través de las tres lenguas y parecen tener cierta universalidad; y (c) hay potencial para mejorar (semi-) automáticamente la compatibilidad y consistencia semántica de las redes de palabras a través de extensiones de sentido basadas en información de polisemia regular (patrones) derivables de otras redes de palabras.

⁷⁵ El artículo consultado no presentan la lista completa.

5.6 Peters (2004)

En esta tesis, el autor explota la estructura jerárquica y las glosas de WordNet con la meta hacer explícitas las relaciones sistemáticas de metonimia que existen entre los sentidos, específicamente, las de polisemia regular. Peters selecciona palabras con sentidos sistemáticamente relacionados, y luego analiza las glosas asociadas a los sentidos bajo consideración, para intentar capturar la relación semántica entre los sentidos. Los patrones de polisemia regular son comparados con Sencor para establecer cuáles patrones realmente ocurren en el texto. Finalmente, las relaciones que han sido validadas son integradas a “fragmentos de conocimiento” que forman una extensión de las estructuras de conocimiento que están disponibles explícitamente en WordNet. El resultado es una base de conocimiento extendida con una cantidad mayor de conocimiento de trasfondo, lo cual ayuda a agilizar la adquisición del conocimiento y puede ser explotado para el proceso de comprensión.

Peters (2004) presenta un inventario de 138 pares de hiperónimos que intervienen en patrones de polisemia sistemática, como los dos ejemplos siguientes. De las 138 relaciones, la evaluación reveló una precisión de 50%.

Tabla 28. Par hiperonímico ‘profession’/‘discipline’ en WordNet

| | |
|----------------------------|--|
| sentido 1: ‘profession’ | an occupation requiring special education (especially in the liberal arts or sciences) |
| sentido 2: ‘discipline’ | a branch of knowledge; "in what discipline is his doctorate?"; "teachers should be well trained in their subject"; "anthropology is the study of human beings" |
| palabras | <i>architecture, law, literature, politics</i> |

Tabla 29. Par hiperonímico ‘game’/‘equipment’ en WordNet

| | |
|---------------------------|--|
| sentido 1: ‘game’ | a contest with rules to determine a winner; "you need four people to play this game" |
| sentido 2: ‘equipment’ | an artifact needed for an undertaking or to perform a service |
| palabras | <i>baseball, basketball, football, handball</i> |

Esta tesis concluye que es viable extraer información implícita en WordNet y vincular esta información de vuelta a la base de conocimiento en una forma explícita. El método puede ser aplicable a cualquier recurso que contenga información taxonómica asociada con sentidos.

Mairal (2011) señala que una debilidad del inventario resultante de los 138 patrones de polisemia regular es que es “producto de un proceso de selección guiado por el juicio subjetivo y la intuición, ya que es el resultado de consultar la literatura pertinente a este tema y los patrones obtenidos por su propia evaluación manual de WordNet”.

5.7 Peters (2006)

Peters continúa el trabajo de Peters (2004) de extracción de conocimiento implícito contenido en la estructura jerárquica de WordNet y EuroWordNet y en las glosas asociadas con cada *synset*. En este caso, intenta extraer automáticamente la formulación explícita de la relación que media entre los sentidos polisémicos regulares. Su metodología siguió tres pasos:

(1) Un proceso automático identifica los candidatos de polisemia regular en WordNet según las distribuciones sistemáticas de los sentidos de los nombres, o sea, se identifican pares de hiperónimos que subsumen las combinaciones de sentidos de las palabras involucradas. Por ejemplo, en dos de sus sentidos, *law* cae bajo ‘profesión’ y ‘disciplina’. Este patrón también lo tienen cuatro otras palabras en WordNet: *architecture*, *literature*, *politics* y *theology*.

(2) Se extraen automáticamente las relaciones entre los sentidos que participan en los patrones. Esta información adicional se obtiene analizando las glosas asociadas con los *synsets* y sus hiperónimos. Todos los miembros de los *synsets* se agrupan en dos grupos de palabras, que se proyectan en las glosas asociadas. Si un verbo ocurre entre pares de palabras de cada grupo, se toma como la relación semántica que se sostiene entre los sentidos. Por ejemplo, el sentido #6 de *law* tiene la glosa ‘*the learned profession that is mastered by graduate study in a law school and that is responsible for the judicial system*’: *he studied law at Yale*. El *synset1* contiene *profession*, el *synset2* contiene *study*. Entre medio está el verbo *is mastered by*, que provee la relación de este patrón de polisemia regular. Al añadir roles temáticos a los conceptos involucrados, se puede aseverar que ‘disciplina’ es el sujeto o instrumento asociado con *is mastered by* y ‘profesión’ es el objeto. En total, se extrajeron 5,000 candidatos de patrones de polisemia regular.

(3) En la tercera fase, se construyen estructuras cada vez más grandes a base de los pares de sentido involucrados en los patrones. La estructura de marco es definida como la

relación que existe entre los elementos de un marco o entre el marco como totalidad y sus elementos.

Barque y Chaumartin (2006) señalan que, si bien el autor da ejemplos de las relaciones, no informa cuántas son ni provee una evaluación de los datos que han sido extraídos.

5.8 Barque y Chaumartin (2006)

Barque y Chaumartin (2006) se acercan al estudio de Peters (2006), pues introducen un método para extraer patrones de polisemia regular de las glosas de WordNet. Su método tiene cuatro pasos:

(1) Extrae *synsets* “auto-referentes”, estos es, *synsets* cuya definición incluye una palabra que comparte la forma con una de las unidades léxicas definidas. Por ejemplo:

- {cerise#1, cherry#4} = the red color of cherries
- {driver#3} = a golfer who hits the golf ball with a driver
- {falsify#4} = falsify knowingly

De este paso, obtuvieron 1,984 *synsets* que probablemente eran ocurrencias de relaciones de polisemia regular.

(2) Describe manualmente los patrones de polisemia a partir de la observación de los 1,984 *synsets* extraídos. Este método depende de dos criterios aplicados a definiciones de WordNet: la posición de inclusión de la unidad léxica#1 (L1) en la definición de la unidad léxica#2 (L2) y los elementos de la definición pertinentes a la relación de polisemia. A partir de esto, se puede atribuir una de tres categorías de polisemia regular (especialización, metáfora o metonimia) a una ocurrencia dada.

(3) Se desambigua, con la ayuda de patrones, el significado de la palabra polisémica de la definición. De los 1,984 candidatos, se sacaron 1,427 ocurrencias de relaciones polisémicas (los otros casos solo ocurren una sola vez). En el patrón ‘ColorOf’, hubo que desambiguar a L2, como se muestra en las definiciones a continuación, entre corchetes:

- {emerald#3} = the green color of an emerald#1[gem]
- {tan#2, topaz#3} = a light brown, the color of topaz#2[gem]
- {copper#4} = a reddish-brown color resembling the color of polished copper#1[metal]

(4) Se generaliza el método a los *synsets* que no son auto-referentes. Encontraron 367 nuevas instancias. La palabra *gold*, por ejemplo, tiene cinco significados en WordNet: ‘monedas’, ‘color’, ‘metal’, ‘gran riqueza’ y ‘algo preciado’. Dos significados tienen el tipo color y metal y pueden, por tanto, estar asociados con la regla ColorOf, incluso si la definición de *gold#2* no incluye la palabra *gold*.

- {amber#1, gold#2} = a deep yellow color (implicit link to gold#3[metal])
- {coral#1} = a variable color averaging a deep pink (implicit link to coral#2 [gem])

Este paso necesita más refinamiento, según los autores.

Este estudio tiene dos tipos de resultados: uno descriptivo, con la clasificación de las relaciones de polisemia regular, y uno metodológico, con la detección automática de las ocurrencias de polisemia regular con precisión aceptable. A continuación se presentan algunos ejemplos de relaciones regulares de metonimia y metáfora (entre paréntesis, se presenta el número de instancias reales comparadas con el número de candidatos detectados):

Relaciones de metonimia regular:

L2 representa a L1

→ playing card represents person or entity (5/6; *queen, king ; ten, nine*)

L2 es producido por L1

→ sound produced by instrument or movement or device (15/15; *drum, whistle; snap; bell*)

Relaciones de metáfora regular:

L2 es similar a L1

→ human communication similar to animal communication (3/4; *to bark, to cackle*)

→ animal part of the body corresponds to human part of the body (3/3; *leg, throat*)

Las 2,351 instancias de relaciones de polisemia regular propuestas por el sistema se evaluaron manualmente; el resultado tuvo una precisión de 91.03%. El “recall” o memoria automática se evaluó manualmente para dos patrones de polisemia regular: la relación metafórica ‘person resembles animal’ dio 25.3%, y la relación metonímica ‘wood derived from tree’ 88.6%. Los autores concluyen que el “recall” también depende de la naturaleza de la relación.

5.9 Barque (2007)

Barque (2007) usa las definiciones lexicográficas de granularidad fina del BDéf⁷⁶, para describir los vínculos entre sentidos polisémicos regulares en palabras del francés pertenecientes al campo semántico de las emociones. Estas definiciones se estructuran explícitamente, lo que permite una descripción más rica de los vínculos de la polisemia y más sutil de las regularidades.

La autora observa que las unidades léxicas que denotan un cierto tipo de emoción están regularmente vinculadas a las unidades léxicas que denotan el objeto de una emoción, por ejemplo:

(1) AMOUR#1 (amor) [Sa naïveté l'empêche de comprendre qu'elle ressent de l'amour pour son cousin. / Ella es demasiado inocente para entender que está enamorada de su primo.] ~ AMOUR#2 [Mon amour a les cheveux noirs. / Mi amor tiene el cabello negro.]

(2) ESPOIR#1 (esperanza) [Internet suscite un immense espoir chez les jeunes scolarisés. / El Internet les da una inmensa esperanza a los jóvenes estudiantes.] ~ ESPOIR#2 [Vous êtes mon dernier espoir. / Tú eres mi última esperanza.]

Para Barque, el vínculo de la polisemia está orientado semánticamente; existe pues una unidad léxica fuente y una unidad léxica semánticamente derivada. En los ejemplos de arriba, AMOUR#1 y ESPOIR#1 son las unidades fuente y AMOUR#2 y ESPOIR#2 son las unidades derivadas semánticamente.

Un “patrón de polisemia” es una herramienta descriptiva que se usa para modelar los vínculos de polisemia en el BDéf. Se compone de dos tipos de información:

(1) *La subespecificación del par de unidades léxicas a las que les aplica.* La caracterización semántica de dos unidades léxicas puede estar más o menos subespecificada. En este caso, la primera especificación obvia es decir que la unidad léxica que denota la emoción debe tener un segundo actante que represente el objeto de la emoción.

(2) *La explicación del vínculo semántico entre la unidad léxica fuente y la derivada.* Esta explicación, que justifica que las dos unidades no sean homónimas, debe estar apoyada por la subespecificación semántica. Si la unidad léxica L1 está metafóricamente vinculada a otra unidad L2, hay que demostrar la analogía entre la subespecificación de L1 y L2. Si la unidad léxica L1 está vinculada a otra unidad L2 metonímicamente, habrá que demostrar la contigüidad entre la subespecificación de L1 y L2. El vínculo ‘emoción’~‘objeto de la

⁷⁶ La *Base de Définitions* (BDéf) es una base de datos formal derivada del *Explanatory Combinatorial Dictionary of Contemporary French* (Altman y Polguere, 2003).

emoción' es un vínculo metonímico dado que la unidad léxica derivada corresponde al segundo actante de la unidad fuente.

La autora distingue entre la metonimia fuerte y la débil. La fuerte se caracteriza por que la unidad léxica derivada incluye el significado de la unidad léxica fuente (Barque y Polguère, 2005); por ejemplo, el vínculo de metonimia entre HOPE#1 (emoción) y HOPE#2 (objeto de la emoción) es fuerte porque la definición de HOPE#2 incluye a HOPE#1. Por el contrario, en la metonimia débil la unidad léxica derivada no incluye el significado de la unidad léxica fuente, como es el caso de ROUGIR#1 (reacción física) y ROUGIR#2 (mal sentimiento). Barque propone además la noción de “dominio de polisemia”, que se puede considerar una expansión de la noción de campo semántico.

Entre las relaciones de metonimia que identifica figuran:

‘emoción’~‘objeto de la emoción’: aplica a FIERTÉ (‘orgullo’), HONTE (‘vergüenza’ ~ ‘humillación’), AMOUR (‘amor’), JOIE (‘gozo’), ESPOIR (‘esperanza’)

‘sentimiento’~‘comunicación lingüística’: aplica a REGRETTER (‘arrepentirse’ ~ ‘deplorar’), SOUHAITER (‘tener esperanza en’ ~ ‘anhelar’),

Entre las relaciones de metáfora figuran:

‘mala sensación física’~‘mal sentimiento’: Aplica DOULEUR (‘dolor’ ~ ‘pena’), IRRITATION (‘irritación’), MALAISE (‘desvanecimiento’ ~ ‘inquietud’), ULCÈRE (‘ulcerado’ ~ ‘horrorizado’)

‘sabor’~‘sentimiento’ : Aplica a SUCRÉ (‘dulce’ ~ ‘azucarado’), ACIDE (‘ácido’), BITTER (‘amargo’), AIGRE (‘agrio’ ~ ‘cortante’), ÂCRE (‘ácido’)

La autora señala que sus patrones de polisemia todavía están en construcción. Desde un punto de vista práctico, podría ser útiles para el desarrollo de una base de datos porque ayudan a sistematizar la descripción proveyendo, de ser necesario, entradas subespecificadas. Además, la organización jerárquica de estos patrones (según los diferentes niveles de subespecificación) podría ser interesante para un estudio interlingüístico.

Conclusión

El potencial para el PLN de identificar patrones regulares de alternancia de sentidos para clases de palabras incluye: (a) simplificar la tarea de asignación de sentidos en la WDS; (b) poder predecir el uso de sentidos noveles a partir de los sentidos establecidos en patrones identificados; y (c) obviar la desambiguación si los sentidos pueden ocurrir simultáneamente, limitándose a identificar clases semánticas subespecificadas.

La mayoría de los acercamientos a la polisemia regular se han aplicado a sustantivos. Algunos han enfrentado dificultades dada la granularidad de los lexicones usados, lo que los ha llevado a seleccionar conceptos de alto nivel que amplían demasiado los posibles conceptos subordinados e hipónimos correspondientes.

Nuestro trabajo mirará la polisemia regular en verbos a la luz de la ontología conceptual de FunGramKB, en un acercamiento semántico y conceptual.

CAPÍTULO VI. METODOLOGÍA

Introducción

En este capítulo caracterizamos el acercamiento metodológico seguido en esta tesis. Aunque exponemos una serie de pasos enumerados, el método no es enteramente lineal y conlleva procesos simultáneos y múltiples vueltas atrás.

Antes de discutir la metodología, repasemos someramente lo discutido hasta ahora sobre la naturaleza de las categorías de REDES y las categorías de FunGramKB. Las restricciones de selección de REDES son primeramente léxicas. Un predicado verbal selecciona sistemáticamente ciertos grupos de argumentos nominales y, por eso, las restricciones de selección se agrupan en *clases léxicas*. Esta información léxica se organiza inductivamente según nociones semánticas; por esa razón, las clases son también *clases semánticas*. Finalmente, la mayoría de las clases se asocia con un concepto del índice conceptual de REDES; de ese modo, las clases cuentan también con una etiqueta *conceptual*. Veamos como ejemplo, en la tabla siguiente, los tres tipos de información –léxica, semántica y conceptual– que ofrece REDES para una clase léxica que se combina con el verbo *cobrar* (en el sentido de ‘adquirir’):

Tabla 30. Descripción de clase léxica en REDES que se combina con *cobrar* ‘adquirir’

| Naturaleza de la información | Componente de REDES | Información específica |
|------------------------------|---------------------|---|
| léxica | argumentos | <i>fuerza, brío, vigor, ánimo, impulso, ímpetu, intensidad, velocidad</i> |
| semántica | descriptor | ‘fuerza, energía o empuje, físicos o anímicos’ |
| conceptual | índice conceptual | FUERZA |

Los conceptos de FunGramKB, por su parte, son nociones cognitivas, que se organizan en una taxonomía de relaciones de subsunción IS-A. Los conceptos tienen una sola lectura recta, que se define conceptualmente a través de los esquemas de MT y PS. Intentamos trasladar las restricciones de selección léxica de REDES de verbo a nombre (o a clase léxica nominal) a preferencias conceptuales del evento conceptual en FunGramKB. Para ello, vinculamos los predicados verbales a conceptos en la subontología de eventos y las clases léxicas nominales a conceptos en la subontología de entidades. Luego, examinamos hasta qué punto la entidad puede ser una preferencia de selección del evento y, por lo tanto,

pasar a ser parte de su definición. Recordemos que las preferencias de selección del concepto son especificaciones semánticas de los participantes en dichos esquemas conceptuales⁷⁷.

Las discrepancias que encontramos entre los contenidos de REDES y FunGramKB son naturales y esperables, pues cada uno está diseñado para propósitos diferentes. Sin embargo, el trabajo de esta tesis se fundamenta en la idea de que hay suficiente afinidad entre las informaciones semántico conceptuales de REDES y FunGramKB para vincular los datos de manera productiva.

En la primera parte del capítulo, presentamos los pasos seguidos para representar conceptualmente en FunGramKB la relación entre el significado y la combinatoria léxica de 10 verbos: *dictar*, *derrochar*, *desaprovechar*, *desperdiciar*, *despilfarrar*, *dilapidar*, *disipar*, *malgastar*, *malograr* y *perder*. En la sección 1.1, establecemos cómo se eligieron los verbos participantes. En la sección 1.2, explicamos cómo vinculamos individualmente cada clase léxica a una entidad ontológica. Repasamos la naturaleza de las clases léxicas de REDES y algunos de los retos que supuso el vincularlas a la ontología conceptual de FunGramKB. En la sección 1.3, establecemos una tipología para caracterizar las clases léxicas con respecto a FunGramKB. En la sección 1.4, analizamos las clases léxicas colectivamente, con el fin de reducir el inventario de clases. En la sección 1.5 discutimos cómo se vincularon los predicados a eventos ontológicos, incluidas sus preferencias de selección. En la sección 1.6, discutimos el manejo de las colocaciones en el módulo léxico de FunGramKB.

En la segunda parte del capítulo, presentamos los pasos seguidos para representar conceptualmente la polisemia regular y la combinatoria léxica de cinco verbos (*tejer*, *tramar*, *urdir*, *cocinarse* y *maquinar*) en FunGramKB. Replicamos la mayoría de los pasos discutidos en la primera parte, pero los entallamos al caso de la polisemia regular.

1. Significado y combinatoria

En esta sección detallamos los pasos seguidos para vincular con la ontología de FunGramKB el significado y combinatoria léxica del verbo *dictar* y de nueve verbos – monosémicos y polisémicos– relacionados semánticamente (*derrochar*, *desaprovechar*, *desperdiciar*, *despilfarrar*, *dilapidar*, *disipar*, *malgastar*, *malograr* y *perder*).

⁷⁷ Todo esto se detallará a lo largo de este capítulo.

1.1 Elegir predicados de REDES relacionados semánticamente

Elegimos un grupo de predicados de REDES y PRÁCTICO relacionados semánticamente que se pudieran vincular –al menos preliminarmente y al menos en uno de sus sentidos– a un mismo concepto en la ontología de FunGramKB. El contar con predicados relacionados semánticamente nos permite trabajar, por un lado, con la relación entre significado y combinatoria léxica de cada predicado (monosémico o polisémico), y, por otro, con los rasgos compartidos por el grupo de predicados, tanto en cuanto a sus significados, como en cuanto a sus patrones de selección léxica.

Para identificar el grupo de predicados, partimos de la noción de ‘malgastar’ y de la palabra *malgastar*. Utilizamos las referencias de “Véase también” de REDES, que remiten a entradas de palabras “relativamente próximas” (Bosque, 2004b: L). También consultamos diccionarios para identificar sinónimos y palabras relacionadas a *malgastar*. Se formuló una larga lista de verbos relacionados entre sí. Sin embargo, resultó que no todos tenían entrada en REDES o PRÁCTICO. La lista final incluyó todos los verbos identificados que tuvieran una entrada analítica o abreviada de selección de argumentos en REDES o PRÁCTICO. Los verbos fueron: *derrochar*, *desaprovechar*, *desperdiciar*, *despilfarrar*, *dilapidar*, *disipar*, *malgastar*, *malograr* y *perder*. Estos nueve predicados se pueden vincular en por lo menos uno de sus sentidos a la noción de ‘malgastar’ y al concepto +WASTE_00 en FunGramKB.

El verbo *dictar*, que se estudió primero y por separado, se eligió por ser claramente polisémico y tener una rica combinatoria léxica en REDES.

1.2 Vincular cada clase léxica con un concepto de la ontología

Las clases léxicas están compuestas de argumentos nominales, por lo que los conceptos correspondientes en la ontología de FunGramKB pertenecen a la subontología de entidades. Para establecer el vínculo entre clase léxica y entidad conceptual, tomamos en cuenta:

- (a) la información léxica, semántica y conceptual de cada clase léxica en REDES y
- (b) la descripción del concepto en FunGramKB y sus lexicalizaciones.

Cada predicado se combina con varias clases léxicas: los nueve predicados relacionados semánticamente seleccionan entre todos un total de 91 clases léxicas, incluidas tanto las clases léxicas “abiertas”, que aparecen en la entradilla, como las clases léxicas en propiedad, que se identifican con letra mayúscula. Por su parte, el verbo *dictar* se combina

con ocho clases léxicas. La distribución de clases léxicas entre los nueve verbos es la siguiente:

Tabla 31. Verbos estudiados y cantidad de clases léxicas con que se combinan

| Verbo | Cantidad de clases léxicas |
|----------------------|----------------------------|
| <i>malgastar</i> | 9 |
| <i>desaprovechar</i> | 6 |
| <i>dilapidar</i> | 10 |
| <i>despilfarrar</i> | 6 |
| <i>desperdiciar</i> | 9 |
| <i>derrochar</i> | 17 |
| <i>malograr</i> | 8 |
| <i>disipar</i> | 13 |
| <i>perder</i> | 16 |
| <i>dictar</i> | 8 |

Vincular las clases léxicas con una entidad conceptual es algo más complicado que vincular los verbos con eventos conceptuales. La idiosincrasia de sus miembros, y los rasgos semánticos *ad hoc* que las describen, hacen que cada una sea prácticamente única, lo que dificulta hacer generalizaciones o visualizar patrones en los datos según los presenta REDES. De ahí la importancia de intentar vincular estos datos con una ontología léxico conceptual construida *a priori* de cualquier aplicación, como FunGramKB.

Como vimos antes, REDES ofrece para cada clase léxica información léxica y semántica, y una asociación conceptual. La información léxica es aportada por los argumentos miembros de la clase; por ejemplo, con el predicado *malgastar* se combina una clase léxica compuesta de *fuerza*, *energía*, *esfuerzo* y *salud*. El descriptor de cada clase aporta la información semántica. En este caso, el descriptor de la clase léxica lee: ‘fuerza o energía, en sentido figurado o físico, y sus efectos’. Finalmente, la vinculación de la clase al índice conceptual de REDES aporta información conceptual. Esta clase está vinculada al concepto FUERZA en REDES. Esta información nos dirige a vincular esta clase léxica al concepto +STRENGTH_00 en FunGramKB, cuya descripción lee ‘la propiedad de ser física o mentalmente fuerte’:

Tabla 32. Descripción de +STRENGTH_00 en FunGramKB

| | |
|-----------------|--|
| Hiperónimo: | +PHYSICAL_ATT_00 |
| PS: | +(e1: +BE_00 (x1: +STRENGTH_00)Theme (x2: +PHYSICAL_ATT_00)Referent) *(e2: +BE_01 (x3: +HUMAN_00 ^ +ANIMAL_00)Theme (x1)Attribute) |
| Descripción: | the property of being physically or mentally strong; "fatigue sapped his strength" |
| Lexicalización: | <i>fuerza</i> (español), <i>force, strength</i> (inglés) |

A continuación, consideramos con más detenimiento el rol de cada nivel de información (semántico, léxico, conceptual) de las clases léxicas en este procedimiento.

1.2.1 El descriptor

Los descriptores de las clases léxicas proveen la información semántica sobre la clase léxica. En el proceso de vincular las clases léxicas con entidades conceptuales, comparamos el contenido del descriptor de REDES con los PS de los conceptos en FunGramKB, tomando en cuenta que el primero está codificado en lenguaje natural y el segundo en COREL. Este proceso presenta varios retos que ilustramos a continuación.

En general, la granularidad de las clases léxicas tiende a ser más fina –en el sentido de ser más detallada y descriptiva– que la de las entidades conceptuales de FunGramKB. Por ejemplo, un descriptor puede identificar una noción y luego cualificarla o especificarla, como la clase que se combina con *derrochar* de ‘sentimientos o sensaciones de complacencia, entusiasmo y afecto, a menudo elevados’ (*alegría, optimismo, entusiasmo, ilusión, pasión, euforia, afusión, satisfacción, felicidad, amor, cariño*). En FunGramKB encontramos entidades conceptuales que tienden a ser unívocas, como, en este caso, +FEELING_00 (‘an emotion that you feel; emotional attribute’⁷⁸), que se corresponde bien con el hiperónimo identificado por REDES (‘sentimientos o sensaciones’), pero no puede capturar la especificación (‘a menudo elevados’).

Otro reto está en que los descriptores de REDES pueden aludir a múltiples nociones a la vez. La clase léxica que acabamos de ver identifica tres nociones, cada una de las cuales podría estar vinculada a un concepto diferente en FunGramKB: ‘sentimientos o sensaciones de complacencia, entusiasmo y afecto, a menudo elevados’. A la hora de vincular esta clase

⁷⁸ Los sentidos que se ofrecen entre paréntesis son abreviaciones de acepciones del DRAE 2001, en el caso del español, y las descripciones que ofrece FunGramKB, en el caso del inglés.

léxica con FunGramKB, identificamos las entidades específicas a las que corresponden las nociones identificadas, por ejemplo, +ENTHUSIASM_00 para ‘entusiasmo’, y luego exploramos la entidad superordinada que pueda dar cabida a todas las nociones, que en este caso sería +FEELING_00.

Las clases léxicas que se refieren a nociones muy abstractas también presentan dificultades frente a la ontología. Un ejemplo de ello es la clase que se combina con *cobrar* de ‘algunas propiedades físicas relativas a la integridad, la delimitación o la existencia de las cosas’ (*realidad, forma, cuerpo, consistencia, corporeidad, encarnación*). Además de que nombra tres nociones distintas, todas son notablemente abstractas. Aquí podríamos nuevamente recurrir al hiperónimo que ofrece REDES (‘algunas propiedades físicas’) y elegir un concepto como +PHYSICAL_ATT_00 (‘physical attribute’), que es apropiado, pero extremadamente amplio.

Finalmente, los descriptores también pueden restringir los miembros de la clase léxica a ciertos ámbitos, como en la clase de ‘oportunidad u ocasión, ventaja o circunstancia favorable, frecuentemente deportiva’ (*oportunidad, ocasión, balón, ventaja, punto, opción, privilegio, posibilidad*), que se combina con *malgastar*. Este tipo de restricción situacional o temática no se puede codificar en la ontología de FunGramKB.

Estos ejemplos ofrecen una muestra de la flexibilidad del lenguaje humano, pero también de la dificultad que esta flexibilidad implica para crear un vínculo con conceptos que hacen referencia unívoca a nociones cognitivas, como los conceptos de la ontología de FunGramKB. Cabe señalar que, en esta tesis, hemos decidido no proponer modificaciones a las entidades conceptuales (lexicalizadas por nombres) existentes en la ontología de FunGramKB, sino que trabajamos con ellas tal cual están descritas actualmente. Esto se contrapone a nuestro procedimiento con los eventos conceptuales a los que vinculamos nuestros verbos, sobre los que sí proponemos cambios, que resultan muchas veces en la creación de nuevos conceptos terminales o subconceptos. Si trabajáramos de esa manera con las entidades con que vinculamos las clases léxicas, podrían surgir nuevos conceptos terminales y subconceptos; por ejemplo, sería posible proponer un concepto terminal, subordinado de +FEELING_00, que especificara que se trata de sentimientos positivos en grado elevado. No asumimos el trabajo de proponer nuevas entidades conceptuales, porque está más allá del alcance de esta tesis y porque FunGramKB –a pesar de tener una granularidad bastante fina en comparación con otras bases de conocimiento– todavía no alcanza parámetros tan específicos como los que mencionamos; incluirlos provocaría una

verdadera explosión de conceptos que haría mucho más lento el funcionamiento del razonador, es decir, de los mecanismos de inferencia.

1.2.2 Los argumentos

Los argumentos constituyen la información léxica de la clase léxica, y los tomamos en cuenta junto con el descriptor al intentar vincular la clase con un concepto de FunGramKB. La consideración de los sustantivos miembros de las clases léxicas conlleva un trabajo minucioso. En primer lugar, no estamos ante una sola unidad léxica, sino ante un conjunto de unidades, muchas de ellas polisémicas. Por lo tanto, hay que considerar cada uno de los argumentos miembros de la clase, para determinar si pueden lexicalizar el concepto identificado o no, y por qué. Frecuentemente, aunque el vínculo entre el descriptor de la clase léxica de REDES y el concepto de FunGramKB sea satisfactorio, muchos argumentos no pueden lexicalizar el concepto. Recordemos que se trata de clases léxicas sensibles a la combinatoria, lo que significa que sus miembros comparten un sentido en común en virtud de la combinación con el predicado, pero no necesariamente los rasgos inherentes que les permitan vincularse a un mismo concepto en la ontología. A la luz de esto, y siguiendo a Bosque (2004b), establecemos que los argumentos pueden pertenecer a tres tipos de paradigmas: semántico, discursivo o pragmático.

Los argumentos de una clase léxica pueden pertenecer al paradigma semántico que define la clase léxica por virtud de sus rasgos inherentes. Si el sentido recto del argumento lexicaliza la entidad conceptual, el vínculo es claro. Por ejemplo, *dinero* ('moneda corriente') lexicaliza a +MONEY_00 ('the most common medium of exchange; functions as legal tender'), mientras que *capacidad* ('aptitud, talento, cualidad que dispone a alguien para el buen ejercicio de algo') y *habilidad* ('capacidad y disposición para algo') lexicalizan a +ABILITY_00 ('possession of the qualities –especially mental qualities– required to do something or get something done'). En otros casos, un sentido secundario del nombre se acerca lo suficiente a la entidad conceptual. Por ejemplo, *fortuna* ('hacienda, capital, caudal') entraña una noción más específica que *dinero*, pero la podemos vincular con +MONEY_00, y *destreza* ('habilidad, arte, primor o propiedad con que se hace algo') es más específica que *capacidad*, pero la vinculamos con +ABILITY_00. Si el argumento léxico expresa un sentido que está fijado en el uso y que está representado como una acepción en los diccionarios, entonces pertenece al paradigma semántico que lexicaliza la entidad conceptual.

Si los miembros de la clase léxica no pertenecen al paradigma semántico, son entonces parte de un paradigma discursivo o un paradigma pragmático. Esto significa que el hablante les asigna un sentido dado en la combinación con el predicado, mediante un proceso de predicación implícita (Bosque, 2004b). Por lo tanto, la pertenencia a una clase léxica no equivale a la posesión inherente de un significado, sino a la manifestación en una pieza léxica de rasgos de ese significado en los contextos que los activan. Veamos a continuación ejemplos de los paradigmas discursivos y pragmáticos en las clases léxicas de REDES.

En el paradigma discursivo incluimos sentidos que se asignan a las palabras a través de procesos de metonimia, metáfora y otras extensiones semánticas. Las extensiones metafóricas y metonímicas –procesos absolutamente regulares en el idioma– constituyen una reconocida dificultad para el PLN, por el reto que supone su vinculación con un inventario de sentidos. Por ejemplo, el verbo *malograr* (‘no llegar al término deseado, fracasar’) selecciona una clase léxica que hemos vinculado con +PLAN_00 (‘a series of steps to be carried out or goals to be accomplished’). A esta clase pertenecen los argumentos *objetivo* (‘fin o intento a que se dirige o encamina una acción u operación’) y *esperanza* (‘estado del ánimo que nos presenta como posible lo que deseamos’), cuyos sentidos rectos no lexicalizan a +PLAN_00. Los diccionarios no definirían *objetivo* y *esperanza* como ‘plan’, pero no es difícil percibir que hayan ganado ese valor en combinación con *malograr*, por un proceso que percibimos como metonímico. Otro ejemplo de un paradigma discursivo figura en una clase léxica que se combina con *malgastar* y que hemos vinculado a +INFORMATION_OBJECT_00. Además de *palabra*, *línea* y *página*, que son subordinados de +INFORMATION_OBJECT_00, la clase incluye los argumentos *saliva*, *óleo*, *tinta*, *cinta*. Es difícil que una ontología conceptual identifique esas palabras por sí solas con +INFORMATION_OBJECT. Sin embargo, en combinación con *malgastar*, por un proceso metafórico, no hay duda de que *saliva*, *óleo*, *tinta*, *cinta* pueden constituir ‘objetos de información’.

A veces REDES explicita el hecho de que algunos miembros de clases léxicas expresan sentidos figurados y generan sentidos figurados en combinación con el predicado: *cobrar* se combina con la clase léxica de sustantivos que designan ‘diversas magnitudes o propiedades físicas de los cuerpos, en especial las relativas al tamaño, más frecuentemente en sentido figurado’ (*dimensión*, *envergadura*, *proporción*, *volumen*, *peso*) y *dictar* con la clase léxica de sustantivos que denotan ‘inclinación o dirección, generalmente en sentido figurado’ (*dirección*, *tendencia*, *rumbo*). Esto presenta un gran reto para una base de conocimiento

como FunGramKB: básicamente, la unidad conceptual misma, por ejemplo, la que corresponde a la noción de ‘dirección’, tendría que tener un PS que admitiera su uso figurado, no solo su uso recto. Si ello se hiciera así, las unidades léxicas podrían vincularse a ese concepto que incluye tanto la noción literal como la noción figurada de ‘dirección’. El hecho es que el manejo de los sentidos figurados en relación con los sentidos rectos o físicos sigue siendo un asunto pendiente de dilucidar en la lexicografía tradicional, y presenta complejidades especiales para la nueva lexicografía, que tiene que ofrecer definiciones cerradas de los conceptos.

En el caso de los paradigmas pragmáticos, las palabras adquieren un sentido dado, no solo por la combinación con un predicado, sino por el contexto situacional. Este es el caso de *balón* y *punto*, que se combinan con *malgastar*, como parte de una clase léxica que expresa ‘posibilidad’:

Tabla 33. Descripción de clase léxica en REDES que se combina con *malgastar*

| Naturaleza de la información | Componente de REDES | Información específica |
|------------------------------|---------------------|---|
| léxica | argumentos | <i>oportunidad, ocasión, balón, ventaja, punto, opción, privilegio, posibilidad</i> |
| semántica | descriptor | oportunidad u ocasión, ventaja o circunstancia favorable, frec. deportivo |
| conceptual | índice conceptual | POSIBILIDAD |

Está claro que *balón* y *punto* pasan a significar ‘posibilidad’ no solo en combinación con *malgastar*, sino también porque ocurren en el discurso deportivo y político, respectivamente, en oraciones como “Jugó con más arietes el colista que el aspirante a la Liga de Campeones, que *malgastó* el *balón* en inocentes disparos lejanos” y “[...] puede caer, pero difícilmente más de los diez *puntos* que tendría que *malgastar* para que el candidato de gobernabilidad no operara en su favor” (REDES, 2004b: sub voce *malgastar*).

Este tipo de coerción semántica resultante de la composicionalidad, al igual que las extensiones metafóricas y metonímicas antes vistas, no se pueden vincular directamente a conceptos ontológicos. En su mayoría, estos argumentos se ubicarán en la plantilla léxica del verbo, en el módulo léxico de FunGramKB, como colocaciones. (Las colocaciones se discuten más a fondo en la sección 1.6 de este capítulo).

En este apartado hemos dado algunos ejemplos de los procesos de metonimia, metáfora y usos pragmáticos que pueden ocurrir en los miembros de las clases léxicas. Sin embargo, en esta tesis no detallaremos los mecanismos específicos –metáfora, metonimia,

extensión— por los cuales las palabras adquieren un sentido figurado. Nos limitaremos a hacer la distinción entre el paradigma semántico, de una parte, y los paradigmas discursivo o pragmático, de otra, cuando analicemos los argumentos miembros de las clases léxicas.

1.2.3 El índice conceptual

Los índices conceptuales de REDES ofrecen una guía útil para orientar la selección de un concepto en FunGramKB. Hay que recordar que los índices conceptuales de REDES fueron construidos *a posteriori* de la recogida de datos y no responden a una estructura taxonómica que relacione formalmente los conceptos entre sí ni que provea definiciones. Como consecuencia, los conceptos del índice conceptual de REDES no necesariamente comparten la misma naturaleza semántica entre sí⁷⁹. Veamos algunos ejemplos.

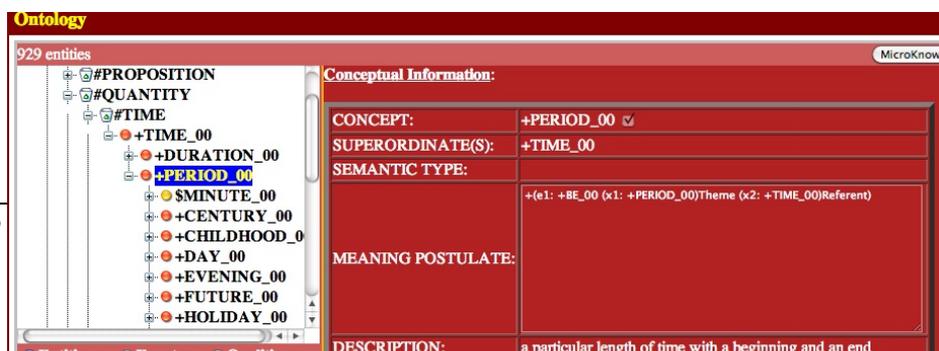
Los conceptos de REDES a veces se corresponden a un género próximo en una taxonomía de relaciones de subsunción. Ese tipo de indicación nos dirige rápidamente a un concepto en REDES. Este es el caso de la siguiente clase léxica, que se combina con *malgastar*. El descriptor ofrece el hiperónimo ‘unidades temporales’, y la clase se identifica con el concepto TIEMPO en REDES:

Tabla 34. Descripción de clase léxica en REDES que se combina con *malgastar*

| Naturaleza de la información | Componente de REDES | Información específica |
|------------------------------|---------------------|--|
| léxica | argumentos | <i>vida, juventud, día</i> |
| semántica | descriptor | unidades temporales, incluidas las que se refieren a la existencia |
| conceptual | índice conceptual | TIEMPO |

En FunGramKB, el concepto correspondiente sería +PERIOD_00, que es subordinado de +TIME_00 y se describe como ‘un largo de tiempo particular, con principio y fin’. Como se ilustra en la imagen siguiente, +PERIOD_00 tiene como conceptos subordinados a periodos de tiempo específicos, como +DAY_00 (‘día’) y +CHILDHOOD_00 (‘niñez’), por lo que entendemos que puede abarcar a los argumentos *vida, juventud, día*:

Imagen 11. +PERIOD_00 en FunGramKB (vista parcial)



79

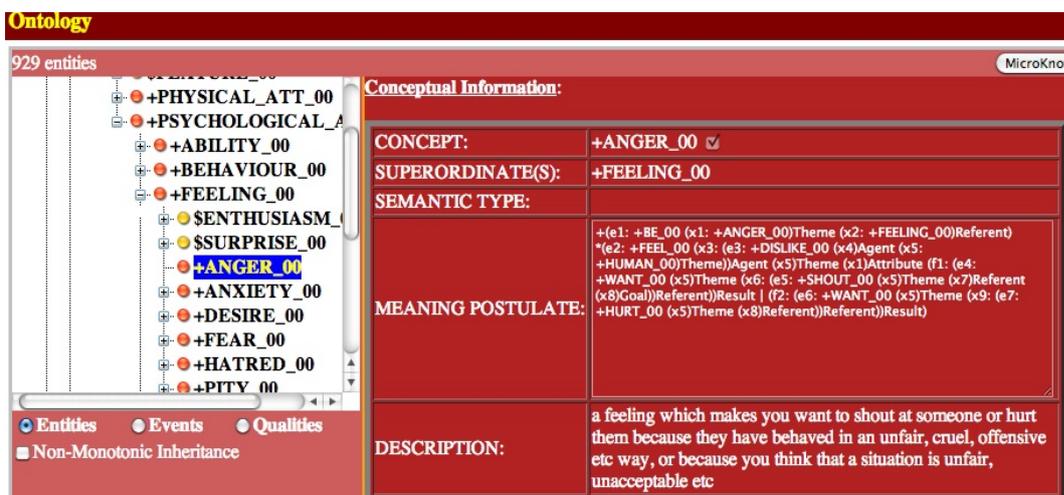
En otras ocasiones el concepto de REDES incluye a la vez lo que serían dos conceptos en FunGramKB: un superordinado y un subordinado. Por ejemplo, REDES asocia la siguiente clase léxica que se combina con *disipar* al concepto SENTIMIENTO HOSTIL, que es una precisión de la noción de ‘sentimiento’.

Tabla 35. Descripción de clase léxica en REDES que se combina con *disipar* (‘hacer desaparecer’)

| Naturaleza de la información | Componente de REDES | Información específica |
|------------------------------|---------------------|---|
| léxica | argumentos | <i>ira, resentimiento, hostilidad, crispación</i> |
| semántica | descriptor | sentimientos de irritación o de animadversión |
| conceptual | índice conceptual | SENTIMIENTO HOSTIL |

Vinculamos esta clase en FunGramKB a la entidad +ANGER_00, cuyo superordinado es +FEELING_00. Aunque REDES lo codifica en dos palabras (SENTIMIENTO y HOSTIL), FunGramKB lo captura en un concepto (+ANGER_00):

Imagen 12. +ANGER_00 en FunGramKB (vista parcial)



Otras veces, el índice conceptual de REDES utiliza un término que describe el papel o finalidad que puede adquirir la entidad descrita. En la tabla siguiente se describe otra clase léxica que se combina con *malgastar* y se identifica en REDES con el concepto RECURSO. Este concepto refleja una posible finalidad del *agua* –uno de los miembros de la clase–, pero su género próximo en una taxonomía de relaciones IS-A sería ‘líquido’. De hecho, una

ontología como FunGramKB no incluye –ni puede incluir– una noción como ‘recurso’, pues demasiadas cosas se podrían considerar ‘recursos’, en el contexto adecuado, como lo evidencian los miembros de la clase: *dinero, agua, corcho, petróleo, gasolina, papel*.

Tabla 36. Descripción de clase léxica en REDES que se combina con *malgastar*

| Naturaleza de la información | Componente de REDES | Información específica |
|------------------------------|---------------------|--|
| léxica | argumentos | <i>dinero, agua, corcho, petróleo, gasolina, papel</i> |
| semántica | descriptor | magnitudes, bienes materiales, recursos |
| conceptual | índice conceptual | RECURSO |

Finalmente, a veces el concepto es un término sinónimo (o incluso idéntico) a los términos que pertenecen a la clase léxica, por ejemplo, ÉXITO, como vemos en la tabla siguiente. Siguiendo este criterio, en principio, el número de clases léxicas podría ser prácticamente igual al número de palabras de una lengua, exceptuando los sinónimos o cuasi-sinónimos.

Tabla 37. Descripción de clase léxica en REDES que se combina con *malgastar*

| Naturaleza de la información | Componente de REDES | Información específica |
|------------------------------|---------------------|---------------------------------|
| léxica | argumentos | <i>éxito, victoria, triunfo</i> |
| semántica | descriptor | resultado feliz de algo |
| conceptual | índice conceptual | ÉXITO |

1.3 Clasificar cada clase léxica según la tipología establecida

Al comparar la información léxica y semántica que aporta REDES para cada clase léxica con la información que se ofrece para cada concepto de la ontología en FunGramKB, establecimos cuatro tipos de clases léxicas que se describen a continuación:

- *Clases léxicas estables* (tipo I);
- *Clases léxicas descritas por superordinados* (tipo II);
- *Clases léxicas mixtas* (tipo III);
- *Clases léxicas que no se pueden vincular a un concepto* (tipo IV).

Aunque el descriptor semántico de REDES nos guía hacia una entidad en FunGramKB, la información léxica es la que determina qué tipo de clase es, en la medida en que los argumentos puedan o no lexicalizar el concepto de la ontología.

La distribución en los cuatro tipos de clases léxicas fue como sigue: las clases mixtas (tipo III) fueron las más frecuentes con 31%; las clases que no se podían describir por una

entidad ontológica (tipo IV) sumaron 27%; las clases estables (tipo I) tuvieron 23%; y las clases descritas por un superordinado (tipo II) tuvieron un 19%.

Cabe señalar aquí que, en el proceso de analizar las clases léxicas de REDES frente al marco ontológico de FunGramKB, pudimos constatar que la granularidad semántica de REDES y FunGramKB no son muy diferentes. Hasta cierto punto, podemos decir que las clases léxicas estables (23%), las clases léxicas descritas por un superordinado (19%) y las clases mixtas (31%) –que juntas suman 73%–, son señal de afinidad semántica. Todas son clases léxicas de REDES que se pueden describir hasta cierto punto mediante una entidad conceptual en FunGramKB. Podemos decir pues que en la mayoría de los casos estudiados (73%), las clases léxicas de REDES encuentran alguna representación en las entidades conceptuales de FunGramKB. Por otro lado, también es cierto que hubo 27% de clases léxicas del tipo IV, que no se pudieron vincular a la ontología tal cual estaban formuladas. En muchos casos, por supuesto, la granularidad de REDES fue más fina que la de FunGramKB, pero, en ocasiones, FunGramKB resultó tener una granularidad más fina que la de REDES. Esto último nos llevó en ocasiones a desdoblarse una clase léxica de REDES en dos entidades de FunGramKB. Veremos ejemplos de la granularidad semántica de cada recurso a lo largo de este capítulo.

Tipo I: Clases léxicas estables

Las clases léxicas que hemos llamado *estables* contienen “paradigmas de base semántica” (Bosque, 2004b), o sea, miembros que comparten inherentemente ciertos rasgos semánticos. En estos casos, el descriptor de REDES parece satisfactoriamente con el PS del concepto identificado en FunGramKB, y los argumentos pueden considerarse lexicalizaciones del mismo. Esto quiere decir que podríamos sustituir la clase y sus argumentos por el concepto, incluso fuera de la combinación particular. En nuestra muestra, las clases léxicas estables representaron un 23%.

Por ejemplo, el predicado *derrochar* ‘manifestar en abundancia’ se combina con una clase compuesta por nombres que designan ‘valor, coraje, a veces extremos’ (*valor, valentía, arrojo, bravura, coraje, heroísmo*), que vinculamos al concepto +COURAGE_00 en FunGramKB. En la tabla siguiente comparamos la información léxico conceptual de REDES con la de FunGramKB:

Tabla 38. Comparación entre clase léxica de REDES y +COURAGE_00 en FunGramKB

| Tipo | de | REDES | FUNGRAMKB |
|------|----|-------|-----------|
|------|----|-------|-----------|

| | | |
|-----------------------|---|---|
| información | | |
| Superodinado: | --- | +BEHAVIOUR_00 |
| Concepto: | CUALIDAD | +COURAGE_00 |
| PS: | --- | +(e1: +BE_00 (x1: +COURAGE_00)Theme (x2: +BEHAVIOUR_00)Referent) *((e2: +DO_00 (x3: +HUMAN_00)Theme (x4)Referent (f1: x1)Manner)(e3: +BE_01 (x4)Theme (x5: +DANGEROUS_00)Attribute)) |
| Descripción: | valor, coraje, a veces extremos | the quality of being brave when you are in danger, in pain, in a difficult situation etc |
| Manifestación léxica: | <i>valor, valentía, arrojo, bravura, coraje, heroísmo</i> | <i>agalla, arrojo, audacia, braveza, bravura, coraje, denuedo, huevo, intrepidez, valentía, valor</i> |

Al comparar la información léxica, semántica y conceptual de REDES con la descripción del concepto +COURAGE_00 en FunGramKB, encontramos suficiente similitud como para vincular la clase al concepto. El concepto CUALIDAD es más amplio que +COURAGE_00, pero el descriptor semántico ‘valor, coraje, a veces extremos’ se acerca a la descripción de la entidad. La mayoría de los miembros de la clase léxica –*valor, valentía, arrojo, bravura, coraje*– son ya lexicalizaciones de +COURAGE_00, y se corresponden por lo tanto con su información conceptual. Aunque la descripción de +COURAGE_00 especifica una situación de peligro o dificultad, que va más allá de la descripción de REDES, la marca de * en el PS nos indica que esta es una condición opcional. Solo habría que añadir *heroísmo* como lexicalización de +COURAGE_00⁸⁰.

También es estable la clase siguiente, que se combina con *perder* ‘fracasar’ y que vinculamos a +TRIAL_00. Ya *juicio* es lexicalización de +TRIAL_00, y podemos añadir *pleito* y *apelación* como lexicalizaciones adicionales. Las descripciones de la clase léxica y de la entidad coinciden lo suficiente como hacer un vínculo estable entre ellas:

Tabla 39. Comparación entre clase léxica de REDES y +TRIAL_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|---------------------|---------------|---|
| Superodinado: | --- | +OCCURRENCE_00 |
| Concepto: | CONFRONTACIÓN | +TRIAL_00 |
| PS: | --- | +(e1: +BE_00 (x1: +TRIAL_00)Theme (x2: +OCCURRENCE_00)Referent) ⁸¹ |

⁸⁰ Como indicamos en el Capítulo IV, la integración de información léxica en los lexicones de FunGramKB es un trabajo en progreso, por lo que no se deben interpretar los datos léxicos como un conjunto cerrado o exhaustivo.

⁸¹ Copiamos este PS según aparece de FunGramKB (fungramkb.com), pero es posible que este incompleto.

| | | |
|------------------------------|---|--|
| Descripción: | juicio, y otros procesos de naturaleza jurídica | a formal legal process in which a judge and jury decide whether someone is guilty of a particular crime by questioning them and considering the evidence |
| Manifestación léxica: | <i>pleito, apelación</i> <i>juicio,</i> | <i>juicio, proceso</i> (español), <i>trial</i> (inglés) |

Hay conceptos en FunGramKB cuya granularidad resulta más fina que la de las clases léxicas de REDES. Esto puede obligar a desdoblar una clase léxica en dos conceptos. Las clases resultantes pueden ser clases estables con respecto al concepto que le corresponde. Por ejemplo, con *disipar* ‘hacer desaparecer’ se combina la clase descrita como ‘situaciones difíciles, peligrosas, adversas o conflictivas’, con los miembros: *peligro, problema, dificultad, amenaza, crisis, polémica, enfrentamiento, tensión*. Esta clase léxica se desdobra en +DANGER_00 (‘the condition of being susceptible to harm or injury: you are in no danger’) y +PROBLEM_00 (‘a factor causing trouble in achieving a positive result or tending to produce a negative result’: *serious difficulties were encountered in obtaining a pure reagent*). El descriptor y los argumentos de la clase léxica de REDES incluyen tanto la noción de ‘peligro’ (*peligro, amenaza*) como la de ‘problema’ (*problema, dificultad, crisis, polémica, enfrentamiento, tensión*). Hay cercanía entre los conceptos, porque comparten el superordinado +STATE_00.

Antes de concluir el apartado sobre las clases léxicas estables, debemos indicar que para determinar definitivamente que una clase léxica es estable y validar el vínculo con el concepto en FunGramKB, tendríamos que escudriñar el comportamiento de la clase léxica más allá de la combinación con el predicado particular. Habría que confirmar, por ejemplo, que los miembros de la clase léxica se comportan del mismo modo en cualquier otra combinación en que participe el concepto +TRIAL_00, más allá de la combinación con *perder* ‘fracasar’. Esta comprobación es necesaria dado el comportamiento variable de los miembros de las clases léxicas, que Bosque ha llamado los “problemas de dispersión de una entrada genérica” (2006: XXXI). Se trata del caso de *guitarra* y *piano*, que discutimos en el capítulo I, que pueden identificarse con la clase de ‘instrumentos de cuerda’, pero no por eso tienen la misma combinatoria siempre: aunque ambos *se afinan* o *se tocan*, uno *se sienta* al piano y *rasguea* la guitarra. En otros casos, hay equivalencia entre las partes semántica y conceptual de la clase léxica y la entidad, pero la entidad puede dar cabida a más lexicalizaciones de las que ocurren en el uso. Bosque (2006) da el ejemplo de los adjetivos *caballar, ecuestre, equino, hípico*, que estarían vinculados a un mismo concepto (‘relativo a

los caballos’), pero no se usan en los mismos contextos: *ganado caballar, club ecuestre, peste equina, federación hípica*. Finalmente, este cotejo habría que extenderlo también a todos los hipónimos léxicos de todos los subordinados conceptuales; recordemos que cuando vinculamos una clase léxica a un concepto, en principio estamos incluyendo todos los conceptos subordinados, por la relación de subsunción con que se estructura la taxonomía. Esta tesis proveerá un primer acercamiento para identificar posibles clases léxicas estables, pero no realizará el análisis exhaustivo requerido para poder concluirlo en definitiva.

Si la clase léxica es estable, y se representa como preferencia de selección en un evento, entonces no será necesario incorporar ninguno de los argumentos en la plantilla léxica del predicado como colocaciones.

Tipo II: Clase léxicas descritas por superordinados

Hay clases léxicas cuyos miembros solo pueden ser descritos colectivamente por medio de un superordinado común. Esto quiere decir que los miembros de la clase están vinculados a diferentes conceptos en FunGramKB, que son subordinados a un mismo concepto superordinado. En nuestra muestra, las clases léxicas descritas por un superordinado sumaron un 19%.

Por ejemplo, el predicado *dilapidar* ‘malgastar’ se combina con la clase *vida, tiempo, historia, año, carrera, hora, instante, momento, futuro*. El descriptor es ‘sustantivos temporales’ y el concepto es TIEMPO en REDES. Los argumentos se vinculan individualmente a los conceptos que se detallan en la tabla y que comparten el superordinado +PERIOD_00:

Tabla 40. Miembros de clase léxica que se combina con *dilapidar* y su vinculación con entidades de FunGramKB

| Argumento | Concepto al que lexicaliza | Concepto Superordinado |
|------------------|-----------------------------------|-------------------------------|
| <i>tiempo</i> | +PERIOD_00 | |
| <i>año</i> | +YEAR_00 | +PERIOD_00 |
| <i>hora</i> | +HOUR_00 | +PERIOD_00 |
| <i>momento</i> | +MOMENT_00 | +PERIOD_00 |
| <i>instante</i> | +MOMENT_00 | +PERIOD_00 |
| <i>futuro</i> | +FUTURE_00 | +PERIOD_00 |
| <i>historia</i> | +PAST_00 | +PERIOD_00 |

Podemos decir entonces que *dilapidar* ‘malgastar’ selecciona a +PERIOD_00. Sin embargo, no todas las lexicalizaciones de +PERIOD_00 ni todos sus subordinados necesariamente se podrán combinar con *dilapidar*. A pesar de esto, hemos decidido que si la representación conceptual de la clase léxica se hace mediante superordinado, daremos por buena la representación conceptual y no integraremos los argumentos al lexicón como colocaciones de *dilapidar*. Esta clase léxica también incluye *vida* y *carrera*, que asumen un valor de ‘periodo de tiempo’ de manera figurada, y ya tienen una asignación conceptual que corresponde a su sentido recto: *vida* a +LIFE_00 y *carrera* a +WORK_00. Los argumentos *vida* y *carrera* sí entran en el lexicón como colocaciones de *dilapidar*, porque no están representados por la entidad +PERIOD_00.

Otro ejemplo lo constituye una clase léxica que se combina con *perder* en el sentido de ‘no llegar a tiempo’. La clase se describe en REDES como ‘medios de transporte’ y la vinculamos en FunGramKB con +VEHICLE_00. Sin embargo, cada uno de sus miembros es en realidad un subordinado de +VEHICLE_00: *avión*[+AIRCRAFT_00], *tren*[+TRAIN_00], *autobús* [+BUS_00], *barco*[+SHIP_00]. En este caso, al establecer como preferencia de selección la entidad +VEHICLE_00, los argumentos no tendrán que figurar como colocaciones en el lexicón.

Tipo III: Clase léxicas mixtas

Las *clases mixtas* son las más frecuentes en nuestra muestra, con un 31%. Los miembros de estas clases manifiestan un denominador común semántico conceptual en virtud de su combinación con un mismo predicado, pero un análisis cuidadoso de los miembros de la clase revela que no todos (o ninguno) pueden lexicalizar el concepto. Esto quiere decir que algunos miembros de la clase léxica responden a paradigmas de base pragmática o discursiva, o sea, adquieren el sentido de la clase de manera figurada o por el contexto comunicativo o situacional. Por lo tanto, no pueden lexicalizar el concepto en la ontología ni mantener la cohesión de clase en otras combinaciones. En estos casos, se valida la selección conceptual, pero parte o toda la selección léxica tendrá que ser codificada en el lexicón como colocaciones de predicados particulares.

Por ejemplo, con *desaprovechar* ‘malgastar’ se combina una clase que vinculamos con la entidad +ABILITY_00. En la próxima tabla, mostramos una comparación entre la información de la clase léxica y la de la entidad conceptual. Algunos argumentos –*talento, capacidad, facultad, facilidad, potencial, cualidad*– lexicalizan el concepto +ABILITY_00, pero otros –*conocimiento, disposición, carrera*– adquieren ese valor de manera figurada en

combinación con *desaprovechar*. Los argumentos *conocimiento*, *disposición* podrían ser subordinados de +ABILITY, pero *carrera* es más complicado: se refiere a una profesión desempeñada a través de una extensión de tiempo; su vínculo con la idea de ‘habilidad’ ocurre a través de una extensión semántica y, por lo tanto, no puede formalizarse en la ontología.

Tabla 41. Comparación entre clase léxica de REDES y +ABILITY_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|------------------------|---|--|
| Superordinado: | --- | +PSYCHOLOGICAL_ATT_00 |
| Concepto: | (CAPACIDAD, FACULTAD) | +ABILITY_00 |
| PS: | --- | +(e1: +BE_00 (x1: +ABILITY_00)Theme (x2: +PSYCHOLOGICAL_ATT_00) Referent) *(e2: pos +DO_00 (x3: +HUMAN_00 ^ +ANIMAL_00)Theme (x4)Referent (f1: x1)Means) |
| Descripción semántica: | --- | possession of the qualities (especially mental qualities) required to do something or get something done; "danger heightened his powers of discrimination" |
| Manifestación léxica: | <i>talento, facultad, conocimiento, disposición, potencial</i> <i>capacidad, facilidad, cualidad, carrera,</i> | <i>aptitud, capacidad, competencia, destreza, don, facultad, talento (español), ability, capability, capacity, competence, faculty, gift, power, talent (inglés)</i> |

En este caso, los argumentos *conocimiento*, *disposición*, *carrera*, que no se pueden vincular formalmente con +ABILITY_00, figurarán como colocaciones en el lexicon.

Tipo IV: Clases léxicas que no se pueden vincular a un concepto

Hubo algunas clases léxicas que no pudimos vincular satisfactoriamente a ningún concepto de FunGramKB. Las nociones semánticas que describen estas clases en REDES no tienen equivalente en la ontología, o están tan dispersas que no se pueden vincular con ningún concepto particular en FunGramKB. En estos casos, todos los argumentos miembros de la clase se ubican como colocaciones en la plantilla léxica. En nuestra muestra, las clases que no se pueden vincular a un concepto sumaron 27%.

Este es el caso de la clase léxica que se combina con *malgastar* y que REDES describe como ‘magnitudes, bienes materiales, recursos’ y asocia al concepto RECURSO; los argumentos que la componen son *agua, petróleo, gasolina, papel, dinero, corcho*. El único

concepto que podría dar cabida a todos esos argumentos, como concepto superordinado, sería +SUBSTANCE_00, pero estaría respondiendo al significado recto de las palabras. La noción de ‘recurso’ la asignan los hablantes a estas sustancias, dado que en el plano pragmático, lo son. No existe, ni puede existir, en FunGramKB un concepto así. Estos argumentos tendrían que ser vaciados en su totalidad en el lexicón, como colocaciones del verbo, para efectos de la codificación en FunGramKB.

Otro caso lo constituye una clase que se combina con *derrochar* ‘manifestar en abundancia’. Se describe como ‘saber hacer, profesionalidad y rasgos de experiencia’ y se asocia al concepto CUALIDAD en REDES. Los argumentos son *soltura, oficio, tablas, elocuencia, dominio*. Lo que dificulta la vinculación con un concepto en este caso es la multiplicidad de nociones del descriptor y la heterogeneidad de los argumentos miembros. En este caso, esta clase no se puede vincular a ninguna entidad, y esos argumentos tienen que ser vaciados en el apartado de colocaciones, en la plantilla léxica de *derrochar*.

1.4 Analizar colectivamente las clases léxicas similares

Cuando varios predicados seleccionan clases léxicas similares, evaluamos las diferentes clases en conjunto. Esto tiene un doble propósito: comparar los patrones de selección léxica conceptual entre predicados relacionados semánticamente y evaluar la posible reducción del inventario de clases léxicas. Por ejemplo, *malgastar, dilapidar, malograr, derrochar, desperdiciar y despilfarrar* seleccionan clases que expresan las nociones de ‘habilidad, capacidad’, a pesar de no ser idénticas. La comparación cuidadosa confirma que esas seis clases léxicas diferentes se pueden vincular a una misma entidad en FunGramKB: +ABILITY_00; también nos muestra en qué se diferencia cada una. Por otro lado, confirma que estos seis verbos, en su sentido de ‘malgastar’ –vinculado a +WASTE_00– comparten una preferencia de selección por +ABILITY_00. Esta evidencia apoya la creación del subconcepto –WASTE_ABILITY, proceso que se explicará a continuación.

El trabajo comparativo con las clases léxicas nos da una idea también de las posibilidades de reducir y formalizar el inventario de clases léxicas. Mientras más clases se puedan vincular a una misma entidad, más se reduce el inventario. Sin embargo, tenemos que mirar la tipología de clases con cuidado, antes de llegar a conclusiones definitivas.

Las clases del tipo I nos ofrecen un primer indicio de lo que podrían ser clases léxico semánticas estables en la lengua, que se podrían representar cabalmente con una entidad

conceptual. Pero para confirmar esto tendríamos que escudriñar el comportamiento de la clase léxica más allá de la combinación con el predicado particular. Habría que mirar todas las otras lexicalizaciones de la entidad, y asegurar que tuvieran el mismo comportamiento que la clase léxica de REDES. Habría que auscultar además el comportamiento de los subordinados conceptuales y sus lexicalizaciones; recordemos que cuando hablamos de un concepto, incluimos a todos los conceptos subordinados, por la estructura de subsunción de la taxonomía.

Las clases del tipo II también requieren un estudio particular: como están representadas por un concepto superordinado, hay que cotejar que este no resulte muy amplio para representar el comportamiento específico de la clase; además, hay que auscultar el comportamiento de las rutas conceptuales subordinadas.

Los miembros de las clases del tipo III, por su parte, están divididas entre un paradigma semántico y un paradigma discursivo o pragmático. Al mirar los miembros de base semántica que se vinculan al concepto, habría que hacer el cotejo descrito para las clases del tipo I. Los miembros de base discursiva o pragmática de las clase tipo III, al igual que los miembros de las clases tipo IV, se ubicarían como colocaciones en la plantilla léxica del predicado en el lexicón.

En esta tesis, ofrecemos un primer acercamiento a la reducción del inventario de clases léxicas de REDES, pero no un análisis exhaustivo.

1.5 Vincular cada predicado con un concepto de la ontología

1.5.1 Establecer los sentidos del predicado

Antes de vincular el predicado formalmente con conceptos de la ontología, identificamos todos sus sentidos. Intentamos proponer el menor número posible de sentidos que capture las distinciones relevantes para una base de conocimiento como FunGramKB. En ello seguimos la recomendación de Nirenburg y Rasking (2004) de que, para efectos del PLN, se deben fusionar tantos significados como sea posible y mantener el número de significados más reducido posible. Para ello tomamos en cuenta:

- (a) la información provista por REDES,
- (b) la información provista por cinco diccionarios del español y
- (c) nuestro conocimiento e intuición lingüísticos a la luz de la información provista por REDES.

Aunque no es su objetivo definir los predicados, cuando es necesario REDES ofrece indicaciones breves sobre sus diferentes sentidos, según comentamos en el capítulo I. Se trata de marcas léxicas simples, de granularidad más gruesa que las de un diccionario tradicional; de hecho, los sentidos que distingue REDES se acercan más a lo que sería apropiado en una base de conocimiento léxico. A veces REDES solo distingue entre el ‘sentido recto’ y el ‘sentido figurado’ del predicado. En otras ocasiones, usualmente cuando el predicado es monosémico, REDES no distingue ningún sentido.

Consideramos también la información provista por cinco diccionarios de la lengua española: DUE, Salamanca, VOX, CLAVE y DRAE, tanto la separación en acepciones, como el contenido de cada definición (semántico y léxico). Los diccionarios se utilizan como guía en la creación de instrumentos para el PLN porque se consideran repositorios fiables de información sobre el significado léxico de las palabras, y en general están avalados por generaciones de hablantes; se consultan varios diccionarios de modo que uno salve las lagunas del otro (Periñán y Arcas, 2005). Los diccionarios nos muestran, además, el manejo lexicográfico que actualmente reciben las unidades en cuestión.

Luego de analizar los artículos lexicográficos, reagrupamos el contenido en el menor número posible de sentidos. Para cada sentido identificado, dividimos el contenido lexicográfico en cuatro categorías: hiperónimos, contornos, colocaciones y sinónimos. La tabla siguiente muestra la información colectada de los cinco diccionarios para el sentido de ‘desperdiciar’ de *perder*:

Tabla 42. Resumen de información lexicográfica de *perder* (‘desperdiciar’)⁸²

| |
|--|
| Hiperónimos: |
| desperdiciar (x3), no aprovechar (debidamente) (x3), disipar o malgastar (x2), no obtener provecho, no ser útil, emplear (mal o de manera inútil), emplear (de mala manera), aplicarse (mal), no conseguir, realizar u obtener, gastarlo o dejarlo pasar |
| Contornos: |
| una persona (x2) // una cosa (x2), una cosa que deseaba, algo, algo que podía y debía ser útil, una oportunidad, cierta cosa que podía serlo, de cierta cosa // sin hacer nada de provecho o sin que lo que se hace conduzca al resultado deseado, para otro fin |
| Colocaciones: |
| <i>el tiempo</i> (x2), <i>nuestro tiempo</i> , <i>tres años</i> , <i>la ocasión</i> , <i>la mejor ocasión</i> , <i>la ocasión de tu vida</i> , <i>el agua</i> , <i>sus palabras</i> (x2), <i>nuestro dinero</i> |
| Sinónimos: |
| <i>desperdiciar</i> (x4), <i>malgastar</i> |

⁸² Usamos las diagonales dobles (//) para separar los diferentes tipos de contornos y corchetes para explicitar el argumento de algún componente que así lo requiera para completar su significado. Cuando se repite un componente, destacamos la frecuencia en el paréntesis (x__)

La información de los diccionarios ofrece también una buena referencia de comparación para los datos REDES y de FunGramKB, con los que tiene puntos en común. El hiperónimo lexicográfico, por ejemplo, se acerca al superordinado conceptual de FunGramKB. El contorno lexicográfico es afín al descriptor de la clase léxica en REDES y a las preferencias de selección de FunGramKB. Los ejemplos de colocaciones frecuentes de los diccionarios pueden compararse con los argumentos miembros de las clases léxicas y a las colocaciones de FunGramKB. Finalmente, los sinónimos nos informan de otros predicados relacionados semánticamente.

1.5.2 Identificar el concepto correspondiente en FunGramKB

A partir de las nociones generales sobre los sentidos de los predicados, procedemos a examinar los conceptos correspondientes en la ontología de FunGramKB. Como son sentidos de predicados verbales, los conceptos pertenecerán a la subontología de eventos. En el caso de palabras monosémicas, el predicado se vincula a un concepto, y en el de las palabras polisémicas, a varios. Por ejemplo, *malgastar* tiene un sentido general de ‘desperdiciar, usar mal o descuidadamente’, que vinculamos con el concepto +WASTE_00, que establece que un Tema humano usa un Referente no especificado de una Manera descuidada:

Tabla 43. Descripción de +WASTE_00 en FunGramKB

| | |
|-----------------|--|
| Hiperónimo: | +USE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2)Referent |
| PS: | +(e1: +USE_00 (x1)Theme (x2)Referent (f1: +CARELESS_00)Manner) |
| Descripción: | use inefficiently or inappropriately; "waste heat" |
| Lexicalización: | <i>malgastar</i> (inglés), <i>waste</i> (español) |

Un verbo polisémico como *perder*, por su parte, se vincula con +WASTE_00, +FAIL_00, +LOSE_00, +LACK_00 y \$MISS_00, además de unos cuantos subconceptos más. Esto se discutirá a fondo en el próximo capítulo.

1.5.3 Comparar las restricciones de selección léxica del verbo con las preferencias de selección del concepto

Para los procedimientos relativos a la codificación de preferencias de selección y subconceptos en FunGramKB, hemos usado como guía general el método establecido por

Jiménez y Pérez (2011), Jiménez y Luzondo (2011) y Luzondo (2011), haciendo algunas modificaciones en correspondencia con nuestros materiales y nuestros objetivos. En esta etapa la comparación se hace más detallada, pues tomamos en cuenta las restricciones de selección léxica de cada predicado, frente a las preferencias de selección descritas en el MT y PS de cada concepto en FunGramKB. Dado que las clases léxicas –que manifiestan las restricciones de selección léxica de cada predicado–, ya han sido vinculadas a entidades conceptuales, podemos compararlas con las preferencias de selección conceptuales que están codificadas en FunGramKB.

Como hemos dicho antes, las preferencias de selección del concepto son especificaciones semánticas de los participantes en dichos esquemas conceptuales. Por ejemplo, el sentido ‘adquirir’ de *cobrar* estaría vinculado al evento conceptual +OBTAIN_00. Como se muestra en la tabla siguiente, las preferencias de selección conceptuales de +OBTAIN_00 son +HUMAN_00 para el Tema y +ARTEFACT_00 para el Referente. El PS de +OBTAIN_00 expresa que esa entidad típicamente humana (x1) realiza algo para tener esa entidad que es típicamente un artefacto (x2), que es una entidad que no tenía previamente:

Tabla 44. Descripción de +OBTAIN_00 en FunGramKB

| | |
|------------------------|---|
| Superordinado | +DO_00 |
| MT: | (x1: +HUMAN_00)Theme (x2: +ARTEFACT_00)Referent |
| PS: | +(e1: +DO_00 (x1)Theme (x2)Referent (f1: (e2: +HAVE_00 (x1)Theme (x2)Referent))Result) +(e3: past n +HAVE_00 (x1)Theme (x2)Referent) |
| Descripción: | come into the possession of something. |
| Lexicalización: | <i>adquirir, obtener, lograr, conseguir</i> (español), <i>acquire, get, obtain</i> (inglés) |

Ahora bien, una de las clases léxicas de *cobrar*[+OBTAIN_00] la vinculamos con el concepto +STRENGHT_00. La siguiente tabla presenta la clase léxica y la entidad conceptual, lado a lado. La información semántica y conceptual de ambas unidades es bastante cercana, y suficiente para establecer un vínculo.

Tabla 45. Comparación entre clase léxica que se combina con *adquirir* y +STRENGHT_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|-----------------------|-------|------------------|
| Superordinado: | --- | +PHYSICAL_ATT_00 |

| | | |
|------------------------------|---|---|
| Concepto: | FUERZA | +STRENGHT_00 |
| PS: | --- | +(e1: +BE_00 (x1: +STRENGTH_00)Theme (x2: +PHYSICAL_ATT_00)Referent) *(e2: +BE_01 (x3: +HUMAN_00 ^ +ANIMAL_00)Theme (x1)Attribute) |
| Descripción: | fuerza, energía o empuje, físicos o anímicos | the property of being physically or mentally strong; "fatigue sapped his strenght" |
| Manifestación léxica: | <i>fuerza, brío, vigor, ánimo, impulso, ímpetu, intensidad, velocidad</i> | <i>fuerza</i> (español), <i>force, strenght</i> (inglés) |

Nos tenemos que preguntar en este momento si la selección léxica de +STRENGHT_00 documentada por REDES para *costrar* está bien representada por la preferencia de selección de +ARTEFACT_00 que establece FunGramKB. La respuesta es negativa, pues los artefactos tienen una ruta conceptual en FunGramKB como entidades materiales mientras que la fuerza sigue la ruta de los atributos físicos abstractos. Parte de nuestro trabajo será entonces lograr una representación conceptual más precisa para *costrar*, con las herramientas que nos da FunGramKB, especialmente a través de la precisión de las preferencias de selección y la consiguiente creación de subconceptos.

A continuación veamos las posibilidades que se nos presentan en este proceso. Si el sentido del predicado, y sus preferencias de selección, se recogen adecuadamente en la descripción conceptual, se vincula el predicado a ese concepto, sin hacer ajustes a sus preferencias de selección. Este es el caso de *perder* en el sentido de ‘dejar de tener algo’, que vinculamos con el evento +LOSE_00, cuyo MT es: (x1: +HUMAN_00^+ANIMAL_00)Theme (x2: +CORPUSCULAR_00)Referent. La clase léxica que se combina con *perder* en este sentido se describe como ‘cualquier objeto físico’, y se compone de *paraguas, pendientes, libro, cartera, foto*. Hemos vinculado esa clase al concepto +SOLID_00, que es subordinado directo de +CORPUSCULAR_00. No hacemos ningún cambio en las preferencias de selección de +LOSE_00, porque +CORPUSCULAR_00 ya incluye a +SOLID_00.

Si el sentido del predicado, o sus preferencias de selección, no se reflejan cabalmente en la descripción conceptual, se propone entonces la creación de un nuevo concepto terminal o de un subconcepto. Los conceptos terminales son necesarios cuando hay una restricción conceptual en el significado de un concepto básico. Estas restricciones se pueden manifestar en restricciones en las preferencias de selección y en los parámetros del PS del concepto. Este es el caso del sentido de ‘no llegar a tiempo’ que adquiere *perder* en combinación con

‘medios de transporte y sus servicios’ (*avión, tren, autobús, barco, transbordo, vuelo*). Lo más cercano en FunGramKB es +FAIL_00, cuyo PS es: +(e1: +TRY_00 (x1)Theme (x2)Referent (f1: (e2: n +DO_00 (x1)Theme (x2)Referent)))Result). Proponemos pues un concepto terminal bajo +FAIL_00 llamado \$MISS_00, que recoja la preferencia de selección por los medios de transporte y sus servicios (+VEHICLE_00 ^ +TRAVEL_00) un parámetro que exprese el hecho de no llegar a tiempo.

En el caso de los subconceptos, la restricción o especificación conceptual ocurre exclusivamente en las preferencias de selección establecidas en el MT de un concepto básico o terminal; no cambia en absoluto el PS del concepto. Los subconceptos no son visibles en la ontología. Por ejemplo, el verbo *perder* adquiere el sentido de fracasar en combinación con ‘confrontaciones, a menudo deportivas, así como algunos de sus lances’ (*batalla, lucha, partido, campeonato, encuentro, combate, guerra, pelea, final, set*). Lo vinculamos nuevamente con +FAIL_00, con las descripciones siguientes:

MT: (x1: +HUMAN_00 ^ +ANIMAL_00)Theme (x2)Referent

PS: +(e1: +TRY_00 (x1)Theme (x2)Referent (f1: (e2: n +DO_00 (x1)Theme (x2)Referent)))Result).

Luego especificamos las preferencias de selección: rellenamos el Referente no especificado por dos entidades que describen las clases léxicas: +BATTLE_00 y +COMPETITION_00. Eso resulta en la propuesta de dos subconceptos nuevos: -FAIL_BATTLE y -FAIL_COMPETITION.

Los subconceptos resultan ser un recurso extremadamente útil para codificar los datos de REDES, pues nos permiten ampliar la extensión de un concepto sin variar su intensidad. Debemos recordar que las preferencias de selección ocurren en una ontología de conceptos cognitivos universales, y no son fenómenos restringidos a una lengua. Sin embargo, aunque nuestros datos pertenecen al español, consideramos que las restricciones de selección léxica que documenta REDES –por haberse demostrado que son sistemáticas en la lengua– son indicativas de preferencias de selección semánticas que merecen ser representadas en la ontología. Por esa misma razón, no hemos fijado un requisito mínimo de lexicalizaciones para la creación de un subconcepto: con que un predicado seleccione una clase léxica es suficiente para formular un nuevo subconcepto. Naturalmente, mientras más lexicalizaciones tenga el subconcepto, más sólida será la justificación para su creación. Por ejemplo, el verbo *dilapidar* es el único cuyas restricciones de selección dan pie para la creación del

subconcepto -WASTE_FAME, mientras que un subconcepto como -WASTE_TIME está lexicalizado por seis verbos: *malgastar, desaprovechar, dilapidar, perder, desperdiciar, malograr*.

1.6 Ubicar como colocaciones los argumentos que no puedan ser descritos conceptualmente

En FunGramKB, la información de colocación léxica se ubica en el módulo léxico, como parte de la información morfosintáctica vinculada con la palabra. Se trata de información particular a cada lengua, que no puede ser descrita en el nivel de la ontología conceptual universal. Como hemos visto antes, las colocaciones se ubican en la plantilla léxica del predicado, como primer o segundo argumento del verbo.

Para efectos de nuestro trabajo, el espacio de colocación de FunGramKB nos permite codificar los argumentos miembros de las clases léxicas que no pueden ser descritos conceptualmente. Para establecer las preferencias de selección siguiendo un protocolo estándar, usamos la tipología de clases léxicas establecida anteriormente. Las clases léxicas de tipo I y II se consideran bien representadas por las entidades conceptuales, por lo que se traducen como preferencias de selección, y ninguno de sus miembros tiene que ser representado como colocación en la entrada léxica del predicado. En las clases léxicas del tipo III, los miembros que no pueden lexicalizar la entidad conceptual son ubicadas como colocaciones del predicado. En las clases del tipo IV, ningún miembro puede lexicalizar el concepto, por lo que todos son identificados como colocaciones en la entrada del predicado en el lexicón. Veamos algunos ejemplos.

En el renglón de colocaciones, se incluyen los argumentos que advienen a la clase léxica por mecanismos pragmáticos o discursivos, y que no pertenecen al paradigma semántico que lexicaliza el concepto. Por ejemplo, el verbo *dictar* se combina con dos clases léxicas que vinculamos con la entidad +PLAN_00. La tabla siguiente muestra la comparación entre las clases léxicas y la entidad conceptual:

Tabla 46. Comparación entre clases léxicas de REDES y +PLAN_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|---------------------|--------------------|--|
| Superodinado: | | +INFORMATION_00 |
| Concepto: | MODELO /// RECURSO | +PLAN_00 |
| PS: | | +(e1: +BE_00 (x1: +PLAN_00)Theme (x2: |

| | | |
|-----------------------|---|--|
| | | +INFORMATION_00)Referent) *(e2: +DO_00 (x3)Theme (x4)Referent (f1: x1)Manner) |
| Descripción: | lo que suele tomarse como referente, prototipo o paradigma de algo /// recursos, procedimientos o formas de actuar aplicables a diversos dominios | a series of steps to be carried out or goals to be accomplished; "they drew up a six-step plan"; "they discussed plans for a new bond issue" |
| Manifestación léxica: | <i>modelo, patrón, pauta</i> /// <i>estrategia, esquema, mecanismo, método, criterio</i> | <i>estrategia, plan, programa, proyecto</i> (español), <i>plan, programme, project, strategy</i> (inglés) |

El PS de +PLAN_00 lee como sigue: existe un plan (Tema) que contiene información (Referente), que, opcionalmente, se puede llevar a cabo. El PS del concepto +PLAN_00 describe satisfactoriamente el descriptor semántico de las clases léxicas, y los argumentos *modelo, patrón, pauta, estrategia* y *esquema* lo lexicalizan. A partir de eso, proponemos la creación del subconcepto –COMMAND_PLAN (esto se discutirá con más detalle en el próximo capítulo). Por su parte, los argumentos *mecanismo, método* y *criterio* adquieren el sentido de +PLAN_00 a través de extensiones semánticas. Como estos argumentos no se pueden vincular a la entidad +PLAN_00, los ubicaremos en la sección de colocaciones en la plantilla léxica del verbo *dictar* en su vínculo con –COMMAND_PLAN:

Tabla 47. Sección de plantilla léxica de *dictar* [–COMMAND_PLAN] en FunGramKB

| | | |
|-------------------|----------------------|------------------------------------|
| Lexical template: | <i>Collocations:</i> | |
| | X = | Y = |
| | | <i>mecanismo, método, criterio</i> |

El ubicar argumentos dados en la sección de colocaciones puede indicar que estos son posibles en el español, pero no necesariamente en otras lenguas. El vincular algunos argumentos en el nivel conceptual indica que son parte de una preferencia de selección conceptual, cognitiva y universal, que debe ser cierta para todas las lenguas. El vincular otros argumentos a la sección de colocaciones en la plantilla léxica del verbo indica que son fenómenos particulares³ de ese idioma. Aparte de la valoración de esa distribución, el hecho de que haya argumentos de REDES que se representen en FunGramKB conceptualmente, y otros léxicamente, asegura que no haya redundancia, pero tampoco se pierda información.

2. Polisemia regular y combinatoria

En la segunda parte de la tesis, exploramos la representación conceptual en FunGramKB de patrones de polisemia regular en verbos, y su relación con la combinatoria

léxica de esos verbos. Seguimos los pasos detallados arriba, con los ajustes que se describen a continuación. En este acercamiento, usamos una aproximación simplificada y manual de los mecanismos automatizados que utiliza Buitelaar (1998a) para identificar patrones de regularidad semántica entre sustantivos del inglés (discutido en el capítulo II). A diferencia de este investigador, que orienta su investigación desde los conceptos nominales de WordNet y luego identifica los nombres que los lexicalizan, nosotros partimos de los verbos y validamos la alternancia conceptualmente, en FunGramKB.

El primer requisito para determinar si existe un patrón de polisemia regular es que haya al menos tres verbos polisémicos que compartan al menos dos de sus sentidos. Como nuestro objetivo es validar el patrón de alternancia semántica a nivel conceptual, esos dos sentidos compartidos se tienen que poder vincular a los mismos dos conceptos en la ontología. Como hemos explicado antes, el patrón de polisemia regular puede ocurrir de manera directa, cuando los verbos mantienen un vínculo directo a los mismos conceptos, por ejemplo *tejer*, *tramar* y *urdir* se vinculan tanto con +WEAVE_00 como con +CONSPIRE_00. El patrón también puede ocurrir de manera indirecta, cuando los verbos tienen alternancia de sentidos con conceptos diferentes que comparten una misma ruta conceptual (Mairal, 2011, comunicación personal). Por ejemplo, *tejer*, *tramar*, *urdir* y *cocinar(se)* alternan entre +CREATE_00 y \$CONSPIRE_00. Aunque *tejer* tiene un vínculo directo con +CREATE_00, *tramar*, *urdir* y *cocinar(se)* lexicalizan conceptos subordinados a +CREATE_00: *tramar* y *urdir* a +WEAVE_00 y *cocinar(se)* a +COOK_00.

A continuación describimos los pasos que seguimos, los cuales, como comentamos arriba, no necesariamente se realizaron de una manera lineal:

2.1 Elegir predicados relacionados semánticamente que sugieran un patrón de polisemia regular

El primer paso fue explorar grupos de verbos relacionados semánticamente y auscultar si su polisemia revelaba patrones de alternancia regular. Identificamos un grupo de cinco verbos –*tejer*, *tramar*, *urdir*, *cocinar(se)* y *maquinar*– que alternaban entre las nociones de ‘tejer’ o ‘confeccionar algo con cuidado’ y ‘conspirar’.

2.2 Vincular cada clase léxica con un concepto de la ontología

Vinculamos cada clase léxica que se combinaba con los verbos, tomando en cuenta la información léxica provista por los argumentos, la información semántica del descriptor y la

información conceptual de la etiqueta conceptual de FunGramKB. Básicamente seguimos los pasos detallados arriba, en la sección 1.2. Luego de tener el análisis de las clases léxicas con respecto a FunGramKB, procedimos a analizar los verbos en sí.

2.3 Vincular cada predicado con un concepto de la ontología

2.3.1 Establecer los sentidos del predicado

Para efectos del estudio de la polisemia regular, hicimos un ajuste en el paso 1.3. Además de establecer los sentidos que se desprendían de REDES y de los datos de combinatoria, tomamos en cuenta sentidos etimológicos que aparecían en los diccionarios tradicionales, pero no en los datos de REDES. Nos referimos a los siguientes sentidos:

- *tramar* ‘atravesar los hilos de la trama por entre los de la urdimbre, para tejer alguna tela’;
- *urdir* ‘preparar los hilos en la urdidera para pasarlos al telar’; y
- *maquinar* ‘trabajar una pieza metalúrgica por medio de una máquina’.

2.3.2 Comparar los sentidos del predicado y sus restricciones de selección léxica, con la descripción de cada concepto y sus preferencias de selección conceptuales

Seguimos en esta etapa los pasos detallados en 1.5, arriba. Al vincular los sentidos de los verbos preliminarmente a eventos conceptuales, emergieron enseguida vínculos de varios verbos con +WEAVE_00, \$CONSPIRE_00 y +CREATE_00, que validaban conceptualmente algunos patrones de polisemia regular.

Procedimos a codificar las preferencias de selección de cada predicado. El diseño ontológico de FunGramKB nos permite auscultar la alternancia de sentidos tomando en cuenta las preferencias de selección codificadas en los subconceptos. Esto introduce un nivel más específico en el cual examinar el patrón de polisemia regular. En algunos casos, los verbos mantenían la alternancia a nivel de subconcepto, mientras que en otros, no.

Durante este proceso, algunos verbos se revelaron como verbos livianos que asumían el sentido léxico de su complemento nominal. Los habíamos vinculado preliminarmente a +CREATE_00, pero al establecer la preferencia de selección, era necesario cambiar el vínculo: +CREATE_00 y +WRITING_00 se convertían en +WRITE_00, mientras que +CREATE_00 y +EXPLANATION_00 se convertían en +EXPLAIN_00.

Para auscultar la existencia de patrones indirectos de polisemia regular, también examinamos las rutas conceptuales hacia arriba de los eventos asignados. Descubrimos que, en efecto, había coincidencias en algunos conceptos superordinados.

2.4 Ubicar como colocaciones los argumentos que no pueden ser descritos conceptualmente

Al igual que con el primer grupo de verbos, aquí también se codificaron como colocaciones los miembros de las clases léxicas que no se podían representar conceptualmente.

Conclusión

Al cabo de estos procesos, obtenemos una descripción de todas las unidades de REDES, y de las relaciones que mantienen entre sí, en términos de la ontología de FunGramKB. Los predicados se vinculan a eventos y las clases léxicas a entidades. Las restricciones de selección léxica se reflejan como preferencias de selección conceptuales en el MT del evento. Los argumentos que no pueden vincularse a entidades se codifican como colocaciones del predicado en el lexicon, en su vínculo con un evento en particular.

El vincular los datos lingüísticos de REDES al marco conceptualista de FunGramKB nos permite explorar relaciones y observar patrones, tanto en los sentidos de los predicados, como en sus preferencias de selección. Por su parte, los patrones de polisemia regular que se descubran a través de este tipo de metodología tienen el mérito de que no pueden ser explicados ni por casualidad (dos o tres palabras como mínimo lo garantizan) ni por la manipulación *ad hoc* de los datos. Dado que la distribución de las palabras en los diferentes conceptos de FunGramKB se realiza por criterios estrictamente extralingüísticos dentro de un enfoque marcadamente conceptualista –y con propósitos independientes al de esta investigación–, los patrones resultantes de polisemia regular solo podrán ser explicados porque exista realmente una regularidad semántica.

CAPÍTULO VII. RESULTADOS Y ANÁLISIS: POLISEMIA Y COMBINATORIA LÉXICA

Introducción

En este capítulo analizamos los resultados obtenidos del proceso de vincular los datos lingüísticos de REDES con el entorno ontológico de FunGramKB.

Resumimos brevemente la metodología seguida, según se discutió en el capítulo VI. El análisis individual de cada verbo de REDES comienza con la vinculación de cada sentido del predicado a un evento en la ontología de FunGramKB; para ello, tomamos en cuenta las indicaciones de REDES sobre los sentidos del verbo, las acepciones lexicográficas de los diccionarios tradicionales y el sentido común e intuición lingüística de la autora. Luego examinamos las clases léxicas con las que se combina el verbo para adquirir cada sentido, e intentamos vincular cada clase léxica a una entidad de la ontología de FunGramKB. La vinculación de clases léxicas a entidades conceptuales cumple dos propósitos: primero, informar de las posibilidades de reducir el inventario de clases léxicas, y segundo, darnos la posibilidad de traducir las restricciones de selección léxica de REDES al plano de las preferencias de selección conceptual de FunGramKB. Entonces, miramos nuevamente la descripción conceptual del evento identificado al inicio, especialmente sus preferencias de selección, y comprobamos si reflejan adecuadamente los datos de REDES. Si no es así, proponemos la creación de un nuevo concepto terminal o subconcepto. Finalmente, examinamos los argumentos miembros de las clases léxicas, para determinar si es necesario ubicar algunos en el módulo léxico, en la plantilla léxica del predicado, como colocaciones. Solo los argumentos que no estén representados conceptualmente como preferencias de selección serán incluidos en la plantilla léxica como colocaciones. Para cada uno de los verbos, terminamos visualizando el entramado de relaciones léxico semánticas que existen entre los sentidos del verbo, las clases léxicas y los argumentos. Después, en el análisis global de los datos, prestamos particular atención a los patrones comunes de selección léxica de los nueve predicados en su sentido compartido (en este caso, +WASTE_00). Comparamos, asimismo, las clases léxicas que se vinculan con una misma entidad, para auscultar la posibilidad de reducir el inventario.

En la sección 1, consideramos en detalle la interacción con FunGramKB de un verbo polisémico, *dictar*, que se combina en REDES con 8 clases léxicas. En la sección 2, miramos la interacción individual con FunGramKB de nueve predicados relacionados semánticamente: *derrochar*, *desaprovechar*, *desperdiciar*, *despilfarrar*, *dilapidar*, *disipar*, *malgastar*,

malograr y perder, y las 91 clases léxicas con que se combinan en total. En la sección 3, analizamos globalmente los nueve verbos, sus sentidos y su combinatoria, como clase de verbos que comparte –en por lo menos uno de sus sentidos– el valor semántico de ‘malgastar’. Además, miramos los patrones de selección léxica compartidos por estos verbos, y analizamos las clases léxicas colectivamente, con el objetivo de reducir el inventario.

1. DICTAR

En esta sección, analizamos los significados y combinatoria léxica del predicado polisémico *dictar*, a la luz del marco ontológico de FunGramKB, siguiendo los pasos detallados antes. REDES identifica cuatro sentidos para el verbo *dictar*: (a) ‘decir o leer algo en voz alta para que alguien lo escriba’; (b) ‘impartir’; (c) ‘promulgar’; (d) ‘precisar, determinar o sugerir’. Los sentidos y la combinatoria léxica de *dictar*, según los presenta REDES, se transcriben en la tabla siguiente⁸³. Como se puede apreciar, las clases léxicas con que se combina *dictar* se distribuyen nítidamente entre los cuatro sentidos:

⁸³ Al igual que en tablas anteriores, eliminamos los ejemplos de uso, por cuestión de espacio.

Tabla 48. Sentidos y combinatoria léxica de *dictar*, en REDES

| Sentido del verbo | CONCEPTO de la CL | Descriptor de la CL | Argumentos miembros de la CL |
|---|-----------------------|--|---|
| ‘decir o leer algo en voz alta para que alguien lo escriba’ | (TEXTO) ⁸⁴ | textos | <i>carta, texto, informe, párrafo</i> |
| ‘impartir’ | --- | lecciones docentes o disertaciones, o el conjunto de ellas | <i>clase, conferencia, lección, charla, curso, ciclo</i> |
| ‘promulgar’ | (NORMA) | resoluciones judiciales o administrativas | <i>ley, decreto, condena, auto, sentencia, fallo, norma</i> |
| ‘precisar, determinar o sugerir’ | DIRECCIÓN | inclinación o dirección, generalmente en sentido figurado | <i>dirección, tendencia, rumbo</i> |
| | MODELO | lo que suele tomarse como referente, prototipo o paradigma de algo | <i>modelo, patrón, pauta</i> |
| | CAMBIO | alteración o modificación | <i>cambio, reforma, regulación, oscilación</i> |
| | CADENCIA | pauta rítmica o proporción en el movimiento | <i>ritmo, compás, velocidad, métrica</i> |
| | RECURSO | recursos, procedimientos o formas de actuar aplicables a diversos dominios | <i>estrategia, esquema, mecanismo, método, criterio</i> |

(a) *dictar* ‘decir o leer algo en voz alta para que alguien lo escriba’

El primer sentido, ‘decir o leer algo en voz alta para que alguien lo escriba’, lo vinculamos preliminarmente en FunGramKB al evento conceptual +SAY_00, cuya descripción se muestra en la siguiente tabla. Los participantes del MT de +SAY_00 son típicamente un Tema ‘humano’, un Referente sin especificar y una Meta ‘humano’:

Tabla 49. Descripción de +SAY_00 en FunGramKB

| | |
|-----------------|--|
| Superordinado: | #COMMUNICATION |
| MT: | (x1: +HUMAN_00)Theme (x2)Referent (x3: +HUMAN_00)Goal |
| PS: | sp ⁸⁵ |
| Descripción: | to speak or tell somebody something, using words |
| Lexicalización: | <i>decir, expresar, hablar, preguntar</i> (español), <i>ask, enquire, express, inquire, mouth, say, speak, talk, tell, utter, verbalize</i> (inglés) |

⁸⁴ REDES no vincula las clases léxicas abiertas al índice conceptual, a diferencia de las clases léxicas cerradas. En los casos en que es posible, la autora hace el vínculo entre clases léxicas abiertas y el índice conceptual, y marca el concepto entre paréntesis. Si el vínculo no es posible, se marca el espacio con tres rayas: ---.

⁸⁵ La abreviatura “sp” significa “Semantic primitive”.

Este sentido de *dictar* se da en combinación con la primera clase léxica, que vinculamos a la entidad conceptual +WRITING_00. En la tabla siguiente, presentamos, lado a lado, la información de REDES sobre la clase léxica, y la de FunGramKB sobre la entidad conceptual. El PS de +WRITING_00 expresa que existe un escrito (Tema) que típicamente es un objeto de información (Referente) que está constituido por letras; opcionalmente, un humano puede leerlo.

Tabla 50. Comparación entre clase léxica de REDES y +WRITING_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|-----------------------|---------------------------------------|---|
| Superordinado: | --- ⁸⁶ | +INFORMATION_OBJECT_00 |
| Concepto: | (TEXTO) | +WRITING_00 |
| PS: | --- | +(e1: +BE_00 (x1: +WRITING_00)Theme (x2: +INFORMATION_OBJECT_00)Referent) +(e2: +COMPRISE_00 (x1)Theme (x3: +LETTER_01)Referent) *(e3: pos +READ_00 (x4: +HUMAN_00)Theme (x1)Referent) |
| Descripción: | textos | reading matter; anything expressed in letters of the alphabet |
| Manifestación léxica: | <i>carta, texto, informe, párrafo</i> | <i>escrito, texto</i> (español), <i>text, writing</i> (inglés) |

Al mirar los argumentos de esta clase léxica uno por uno, comprobamos que *texto* lexicaliza a +WRITING_00, pero los demás argumentos son subordinados de +WRITING_00: *carta*[+LETTER_00], *informe*[+DOCUMENT_00], *párrafo*. Se trata de una clase del tipo II, o sea, una clase léxica cuyos miembros solo pueden ser descritos colectivamente por medio de un superordinado común.

Resulta claro que el sentido de *dictar* de ‘decir o leer algo en voz alta para que alguien lo escriba’ requiere una precisión del concepto +SAY_00. Proponemos la creación de un nuevo concepto terminal bajo +SAY_00, llamado \$DICTATE_00⁸⁷. Recordemos que cuando hay una restricción conceptual en el significado de un concepto básico, se puede crear un nuevo concepto terminal (Mairal y Periñán, 2009). Los conceptos terminales constituyen el último nivel visible en la ontología y son precedidos por el signo de \$. Esto se traduce en

⁸⁶ También usamos las tres rayas (---) para indicar que no existe esa información. La tabla está diseñada para acomodar los conceptos superordinados y los PS de los conceptos de FunGramKB, pero REDES no incluye ese tipo de información.

⁸⁷ Los conceptos terminales y subconceptos usualmente se nombran a partir de la unidad más prototípica que los lexicaliza, o de una breve paráfrasis transparente (por ejemplo, \$SPORT_00 o TAKE_SHOES_00, respectivamente). Se usa el inglés como metalengua para la conveniencia de los ingenieros, lingüistas y usuarios, aunque una etiqueta arbitraria –como 8X92D_00– sería igualmente significativa a la máquina.

una restricción de los participantes en el MT del concepto básico y en una posible adición de parámetros en el PS. En este caso, es necesario añadir un enunciado que establezca que la Meta ‘humano’ (x3) escribe lo que dice el Tema ‘humano’ (x2). Asimismo, el concepto terminal \$DICTATE debe especificar la preferencia de selección del Referente a +WRITING_00, a partir de la información de REDES:

Tabla 51. Descripción de \$DICTATE_00 en FunGramKB (concepto terminal propuesto)

| | |
|-----------------|--|
| Superordinado: | +SAY_00 |
| MT: | (x1: +HUMAN_00)Theme (x2: +WRITING_00)Referent (x3: +HUMAN_00)Goal |
| PS: | +(e1: +SAY_00 (x1)Theme (x2)Referent (x3)Goal (f1: (e2: +WRITE_00 (x3)Theme (x2)Referent))Reason) |
| Descripción: | to speak or tell somebody something, using words, with the purpose of that person writing it down. |
| Lexicalización: | <i>dictar</i> (español), <i>dictate</i> (inglés) |

Finalmente, dado que la selección léxica estará representada como preferencia conceptual en el MT de \$DICTATE, y dado que el vínculo de la clase léxica con +WRITING_00 es del tipo II, no hay que ubicar los miembros de la clase léxica como colocaciones en la entrada léxica de *dictar*[\$DICTATE].

(b) *dictar* ‘impartir’

El segundo sentido de *dictar* es ‘impartir’. Vinculamos este sentido al evento +TEACH_00, cuya descripción se muestra en la tabla siguiente. Cabe señalar que +TEACH_00 es subordinado de +EXPLAIN_00, que es subordinado de +SAY_00, así que comparte la ruta conceptual de \$DICTATE_00:

Tabla 52. Descripción de +TEACH_00 en FunGramKB

| | |
|-----------------|--|
| Superordinado: | +EXPLAIN_00 |
| MT: | (x1: +HUMAN_00)Theme (x2)Referent (x3: +HUMAN_00 ^ +ANIMAL_00)Goal |
| PS: | +(e1: +EXPLAIN_00 (x1)Theme (x2)Referent (x3)Goal (f1: (e2: +LEARN_00 (x3)Theme (x2)Referent))Purpose) |
| Descripción: | impart skills or knowledge to; "I taught them French"; "He instructed me in building a boat" |
| Lexicalización: | <i>enseñar, entrenar, instruir</i> (español), <i>instruct, teach, train</i> (inglés) |

El PS de +TEACH_00 expresa que típicamente un ‘humano’ (Tema) explica algo no especificado (Referente) a un ‘humano’ o ‘animal’ (Meta), con el propósito de que ese

‘humano’ o ‘animal’ (Meta) aprenda ese algo (Referente). Al considerar los argumentos de la clase léxica con que se combina *dictar* en este sentido, concluimos que es una clase del tipo IV, que no se puede vincular satisfactoriamente a una entidad. En este caso, la dificultad radica en que la clase es mucho más específica que los conceptos en la ontología; la entidad más cercana es +SPEECH_00, pero lo ideal sería contar con un subordinado que capturara la noción de ‘conferencia, charla’:

Tabla 53. Comparación entre clase léxica de REDES y +SPEECH_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB (*no hay vínculo) |
|-----------------------|--|---|
| Superordinado: | --- | +INFORMATION_OBJECT_00 |
| Concepto: | --- | +SPEECH_00 |
| PS: | --- | +(e1: +BE_00 (x1: +SPEECH_00)Theme (x2: +INFORMATION_OBJECT_00)Referent) +(e2: +COMPRISE_00 (x1)Theme (x3: +SOUND_00)Referent) |
| Descripción: | lecciones docentes o disertaciones, o el conjunto de ellas | communication by word of mouth; "his speech was slurred"; "he uttered harsh language"; "he recorded the language of the streets" |
| Manifestación léxica: | <i>clase, conferencia, lección, charla, curso, ciclo</i> | <i>habla</i> (español), <i>speech</i> (inglés) |

En vista de que no hay vínculo conceptual satisfactorio para esta clase léxica, en la plantilla léxica de *dictar*[+TEACH], bajo el segundo argumento, aparecerían las colocaciones *clase, conferencia, lección, charla, curso, ciclo*. Esto quiere decir que la combinación entre *dictar* y cualquiera de esos argumentos estaría vinculada a +TEACH_00.

(c) *dictar* ‘promulgar’

El tercer sentido de *dictar* es ‘promulgar’, que hemos vinculado al evento conceptual +COMMAND_00, cuya descripción en FunGramKB se presenta en la tabla siguiente. El PS de +COMMAND_00 expresa que un ‘humano’ (Tema) le pide algo no especificado (Referente) a un ‘humano’ o ‘animal’ (Meta), que este último tiene que llevar a cabo obligatoriamente:

Tabla 54. Descripción de +COMMAND_00 en FunGramKB

| | |
|----------------|--|
| Superordinado: | +REQUEST_01 |
| MT: | (x1: +HUMAN_00)Theme (x2)Referent (x3: +HUMAN_00 ^ +ANIMAL_00)Goal |
| PS: | +(e1: +REQUEST_01 (x1)Theme (x4: (e2: obl +DO_00 (x3)Theme |

| | |
|------------------------|---|
| | (x2)Referent))Referent (x3)Goal) |
| Descripción: | issue commands or orders for |
| Lexicalización: | <i>mandar, ordenar</i> (español), <i>bid, charge, command, dictate, enjoin, order, prescribe</i> (inglés) |

Dictar asume el valor de +COMMAND_00 en combinación con la clase léxica que presentamos en la tabla a continuación, que hemos vinculado a la entidad +LAW_00. El PS de +LAW expresa que existe una ‘ley’ (Tema) que constituye un grupo (Referente) que se compone de reglas; opcionalmente, la ley es creada por el gobierno y, opcionalmente, los humanos están obligados a obedecerla:

Tabla 55. Comparación entre clase léxica de REDES y +LAW_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|------------------------------|--|---|
| Superordinado: | -- | +GROUP_00 |
| Concepto: | (NORMA) | +LAW_00 |
| PS: | --- | +((e1: +BE_00 (x1: +LAW_00)Theme (x2: +GROUP_00)Referent)(e2: +COMPRISE_00 (x2)Theme (x3: +RULE_00)Referent)) *(e3: +CREATE_00 (x4: +GOVERNMENT_00) Theme (x1) Referent) *(e4: obl +OBEY_00 (x5: +HUMAN_00) Theme (x1)Referent) |
| Descripción: | resoluciones judiciales o administrativas | the collection of rules imposed by authority; "civilization presupposes respect for the law" |
| Manifestación léxica: | <i>ley, decreto, norma condena, auto, sentencia, fallo</i> | <i>derecho, legislación, ley</i> (español), <i>law, legislation</i> (inglés) |

Al mirar los argumentos miembros de la clase léxica, se puede decir que *ley, decreto, norma* lexicalizan a +LAW_00, mientras que los demás –*condena, auto, sentencia, fallo*– son nociones subordinadas. Se trata de una clase del tipo II: una clase léxica que solo puede ser descrita por un concepto superordinado. No hay que especificar los miembros de la clase léxica como colocaciones, porque están representados en la entidad conceptual.

El evento +COMMAND_00 captura este significado de *dictar*, pero no sus preferencias de selección. Por eso, proponemos la creación del subconcepto –COMMAND_LAW, bajo COMMAND_00, que especificaría la preferencia de selección del rol de Referente como +LAW_00 y limitaría la Meta a humanos. Hemos añadido, además de *dictar* y *dictate*, otras lexicalizaciones posibles de este concepto, en ambas lenguas:

Tabla 56. Descripción de –COMMAND_LAW (subconcepto propuesto)

| | |
|-----------------|--|
| Superordinado: | +COMMAND_00 |
| MT: | (x1: +HUMAN_00)Theme (x2: LAW_00)Referent (x3: +HUMAN_00)Goal |
| Lexicalización: | <i>dictar, promulgar, decretar</i> (español), <i>dictate, enact, pass</i> (inglés) |

(d) *dictar* ‘precisar, determinar, sugerir’

El cuarto sentido que identifica REDES para *dictar* es ‘precisar, determinar, sugerir’, y ocurre en combinación con cinco clases léxicas en propiedad. Este es el sentido dominante de *dictar*, en el sentido de que ocurre en combinación con la mayoría de las clases léxicas⁸⁸. Hemos vinculado este sentido al evento +COMMAND_00 también, cuya descripción presentamos arriba. La distinción hecha por REDES de esos dos sentidos para *dictar* (‘promulgar’ y ‘precisar, determinar, sugerir’) se funde en un mismo concepto en FunGramKB, aunque, como veremos, se establecerán varios subconceptos según las diferentes preferencias de selección.

Veamos cada caso en detalle. Vinculamos la primera clase léxica a la entidad +DIRECTION_00. La tabla siguiente presenta una comparación entre la clase léxica y la entidad:

Tabla 57. Comparación entre clase léxica de REDES y +DIRECTION_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|-----------------------|---|---|
| Superordinado: | --- | +ARTIFICIAL_AREA_00 & +NATURAL_AREA_0 |
| Concepto: | DIRECCIÓN | +DIRECTION_00 |
| PS: | --- | +(e1: +BE_00 (x1: +DIRECTION_00)Theme (x2: +ARTIFICIAL_AREA_00 ^ +NATURAL_AREA_00)Referent) *(e2: +MOVE_00 (x3)Agent (x4)Theme (x5)Location (x6)Origin (x1)Goal) |
| Descripción: | inclinación o dirección, generalmente en sentido figurado | the spatial relation between something and the course along which it points or moves; "he checked the direction and velocity of the wind" |
| Manifestación léxica: | <i>dirección, tendencia, rumbo</i> | <i>dirección, rumbo, sentido</i> (español), <i>direction</i> (inglés) |

⁸⁸ Observamos consistentemente en los resultados que los verbos tienen uno o dos patrones de combinatoria dominantes, que se corresponden con uno o dos sentidos sominantes, mientras que los otros son minoritarios. Esto coincide con las orbservaciones de Hanks (2013, §4.3): “It sometimes happens that frequencies are roughly evenly distributed across a number of patterns of a word, but more often one or two patterns are dominant, while other patterns may be quite rare”.

Los argumentos miembros de la clase léxica lexicalizan el concepto +DIRECTION_00; en ese sentido, es una clase estable (tipo I). Cabe señalar, sin embargo, que el concepto +DIRECTION_00 de FunGramKB, que se refiere a la relación espacial entre dos puntos, no es capaz de capturar la valoración figurada del concepto DIRECCIÓN de REDES. Dado eso, sus miembros responden a la entidad conceptual parcialmente (en sus sentidos rectos), pero no en su totalidad (en su sentido figurado). No obstante lo anterior, hemos identificado la clase léxica como Tipo I, y no hay que especificar las colocaciones en el léxico.

A partir del vínculo con la entidad +DIRECTION_00, proponemos la creación del subconcepto -COMMAND_DIRECTION. Este subconcepto establece la preferencia de selección del Referente por +DIRECTION_00 y restringe la de la Meta a +HUMAN_00:

Tabla 58. Descripción de -COMMAND_DIRECTION (subconcepto propuesto)

| | |
|-----------------|--|
| Superordinado: | +COMMAND_00 |
| MT: | (x1: +HUMAN_00)Theme (x2: DIRECTION_00)Referent (x3: +HUMAN_00)Goal |
| Lexicalización: | <i>dictar, determinar, sentar</i> (español), <i>dictate, determine, influence</i> (inglés) |

Las próximas dos clases léxicas de REDES las hemos vinculado a una misma entidad, +PLAN_00. El PS de +PLAN_00 lee como sigue: existe un plan (Tema) que contiene información (Referente), que, opcionalmente, se puede llevar a cabo y de cierta manera. El concepto +PLAN_00 describe satisfactoriamente el descriptor semántico de las clases léxicas, y los argumentos *modelo, patrón, pauta, estrategia y esquema* lo lexicalizan; sin embargo, los argumentos *mecanismo, método y criterio* adquieren ese sentido de manera figurada. La clase pertenece pues al tipo III; es una clase mixta:

Tabla 59. Comparación entre clase léxica de REDES y +PLAN_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|---------------------|--|--|
| Superordinado: | | +INFORMATION_00 |
| Concepto: | MODELO /// RECURSO ⁸⁹ | +PLAN_00 |
| PS: | | +(e1: +BE_00 (x1: +PLAN_00)Theme (x2: +INFORMATION_00)Referent) *(e2: +DO_00 (x3)Theme (x4)Referent (f1: x1)Manner) |
| Descripción: | lo que suele tomarse como referente, prototipo o paradigma de algo /// | a series of steps to be carried out or goals to be accomplished; "they drew |

⁸⁹ Separamos la información de cada clase léxica con ///.

| | | |
|----------------------|---|---|
| | recursos, procedimientos o formas de actuar aplicables a diversos dominios | up a six-step plan"; "they discussed plans for a new bond issue" |
| Manifestación léxica | <i>modelo, patrón, pauta /// estrategia, esquema, mecanismo, método, criterio</i> | <i>estrategia, plan, programa, proyecto</i> (español), <i>plan, programme, project, strategy</i> (inglés) |

La entrada del subconcepto –COMMAND_PLAN especificaría la preferencia de selección del Referente a +PLAN_00:

Tabla 60. Descripción de –COMMAND_PLAN (subconcepto propuesto)

| | |
|-----------------|--|
| Superordinado: | +COMMAND_00 |
| MT: | (x1: +HUMAN_00)Theme (x2: PLAN_00)Referent (x3: +HUMAN_00)Goal |
| Lexicalización: | <i>dictar, sentar</i> (español), <i>dictate</i> (inglés) |

En el lexicón, la plantilla léxica de *dictar*[-COMMAND_PLAN] incluiría las colocaciones que no están representadas por +PLAN: *mecanismo, método, criterio*.

La próxima clase léxica pertenece al tipo IV, pues no se puede vincular a ninguna entidad. Es posible que la noción de CAMBIO propuesta por REDES no pueda ser contenida en una ontología como la de FunGramKB. La entidad más cercana es +OCCURRENCE_00, pero resulta demasiado amplia:

Tabla 61. Comparación entre clase léxica de REDES y +OCCURRENCE_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB (*no hay vínculo) |
|-----------------------|--|---|
| Superordinado: | --- | #PROCESS |
| Concepto: | CAMBIO | +OCCURRENCE_00 |
| PS: | --- | +(e1: +DO_00 (x1)Theme (x2)Referent) |
| Descripción: | alteración o modificación | something that happens |
| Manifestación léxica: | <i>cambio, reforma, regulación, oscilación</i> | <i>acción, acontecimiento, actividad, acto, evento, incidente, lance, proceso, suceso</i> (español), <i>act, action, activity, event, happening, occurrence, process</i> (inglés) |

En este caso, como los argumentos no lexicalizan el concepto, deberían ir como colocaciones en el léxico, en la plantilla léxica de *dictar*[+COMMAND_00].

La última clase léxica bajo este sentido es también del tipo IV: no se puede vincular satisfactoriamente a ninguna entidad. La entidad más cercana es +MUSIC_00, pero este concepto resulta mucho más amplio que la noción que denota la clase léxica, como se percibe en la siguiente tabla comparativa:

Tabla 62. Comparación entre clase léxica de REDES y +MUSIC_00 en FunGramKB

| | | |
|------------------------------|---|---|
| Tipo de información | REDES | FUNGRAMKB (*no hay vínculo) |
| Superodinado: | --- | +SOUND_00 |
| Concepto: | CADENCIA | +MUSIC_00 |
| PS: | --- | +(e1: +BE_00 (x1: +MUSIC_00)Theme (x2: +SOUND_00)Referent) *(e2: +BE_01 (x1)Theme (x3: +PLEASANT_00)Attribute) *(e3: +CREATE_00 (x4: +HUMAN_00)Theme (x1)Referent (f1:+MUSICAL_INSTRUMENT_00)Instrument) |
| Descripción: | pauta rítmica o proporción en el movimiento | the sounds produced by singers or musical instruments (or reproductions of such sounds) |
| Manifestación léxica: | <i>ritmo, compás, velocidad, métrica</i> | <i>música</i> (español), <i>music</i> (inglés) |

Los miembros de la clase léxica *–ritmo, compás, velocidad, métrica–* figurarán como colocaciones en la plantilla léxica de *dictar*[+COMMAND_00].

El sentido dominante de *dictar*, a la luz de su combinatoria léxica, es +COMMAND_00. En la próxima tabla, se muestran las rutas conceptuales de las clases léxicas con que se combina *dictar* en este sentido para examinar si existe alguna tendencia; sin embargo, no es evidente ningún patrón particular en la selección:

Tabla 63. Rutas conceptuales de conceptos vinculados a clases léxicas seleccionadas por *dictar*[+COMMAND_00]

| | | | | | |
|---|--------------|-------------|--------------|---------------|------------|
| Ruta conceptual -----> -----> -----> -----> -----> -----> | | | | | |
| #Physical | #Object | #Collection | +Group | +Law | |
| | | #Region | +Place | +Natural_Area | +Direction |
| #Abstract | #Proposition | +Thought | +Information | +Plan | |

1.1 Resumen *DICTAR*

La próxima tabla muestra el resumen de los vínculos hechos en FunGramKB de los datos de REDES sobre el verbo *dictar*. En la columna de la izquierda se muestra el concepto (básico, terminal o subconcepto) al que se vincula *dictar*, y en la de la derecha, las colocaciones que deben figurar en la plantilla léxica del predicado, en la vinculación con cada evento; recordemos que solo los argumentos que no están representados en las preferencias de selección conceptuales deberán figurar como colocaciones:

Tabla 64. Resumen: *dictar* en FunGramKB (codificación propuesta)

| Vínculo a evento conceptual | Colocaciones en plantilla léxica de <i>dictar</i> |
|-----------------------------|---|
| \$DICTATE_00 | --- |
| +TEACH_00 | <i>clase, conferencia, lección, charla, curso, ciclo</i> |
| +COMMAND_00 | <i>cambio, reforma, regulación, oscilación; ritmo, compás, velocidad, métrica</i> |
| -COMMAND_LAW | --- |
| -COMMAND_DIRECTION | --- |
| -COMMAND_PLAN | <i>mecanismo, método, criterio</i> |

La próxima tabla resume la tipología de las clases (columna 1), los conceptos asignados a cada clase léxica en FunGramKB (columna 2) y en REDES (columna 3). En la columna 4, se muestran los argumentos que pertenecen al paradigma semántico de la clase léxica, y en la columna 5, los que pertenecen a los paradigmas discursivo o pragmático:

Tabla 65. Tipología de clases léxicas que se combinan con *dictar* con respecto a FunGramKB

| Tipo de clase | CONCEPTO en FunGramKB | CONCEPTO en REDES | Argumentos CL - paradigma semántico | Argumentos CL - paradigma discursivo o pragmático |
|---------------|-----------------------|-------------------|---|--|
| I | +DIRECTION_00 | DIRECCIÓN | <i>dirección, tendencia, rumbo</i> | --- |
| II | +LAW_00 | (NORMA) | <i>ley, decreto, condena, auto, sentencia, fallo, norma</i> | --- |
| II | +WRITING_00 | (TEXTO) | <i>carta, texto, informe, párrafo</i> | --- |
| III // III | +PLAN_00 | MODELO // RECURSO | <i>modelo, patrón, pauta // esquema, estrategia</i> | <i>mecanismo método, criterio</i> |
| IV | --- | --- | --- | <i>clase, conferencia, lección, charla, curso, ciclo</i> |
| IV | --- | CAMBIO | --- | <i>cambio, reforma, regulación, oscilación</i> |
| IV | --- | CADENCIA | --- | <i>cambio, reforma, regulación, oscilación</i> |

2. PERDER, DISIPAR, DERROCHAR, MALOGRAR, MALGASTAR, DILAPIDAR, DESAPROVECHAR, DESPILFARRAR, DESPERDICAR

En esta sección, discutiremos nueve verbos relacionados semánticamente. Todos comparten, en al menos uno de sus sentidos, el valor de ‘malgastar’, que hemos vinculado al evento +WASTE_00: *derrochar, desaprovechar, desperdiciar, despilfarrar, dilapidar, disipar, malgastar, malograr y perder*. Primero discutiremos cada verbo individualmente, tomando en cuenta los sentidos y la combinatoria léxica de cada uno. Comenzaremos con los verbos polisémicos –*perder, disipar, derrochar, malograr*–, y luego veremos los monosémicos –*malgastar, desaprovechar, dilapidar, despilfarrar, desperdiciar*–. En la próxima sección, analizaremos los nueve verbos colectivamente, para examinar si comparten patrones en su selección léxica.

2.1 PERDER

REDES identifica cuatro sentidos para el verbo *perder*: (a) ‘no alcanzar a tiempo’; (b) ‘desaprovechar’; (c) ‘resultar vencido en’; (d) ‘quedar desprovisto de’. Existe una correspondencia clara entre cada uno de estos sentidos de *perder* y las clases léxicas con que se combina el verbo. Los sentidos y la combinatoria léxica de *perder*, según los presenta REDES, se transcriben en la tabla siguiente:

Tabla 66. Sentidos y combinatoria léxica de *perder*, en REDES

| Sentido del verbo en REDES | CONCEPTO asignado a la CL en REDES | Descriptor de la CL en REDES | Argumentos miembros de la CL en REDES |
|----------------------------|------------------------------------|--|--|
| 'no alcanzar a tiempo' | --- | medios de transporte | <i>avión, tren, autobús, barco</i> |
| | --- | sus servicios | <i>transbordo, vuelo</i> |
| 'desaprovechar' | (TIEMPO) | sustantivos temporales | <i>día, mes, año, tiempo</i> |
| | --- | circunstancias favorables | <i>ocasión, oportunidad, posibilidad</i> |
| 'resultar vencido en' | CONFRONTACIÓN | juicio, y otros procesos de naturaleza jurídica | <i>pleito, juicio, apelación</i> |
| | --- | confrontaciones, a menudo deportivas, así como algunos de sus lances | <i>batalla, lucha, partido, campeonato, encuentro, combate, guerra, pelea, final, set</i> |
| 'quedar desprovisto de' | (CUALIDAD) | propiedades de las personas y las cosas | <i>color, brillantez</i> |
| | --- | materias y sustancias | <i>gasolina, pelo</i> |
| | (CAPACIDAD) | capacidades | <i>vista, oído, olfato</i> |
| | --- | deseos o necesidades | <i>el apetito, el gusto, el sueño</i> |
| | --- | cualquier objeto físico | <i>paraguas, pendientes, libro, cartera, foto</i> |
| | CURSO | rumbo o curso, frec. en sentido figurado | <i>rumbo, norte, camino, dirección</i> |
| | CUALIDAD | rasgos y propiedades fundamentales del individuo | <i>memoria, equilibrio, personalidad, identidad, vida, compostura, lucidez</i> |
| | INCLINACIÓN | energía, deseo, impulso | <i>deseo, interés, gana, ilusión, fuerza, energía, impulso, esperanza, curiosidad, aliento</i> |
| | SENTIMIENTO | diversas sensaciones y sentimientos | <i>miedo, temor, vergüenza, paciencia, nervios, alegría, optimismo, cariño, amor</i> |
| | PODER | poder, control o capacidad para ejercitarlos | <i>poder, control, autoridad, dominio, hegemonía, soberanía</i> |
| COSTUMBRE | costumbre | <i>costumbre, hábito</i> | |

Examinamos a continuación los cuatro sentidos de *perder*, junto con su combinatoria léxica, y su representación en el marco de la base de conocimiento FunGramKB.

(a) *perder* ‘no alcanzar a tiempo’

Para significar ‘no llegar a tiempo’, *perder* se combina con dos clases léxicas: (a) ‘medios de transporte’, *avión, tren, autobús, barco*, y (b) ‘sus servicios’, *transbordo, vuelo*. Estas dos clases léxicas las hemos vinculado a +VEHICLE_00 y a +TRAVEL_00, respectivamente; las próximas dos tablas presentan una comparación entre la clase léxica y el concepto. En ambos casos, se trata de clases descritas por un concepto superordinado (tipo II). Los miembros de la clase léxica que identificamos con +VEHICLE_00 se corresponden con conceptos subordinados (*tren*[+TRAIN_00], *autobús*[+BUS_00], etc). Los de la clase que vinculamos con +TRAVEL_00 también son nociones subordinadas a ese concepto:

Tabla 67. Comparación entre clase léxica de REDES y +VEHICLE_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|------------------------------------|--|
| Superordinado: | --- | +TRANSPORT_00 |
| Concepto: | --- | +VEHICLE_00 |
| PS: | --- | +(e1: +BE_00 (x1: +TRANSPORT_00)Theme (x2: +ARTIFICIAL_OBJECT_00 & +CORPUSCULAR_00 & +SOLID_00)Referent) |
| Descripción: | medios de transporte | conveyance that transports people or objects |
| Manifestación léxica | <i>avión, tren, autobús, barco</i> | (español), <i>vehicle</i> (inglés) |

Tabla 68. Comparación entre clase léxica de REDES y +TRAVEL_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|--------------------------|--|
| Superordinado: | --- | +MOTION_00 |
| Concepto: | --- | +TRAVEL_00 |
| PS: | --- | +(e1: +BE_00 (x1: +TRAVEL_00)Theme (x2: +MOTION_00)Referent) +(e2: +TRAVEL_01 (x3)Agent (x4)Theme (x5)Location (x6)Origin (x7)Goal (f1: x1)Scene) |
| Descripción: | sus servicios | the act of going from one place to another; "he enjoyed selling but he hated the travel" |
| Manifestación léxica | <i>transbordo, vuelo</i> | <i>viaje</i> (español), <i>journey, travel, travelling, trip</i> (inglés) |

El sentido de ‘no llegar a tiempo’ no está cabalmente representado por ningún concepto en FunGramKB, por lo que proponemos que se añada, bajo +FAIL_00, un concepto terminal llamado \$MISS_00. Lo ubicamos bajo +FAIL_00, porque al mirar este sentido de *perder* (‘no llegar a tiempo’) en los diccionarios, los hiperónimos expresan que ‘no se logra hacer algo’ e incluyen ‘fracasar’; además, en los diccionarios en inglés, el hiperónimo más frecuente de *miss*, en ese sentido, es ‘fail’.

La propuesta establecería las siguientes preferencias de selección para el concepto terminal \$MISS_00, que se ubicaría bajo +FAIL_00: en el MT, el Tema se especifica como +HUMAN_00; el Referente se especifica como +VEHICLE_00 o +TRAVEL_00. En el PS, se añade un parámetro que establece que el Tema no llega al Referente a tiempo. Las dos tablas siguientes presentan el concepto básico +FAIL_00 y el concepto terminal propuesto \$MISS_00:

Tabla 69. Descripción de +FAIL_00 en FunGramKB

| | |
|-----------------|---|
| Superordinado: | +TRY_00 |
| MT: | (x1: +HUMAN_00 ^ +ANIMAL_00)Theme (x2)Referent |
| PS: | +(e1: +TRY_00 (x1)Theme (x2)Referent (f1: (e2: n +DO_00 (x1)Theme (x2)Referent))Result) |
| Descripción | be unsuccessful; "Where do today's public schools fail?"; "The attempt to rescue the hostages failed miserably" |
| Lexicalización: | <i>fallar fracasar, malograr</i> (español), <i>fail</i> (inglés) |

Tabla 70. Descripción de \$MISS_00 (concepto terminal propuesto)

| | |
|-----------------|--|
| Superordinado: | +FAIL_00 |
| MT: | (x1: +HUMAN_00)Theme (x2: +VEHICLE_00 ^ +TRAVEL_00)Referent |
| PS: | +(e1: +FAIL (x1)Theme (x2)Referent (f1: (e2: +ARRIVE_00 (x3)Agent (x1)Theme (x4)Location (x5)Origin (x6)Goal (f2: +LATE_00)Time))Result) |
| Descripción | fail to arrive on time: I missed the bus, You missed the beginning of the movie |
| Lexicalización: | <i>perder</i> (español), <i>miss</i> (inglés) |

Finalmente, en el nivel léxico, en la plantilla léxica de *perder*[\$MISS_00], en el apartado de colocaciones, no hay que detallar los argumentos con los que este verbo adquiere este sentido, dado que son clases léxicas del tipo II y la preferencia de selección en el nivel conceptual los representa.

(b) *perder* ‘desaprovechar’

El segundo sentido que *perder* asume es ‘desaprovechar’, que vinculamos a +WASTE_00. El PS de +WASTE expresa que una entidad típicamente humana (Tema) usa algo (Referente) de manera ineficiente o inapropiada. El Tema (‘entidad que lleva a cabo una acción, con o sin volición’) está especificado como +HUMAN_00, mientras que el Referente (‘entidad que está directamente involucrada en el evento causado por otra entidad’) no tiene preferencia de selección.

Tabla 71. Descripción de +WASTE_00 en FunGramKB

| | |
|------------------------|--|
| Hiperónimo: | +USE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2)Referent |
| PS: | +(e1: +USE_00 (x1)Theme (x2)Referent (f1: +CARELESS_00)Manner) |
| Descripción: | use inefficiently or inappropriately; "waste heat" |
| Lexicalización: | <i>malgastar</i> (español), <i>waste</i> (inglés) |

El verbo *perder* asume este sentido en combinación con dos clases léxicas. La primera la hemos identificado con la entidad +PERIOD_00. Esta clase es del tipo II, pues está descrita por un concepto superordinado. Tres de los cuatro miembros se corresponden con conceptos subordinados de +PERIOD_00 (*day*[+DAY_00], *mes*[+MONTH_00], *año*[+YEAR_00]), y el cuarto (*tiempo*) lexicaliza a +PERIOD_00. Se representa la clase léxica con el concepto superordinado; no es necesario, pues, ubicar ningún argumento como colocación en la plantilla léxica:

Tabla 72. Comparación entre clase léxica de REDES y +PERIOD_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------------|------------------------------|---|
| Superordinado: | --- | +TIME_00 |
| Concepto: | --- | +PERIOD_00 |
| PS: | --- | +(e1: +BE_00 (x1: +PERIOD_00)Theme (x2: +TIME_00)Referent) |
| Descripción: | sustantivos temporales | a particular length of time with a beginning and an end |
| Manifestación léxica | <i>día, mes, año, tiempo</i> | <i>edad, época, era, intervalo, periodo, tiempo</i> (español), <i>age, epoch, era, interval, period, span, stretch, term, time</i> (inglés) |

A partir de la selección léxica de *perder* en este sentido, proponemos el subconcepto – WASTE_PERIOD:

Tabla 73. Descripción de -WASTE_PERIOD (subconcepto propuesto)

| | |
|---------------------|---|
| Hiperónimo: | +USE_00 |
| Marco temático (MT) | (x1: +HUMAN_00)Theme (x2: +PERIOD_00)Referent |
| Descripción | use time inefficiently or inappropriately |

La segunda clase con que se combina *perder*[+WASTE_00], que REDES describe como ‘circunstancias favorables’, tiene los miembros *oportunidad, posibilidad*. Esta clase es del tipo IV: no se puede vincular a ninguna entidad en FunGramKB. La noción de ‘posibilidad’ no tiene representación en FunGramKB, tal vez porque cualquier cosa puede ser una posibilidad en un contexto comunicativo o pragmático dado. En este caso, los argumentos deben ubicarse como colocaciones en la entrada léxica de *perder*[+WASTE_00].

(c) *perder* ‘resultar vencido en’

Hemos vinculado el tercer sentido del verbo *perder*, ‘resultar vencido en’, con el evento +FAIL_00, cuya descripción en FunGramKB repetimos en la tabla siguiente. El PS de +FAIL_00 expresa que un humano o animal (Tema) intenta hacer algo no especificado (Referente), y el resultado es que ese humano o animal no logra hacer ese algo. El verbo *perder* asume este sentido en combinación con dos clases léxicas, que hemos desdoblado en tres entidades en FunGramKB: +TRIAL_00, +BATTLE_00 y +COMPETITION_00:

Tabla 74. Descripción de +FAIL_00 en FunGramKB

| | |
|-----------------|---|
| Superordinado: | +TRY_00 |
| MT: | (x1: +HUMAN_00 ^ +ANIMAL_00)Theme (x2)Referent |
| PS: | +(e1: +TRY_00 (x1)Theme (x2)Referent (f1: (e2: n +DO_00 (x1)Theme (x2)Referent))Result) |
| Descripción | be unsuccessful; "Where do today's public schools fail?"; "The attempt to rescue the hostages failed miserably" |
| Lexicalización: | <i>fallar fracasar, malograr</i> (español), <i>fail</i> (inglés) |

La próxima tabla compara la información de la primera clase léxica con la descripción del concepto en FunGramKB. Esta clase pertenece al tipo I: es una clase léxica estable con respecto a FunGramKB. Todos los miembros de la clase pueden lexicalizar el concepto, y el concepto puede sustituir la clase:

Tabla 75. Comparación entre clase léxica de REDES y +TRIAL_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|---|--|
| Superodinado: | | +OCCURRENCE_00 |
| Concepto: | CONFRONTACIÓN | +TRIAL_00 |
| PS: | | PS: +(e1: +BE_00 (x1: +TRIAL_00)Theme (x2: +OCCURRENCE_00)Referent) |
| Descripción: | juicio, y otros procesos de naturaleza jurídica | a formal legal process in which a judge and jury decide whether someone is guilty of a particular crime by questioning them and considering the evidence |
| Manifestación léxica | <i>pleito, juicio, apelación</i> | <i>juicio, proceso</i> (español), <i>trial</i> (inglés) |

La segunda clase léxica se divide en dos conceptos: +BATTLE_00 y +COMPETITION_00. La próxima tabla muestra la información que ofrece REDES sobre la clase léxica:

Tabla 76. Descripción de clase léxica en REDES que se combina con *perder* ‘resultar vencido’

| Tipo de información | Datos de REDES |
|------------------------|---|
| Concepto: | CONFRONTACIÓN |
| Descriptor: | confrontaciones, a menudo deportivas, así como algunos de sus lances |
| Miembros clase léxica: | <i>batalla, lucha, partido, campeonato, encuentro, combate, guerra, pelea, final, set</i> |

Las próximas dos tablas presentan la descripción conceptual de las entidades +BATTLE_00 y +COMPETITION_00. Los argumentos *batalla, lucha, combate, guerra, pelea* lexicalizan a +BATTLE_00 y *partido, campeonato, encuentro, final, set* lexicalizan a +COMPETITION_00. Ambas son clases estables (tipo I):

Tabla 77. Descripción de +BATTLE_00 en FunGramKB

| | |
|-----------------|---|
| Superordinado: | +VIOLENCE_00 |
| Concepto: | + BATTLE_00 |
| PS: | +(e1: +BE_00 (x1: +BATTLE_00)Theme (x2: +OCCURRENCE_00)Referent) *(e2: +FIGHT_00 (x3: +ARMY_00)Theme (x4: +ARMY_00)Referent (f1: +LAND_00 ^ +SEA_00 ^ +SKY_00)Location (f2: +WEAPON_00)Instrument(f3: (e3: pos +DIE_00 (x5: m +HUMAN_00)Theme))Result (f4: x1)Scene) |
| Descripción: | a hostile meeting of opposing military forces in the course of a war; "Grant won a decisive victory in the battle of Chickamauga" |
| Lexicalización: | <i>batalla, combate, contienda, hostilidades, lucha, ofensiva</i> (español), <i>battle, combat, engagement, fight, hostilities, squirmish</i> (inglés) |

Tabla 78. Descripción de +COMPETITION_00 en FunGramKB

| | |
|-----------------|---|
| Superordinado | +OCCURRENCE_00 |
| Concepto: | +COMPETITION_00 |
| PS: | +(e1: +BE_00 (x1: +COMPETITION_00)Theme (x2: +OCCURRENCE_00)Referent) *(e2: +COMPETE_00 (x3: +HUMAN_00 ^ +TEAM_00)Theme (x4: +HUMAN_00 ^ +TEAM_00)Referent (f1: x1)Scene (f2: (e3: +WIN_00 (x3)Theme (x5: +PRIZE_00)Referent))Purpose) |
| Descripción: | an occasion on which a winner is selected from among two or more contestants |
| Lexicalización: | <i>certamen, competición, concurso, juego, mano a mano, partido, torneo</i> (español), <i>competition, contest, game, head-to-head, match, tournament</i> (inglés) |

Para codificar este sentido en FunGramKB, proponemos la creación de los subconceptos -FAIL_TRIAL, -FAIL_BATTLE y -FAIL_COMPETITION, que se presentan en las siguientes tres tablas. Como las restricciones de selección de *perder* en el sentido de +FAIL_00 son clases léxicas del tipo I, no es necesario añadir ningún argumento en el espacio de las colocaciones en las plantillas léxicas de los subconceptos: *perder*[-FAIL_TRIAL], *perder*[-FAIL_BATTLE] o *perder* [-FAIL_COMPETITION].

Tabla 79. Descripción de -FAIL_TRIAL (subconcepto propuesto)

| | |
|----------------|---|
| Superordinado: | +TRY_00 |
| MT: | (x1: +HUMAN_00)Theme (x2:TRIAL)Referent |
| Descripción | be unsuccessful at a trial |

Tabla 80. Descripción de -FAIL_BATTLE (subconcepto propuesto)

| | |
|----------------|--|
| Superordinado: | +TRY_00 |
| MT: | (x1: +HUMAN_00)Theme (x2:BATTLE)Referent |
| Descripción | be unsuccessful at a battle |

Tabla 81. Descripción de -FAIL_COMPETITION (subconcepto propuesto)

| | |
|----------------|---|
| Superordinado: | +TRY_00 |
| MT: | (x1: +HUMAN_00)Theme (x2:COMPETITION)Referent |
| Descripción | be unsuccessful at a competition |

(d) *perder* ‘quedar desprovisto de’

El cuarto sentido de *perder* según REDES, ‘quedar desprovisto de’, lo desdoblamos en dos eventos conceptuales que son superordinado y subordinado: +LACK_00 y +LOSE_00, respectivamente. Aunque REDES agrupa estas clases léxicas bajo la noción de ‘quedar desprovisto de’, la ontología de eventos de FunGramKB nos lleva a separarlas, porque el Referente de +LACK_00 no está especificado, y provee el espacio para nociones abstractas, mientras que el de +LOSE_00 especifica el Referente como un objeto material y tiene un parámetro muy específico.

La descripción de +LOSE_00 en FunGramKB se presenta en la tabla siguiente; el PS de +LOSE_00 expresa que un humano o animal (Tema) carece de un objeto corpuscular (Referente), a causa de que lo colocó en un lugar y no recuerda dónde:

Tabla 82. Descripción de +LOSE_00 en FunGramKB

| | |
|------------------------|---|
| Superordinado: | +LACK_00 |
| Concepto: | +LOSE_00 |
| MT: | (x1: +HUMAN_00^+ANIMAL_00)Theme (x2: +CORPUSCULAR_00)Referent |
| PS: | +(e1: +LACK_00 (x1)Theme (x2)Referent (f1: (e2: past +PUT_00 (x1)Agent (x2)Theme (x3)Origin (x4)Goal))Reason (f2: (e3: n +REMEMBER_00 (x1)Agent (x1)Theme (x4)Referent))Reason) |
| Descripción: | place (something) where one cannot find it again; "I misplaced my eyeglasses" |
| Lexicalización: | <i>extraer, perder, traspapelar</i> (español), <i>lose, mislay, misplace</i> (inglés) |

La clase léxica que activa este sentido de *perder* se vincula a +SOLID_00. Aquí se incluye cualquier objeto físico, por lo que claramente se trata de una clase del tipo II, descrita por un concepto superordinado:

Tabla 83. Comparación entre clase léxica de REDES y +SOLID_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|-----------------------------|---|--|
| Superordinado: | --- | +CORPUSCULAR_00 ^ +SUBSTANCE_00 |
| Concepto: | --- | +SOLID_00 |
| PS: | --- | +(e1: +BE_00 (x1: +SOLID_00)Theme (x2: +SUBSTANCE_00 ^ +CORPUSCULAR_00)Referent) |
| Descripción: | cualquier objeto físico | a substance that is a solid at room temperature and pressure |
| Manifestación léxica | <i>paraguas, pendientes, libro, cartera, foto</i> | (español), <i>solid</i> (inglés) |

Para codificar este sentido en FunGramKB, vinculamos *perder* a +LOSE_00. Nótese que +SOLID_00 es subordinado de +CORPUSCULAR, la preferencia se selección de +LOSE_00. Por lo tanto, no tenemos que añadir los miembros de la clase léxica al espacio de colocaciones en la plantilla léxica de *perder*[+LOSE_00], porque la preferencia de selección +CORPUSCULAR_00 ya las incluye. Como es una clase del tipo II, no hay que detallar las colocaciones en la entrada léxica de *perder*[+LOSE_00].

La descripción de +LACK_00 en FunGramKB se presenta en la tabla siguiente. El PS de +LACK_00 expresa que un Tema no especificado no tiene un Referente no especificado en un lugar no especificado; el Tema necesita ese Referente:

Tabla 84. Descripción de +LACK_00 en FunGramKB

| | |
|-----------------------|--|
| Superordinado: | +HAVE_00 |
| Concepto: | +LACK_00 |
| | (x1)Theme (x2)Referent |
| PS: | +(e1: n +HAVE_00 (x1)Theme (x2)Referent (f1)Location) +(e2: +NEED_00 (x1)Theme (x2)Referent) |
| Descripción | be without; "This soup lacks salt"; "There is something missing in my jewellery box!" |
| Lexicalización | <i>carecer, escasear, faltar, necesitar, precisar, requerir</i> (español), <i>to be defficient in, be short of, be without, lack, miss, need, require, want</i> (inglés) |

El vínculo de *perder* con +LACK_00 se da con las nueve clases léxicas que presentamos a continuación. Este es el sentido dominante de *perder*: la vasta mayoría de las clases léxicas con que se combina *perder* activan este sentido.

La clase léxica vinculada a +PHYSICAL_ATTRIBUTE_00 pertenece al tipo II: se vincula con un concepto superordinado, pues sus miembros se corresponden con conceptos subordinados (por ejemplo, *color*[+COLOUR_0]0) o que podrían serlo:

Tabla 85. Comparación entre clase léxica de REDES y +PHYSICAL_ATT_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|------------------------------|---|--|
| Superordinado: | --- | +ATTRIBUTE_00 |
| Concepto: | (CUALIDAD) | +PHYSICAL_ATT_00 |
| PS: | --- | +(e1: +BE_00 (x1: PHYSICAL_ATT_00)Theme (x2: +ATTRIBUTE_00)Referent) *(e2: +BE_01 (x3: +CORPUSCULAR_00 ^ +SUBSTANCE_00)Theme (x1)Attribute) |
| Descripción: | propiedades de las personas y las cosas | physical attribute |
| Manifestación léxica: | <i>color, brillantez</i> | --- |

El concepto superordinado +PHYSICAL_ATT_00 es bastante amplio; bajo él, se organizan conceptos que se corresponden con atributos físicos permanentes, más difíciles de combinar con *perder*, como +WIDTH_00 o +LENGHT_00. Sin embargo, en principio, cualquier atributo físico se puede *perder*, de manera que consideramos que este concepto puede representar la clase léxica adecuadamente.

Las próximas dos clases léxicas se vinculan a la entidad +SENSE_00. Estas clases se describen con un concepto superordinado (tipo II), porque sus miembros se corresponderían con conceptos subordinados a la noción de ‘facultad a través de la que se percibe el mundo exterior’.

Tabla 86. Comparación entre clase léxica de REDES y +SENSE_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|---|---|
| Superordinado: | --- | +ABILITY_00 |
| Concepto: | --- | +SENSE_00 |
| PS: | --- | +(e1: +BE_00 (x1: +SENSE_00)Theme (x2: +ABILITY_00)Referent) *(e2: +PERCEIVE_00 (x3)Theme (x4)Referent (f1: x1)Means) |
| Descripción: | deseos o necesidades /// capacidades | the faculty through which the external world is perceived |
| Manifestación léxica | <i>apetito, gusto, sueño</i> /// <i>vista, oído, olfato</i> | <i>sentido</i> (español), <i>sense</i> (inglés) |

La próxima clase se vincula con la entidad +DESIRE_00. Esta clase es del tipo III: una clase mixta en la que algunos miembros pueden lexicalizar el concepto +DESIRE_00 (*deseo, interés, gana, ilusión*) y los demás (*fuerza, energía, impulso, esperanza, curiosidad, aliento*) adquieren ese sentido figuradamente, en función de la combinación con *perder*[+LACK_00]:

Tabla 87. Comparación entre clase léxica de REDES y +DESIRE_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|--|---|
| Superordinado: | --- | +FEELING_00 |
| Concepto: | INCLINACIÓN | +DESIRE_00 |
| PS: | --- | +(e1: +BE_00 (x1: +DESIRE_00)Theme (x2: +FEELING_00)Referent) *(e2: +FEEL_00 (x3: (e3: +DESIRE_01 (x4: +HUMAN_00)Theme (x5)Referent))Agent (x4)Theme (x1)Attribute) |
| Descripción: | energía, deseo, impulso | a strong hope or wish |
| Manifestación léxica | <i>deseo, interés, gana, ilusión, fuerza, energía, impulso, esperanza, curiosidad, aliento</i> | <i>deseo</i> (español), <i>desire, wish</i> (inglés) |

En este caso, incluimos los argumentos que adquieren el valor figuradamente como colocaciones en la plantilla léxica: *fuerza, energía, impulso, esperanza, curiosidad, aliento*.

La próxima clase se vincula con la entidad +COGNITIVE_ATTRIBUTE. En la tabla siguiente se presenta la comparación entre la clase léxica y el concepto. Se trata de una clase mixta (tipo III), que tiene miembros que pueden ser subordinados del concepto (como *memoria*[+MEMORY_00] y *lucidez*[+INTELLIGENCE_00]), pero los demás (*equilibrio*, *personalidad*, *identidad*, *vida*, *compostura*) no son lexicalizaciones exclusivas de +COGNITIVE_ATTRIBUTE. Estos argumentos se ubicarán entonces en la plantilla léxica, como colocaciones de *perder*[+LACK_00].

Tabla 88. Comparación entre clase léxica de REDES y +COGNITIVE_ATT en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|--|--|
| Superordinado: | --- | +ABILITY_00 |
| Concepto: | CUALIDAD | +COGNITIVE_ATT |
| PS: | --- | +(e1: +BE_00 (x1: +COGNITIVE_ATT_00)Theme (x2: +ABILITY_00)Referent) +(e2: +BE_01 (x3: HUMAN_00)Theme (x1)Attribute) |
| Descripción: | rasgos y propiedades fundamentales del individuo | cognitive attribute |
| Manifestación léxica | <i>memoria</i> , <i>equilibrio</i> , <i>personalidad</i> , <i>identidad</i> , <i>vida</i> , <i>compostura</i> , <i>lucidez</i> | --- |

La próxima clase léxica se vincula con +FEELING_00. Este es un caso del tipo II, una clase descrita por un concepto superordinado. Muchos de sus miembros se corresponden con conceptos subordinados a +FEELING_00: *miedo*[+FEAR_00], *temor*[+FEAR_00], *vergüenza*[+SHAME_00], *nervios*[+ANXIETY_00], *alegría*[+PLEASURE_00], *optimismo*[+PLEASURE_00], *cariño*[+PLEASURE_00], *amor*[+PLEASURE_00]:

Tabla 89. Comparación entre clase léxica de REDES y +FEELING_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|--|---|
| Superodinado: | --- | +PSYCHOLOGICAL_ATT_00 |
| Concepto: | SENTIMIENTO | +FEELING_00 |
| PS: | --- | +(e1: +BE_00 (x1: +FEELING_00)Theme (x2: +PSYCHOLOGICAL_ATT_00)Referent) +(e2: +FEEL_00 (x3)Agent (x4: +HUMAN_00)Theme (x1)Attribute) |
| Descripción: | diversas sensaciones y sentimientos | an emotion that you feel; emotional attribute |
| Manifestación léxica | <i>miedo, temor, vergüenza, paciencia, nervios, alegría, optimismo, cariño, amor</i> | <i>ánimo, emoción, sentimiento</i> (español), <i>emotion, feeling, humour, mood, sentiment, spirit, temper</i> (inglés) |

La próxima clase léxica pertenece al tipo IV: no se puede vincular satisfactoriamente a ninguna entidad. El concepto PODER de REDES no tiene un reflejo paralelo en FunGramKB, tal vez porque es una noción demasiado amplia, que puede referirse a un sinnúmero de cosas. La entidad más cercana sería tal vez +BEHAVIOUR_00, de la cual la clase léxica podría ser subordinada, junto con +COURAGE_00, +EVIL_00 y +HUMOUR_00.

Tabla 90. Comparación entre clase léxica de REDES y +BEHAVIOUR_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB (*no hay vínculo) |
|-----------------------|---|--|
| Superodinado: | --- | +PSYCHOLOGICAL_ATT_00 |
| Concepto: | PODER | +BEHAVIOUR_00 |
| PS: | --- | +(e1: +BE_00 (x1: +BEHAVIOUR_00)Theme (x2: +PSYCHOLOGICAL_ATT_00)Referent) +(e2: +DO_00 (x3: +HUMAN_00 ^ +ANIMAL_00)Theme (x4)Referent (f1: x1)Manner) |
| Descripción: | poder, control o capacidad para ejercitarlos | behavioural attribute |
| Manifestación léxica: | <i>poder, control, autoridad, dominio, hegemonía, soberanía</i> | <i>actitud, comportamiento, conducta, proceder</i> (español), <i>attitude, behaviour, comportment, conduct, demeanour</i> (inglés) |

El verbo *perder* también se vincula con el evento +LACK_00 en combinación con una clase que vinculamos con la entidad +CUSTOM_00. Esta clase es estable (tipo I) en relación con el concepto en FunGramKB. Los dos argumentos que identifica REDES pueden lexicalizar el concepto +CUSTOM_00, y la descripción de +CUSTOM_00 describe bien la clase léxica:

Tabla 91. Comparación entre clase léxica de REDES y +CUSTOM_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|--------------------------|---|
| Superordinado: | --- | +OCCURRENCE_00 |
| Concepto: | COSTUMBRE | +CUSTOM_00 |
| PS: | --- | +(e1: +BE_00 (x1: +CUSTOM_00)Theme (x2: +OCCURRENCE_00)Referent) +((e2: pres +DO_00 (x3: +HUMAN_00)Theme (x4)Referent (f1: x1)Scene)(e3: past +DO_00 (x5: +HUMAN_00)Theme (x4)Referent (f2: +MUCH_00)Frequency)) |
| Descripción: | costumbre | a specific practice of long standing |
| Manifestación léxica | <i>costumbre, hábito</i> | <i>costumbre, hábito, tradición, usanza, uso</i> (español), <i>custom, habit, tradition</i> (inglés) |

El verbo *perder*[+LACK_00] también se combina con dos clases léxicas que designan objetos materiales. La primera clase léxica es muy general: se trata de ‘materias y sustancias’ y la hemos vinculado con +SUBSTANCE_00. Esta clase pertenece al tipo II, por ser una clase descrita por un concepto superordinado: *gasolina* se vincula a +PETROL_00 y *pelo* a +SOLID_00, dos subordinados de +SUBSTANCE_00:

Tabla 92. Comparación entre clase léxica de REDES y +SUBSTANCE_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|-----------------------|---|
| Superordinado: | --- | +ARTIFICIAL_OBJECT_00 ^ +NATURAL_OBJECT_00 |
| Concepto: | --- | +SUBSTANCE_00 |
| PS: | --- | +(e1: +BE_00 (x1: +SUBSTANCE_00)Theme (x2: +ARTIFICIAL_OBJECT_00 ^ +NATURAL_OBJECT_00)Referent) |
| Descripción: | materias y sustancias | --- |
| Manifestación léxica | <i>gasolina, pelo</i> | <i>matter, substance</i> (inglés) |

La segunda clase léxica que tiene una ruta conceptual material la vinculamos con +DIRECTION_00. Aunque el concepto describe bien los sentidos rectos de los miembros de la clase léxica, no captura tan bien sus sentidos figurados. Sin embargo, la consideramos una clase estable (tipo I), porque los argumentos sí lexicalizan el concepto en su sentido recto:

Tabla 93. Comparación entre clase léxica de REDES y +DIRECTION_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|---|--|
| Superodinado: | --- | +NATURAL_AREA_00 |
| Concepto: | CURSO | +DIRECTION_00 |
| PS: | --- | PS: +(e1: +BE_00 (x1: +DIRECTION_00)Theme (x2: +ARTIFICIAL_AREA_00 ^ +NATURAL_AREA_00)Referent) *(e2: +MOVE_00 (x3)Agent (x4)Theme (x5)Location (x6)Origin (x1)Goal) |
| Descripción: | rumbo o curso, frecuentemente en sentido figurado | the spatial relation between something and the course along which it points or moves; "he checked the direction and velocity of the wind" |
| Manifestación léxica | <i>camino, dirección, rumbo, norte</i> | <i>dirección, rumbo, sentido</i> (español), <i>direction</i> (inglés) |

Habiendo repasado todas las selecciones léxicas que hace *perder* en su sentido de +LACK_00, en la próxima tabla presentamos el resumen de las rutas conceptuales de las clases léxicas con que se combina. Queda claro que hay una preferencia de selección por los atributos, especialmente los atributos psicológicos, en la selección sistemática de *perder*[+LACK_00]:

Tabla 94. Rutas conceptuales de conceptos vinculados a clases léxicas que se combinan con *perder*[+LACK_00]

| Rutas conceptuales -----> -----> -----> -----> -----> | | | | | |
|---|------------|------------------------|--------------------|---------------|-----------------------|
| #Abstract | #Attribute | +Attribute | +Physical_Att | | |
| | | | +Psychological-Att | +Ability | +Sense |
| | | | | +Ability | +Cognitive_Att |
| | | | | +Feeling | |
| | | | | +Feeling | +Desire |
| #Physical | #Object | #Self_connected_object | +Natural_object | +Corpuscular | +Substance |
| | | #Region | +Place | +Natural_Area | +Direction |
| | #Process | +Occurrence | +Custom | | |

Ante el hecho de que +ABILITY_00 y +FEELING_00 son subordinados directos de PSYCHOLOGICAL_ATT, proponemos la creación del subconcepto – LACK_PSYCHOLOGICAL_ATT. Esta es una preferencia de selección que engloba las ‘habilidades’ y ‘sentimientos’. A este subconcepto, vamos a asignar las colocaciones de la clase léxica *poder, control, autoridad, dominio, hegemonía, soberanía*, que no se pudo vincular a ninguna entidad satisfactoriamente.

Tabla 95. Descripción de -LACK_PSYCHOLOGICAL_ATT (concepto propuesto)

| | |
|----------------|---|
| Superordinado: | +HAVE_00 |
| Concepto: | +LACK_00 |
| MT: | (x1)Theme (x2: PSYCHOLOGICAL_ATT)Referent |
| Descripción | be without; “she has lost her hearing” |

A partir de las otras selecciones léxicas, proponemos la creación de – LACK_SUBSTANCE, –LACK_DIRECTION y –LACK_CUSTOM. Las tablas siguientes presentan estos subconceptos propuestos:

Tabla 96. Descripción de -LACK_SUBSTANCE (concepto propuesto)

| | |
|----------------|-----------------------------------|
| Superordinado: | +HAVE_00 |
| Concepto: | +LACK_00 |
| MT: | (x1)Theme (x2: SUBSTANCE)Referent |
| Descripción | be without; “she is losing hair” |

Tabla 97. Descripción de -LACK_DIRECTION (concepto propuesto)

| | |
|----------------|------------------------------------|
| Superordinado: | +HAVE_00 |
| Concepto: | +LACK_00 |
| MT: | (x1)Theme (x2: DIRECTION)Referent |
| Descripción | be without; “she has lost her way” |

Tabla 98. Descripción de -LACK_CUSTOM (concepto propuesto)

| | |
|----------------|---------------------------------------|
| Superordinado: | +HAVE_00 |
| Concepto: | +LACK_00 |
| MT: | (x1)Theme (x2: CUSTOM)Referent |
| Descripción | be without; “she lost her traditions” |

2.1.1 Resumen PERDER

En la tabla siguiente se presenta la representación completa de *perder* en FunGramKB. En la primera columna están los eventos conceptuales a los que se vinculan los

sentidos del verbo *perder* y en la segunda las colocaciones que se deben incorporar en el léxico en la plantilla léxica de *perder*. Recordemos que las colocaciones solo incluyen los argumentos que no se pudieron representar conceptualmente:

Tabla 99. Resumen: *PERDER* en FunGramKB (vinculación propuesta)

| Vínculo a evento conceptual | Colocaciones en plantilla léxica de <i>perder</i> |
|-----------------------------|--|
| \$MISS_00 | --- |
| +WASTE_00 | <i>oportunidad, posibilidad</i> |
| -WASTE_PERIOD | --- |
| +FAIL_00 | --- |
| -FAIL_TRIAL | --- |
| -FAIL_BATTLE | --- |
| -FAIL_COMPETITION | --- |
| +LOSE_00 | --- |
| -LACK_ PSYCHOLOGICAL_ATT | <i>poder, control, autoridad, dominio, hegemonía, soberanía; fuerza, energía, impulso, esperanza, curiosidad, aliento; equilibrio, personalidad, identidad, vida, compostura</i> |
| -LACK__SUBSTANCE | --- |
| -LACK__DIRECTION | --- |
| -LACK__CUSTOM | --- |

La próxima tabla resume la tipología de las clases que se combinan con *perder*, y los conceptos asignados en FunGramKB y REDES a cada clase léxica. Recordemos que las clases léxicas tipo I y II son representadas por preferencias de selección conceptuales. Mientras, las clases tipo III tendrá algunos miembros que no serán representados por una entidad y que, por lo tanto, necesitarán figurar como colocaciones en la plantilla léxica; los que se vinculan a la entidad figuran en la columna de argumentos del paradigma semántico, y los que no, en la columna de argumentos del paradigma discursivo o pragmático. Los miembros de las clases del tipo IV tendrán que figurar en su totalidad como colocaciones; se presentan en la tabla como argumentos del paradigma discursivo o pragmático:

Tabla 100. Tipología de clases léxicas que se combinan con *perder* con respecto a FunGramKB

| Tipo de clase | CONCEPTO en FunGramKB | CONCEPTO en REDES | Argumentos CL – paradigma semántico | Argumentos CL – paradigma discursivo pragmático |
|---------------|-------------------------|-------------------|--|---|
| I | +TRIAL_00 | CONFRONTACIÓN | <i>pleito, juicio, apelación</i> | --- |
| I | +BATTLE_00 | CONFRONTACIÓN | <i>batalla, lucha, encuentro, combate, guerra, pelea,</i> | --- |
| I | +COMPETITION_00 | CONFRONTACIÓN | <i>campeonato, final, set, partido,</i> | --- |
| I | +CUSTOM_00 | COSTUMBRE | <i>costumbre, hábito</i> | --- |
| I | +DIRECTION_00 | CURSO | <i>camino, dirección, rumbo, norte</i> | --- |
| II | +VEHICLE_00 | --- | <i>avión, tren, autobús, barco;</i> | --- |
| II | +TRAVEL_00 | --- | <i>transbordo, vuelo</i> | --- |
| II | +PERIOD_00 | --- | <i>día, mes, año, tiempo</i> | --- |
| II | +SOLID_00 | --- | <i>paraguas, pendientes, libro, cartera, foto</i> | --- |
| II | +PHYSICAL_ATTRIBUTES_00 | (CUALIDAD) | <i>color, brillantez</i> | --- |
| II | +SENSE_00 | --- | <i>vista, oído, olfato; apetito, gusto, sueño</i> | --- |
| II | +FEELING_00 | SENTIMIENTO | <i>miedo, temor, vergüenza, paciencia, nervios, alegría, optimismo, cariño, amor</i> | --- |
| II | +SUBSTANCE_00 | --- | <i>gasolina, pelo</i> | --- |
| III | +COGNITIVE_ATTRIBUTES | CUALIDAD | <i>memoria, lucidez</i> | <i>equilibrio, personalidad, identidad, vida, compostura,</i> |
| III | +DESIRE_00 | INCLINACIÓN | <i>deseo, interés, gana, ilusión,</i> | <i>fuerza, energía, impulso, esperanza, curiosidad, aliento</i> |
| IV | --- | --- | --- | <i>ocasión, oportunidad, posibilidad</i> |
| IV | --- | PODER | --- | <i>ocasión, oportunidad, posibilidad, ocasión, oportunidad, posibilidad</i> |

2.2 DISIPAR

La tabla siguiente presenta la información que ofrece REDES para el verbo *disipar*. REDES distingue los dos sentidos del verbo *disipar* como “literal” y “figurado”. Nosotros identificamos el sentido literal con la noción de ‘esparcir, separar’. El sentido figurado lo desdoblamos en dos nociones, ‘malgastar’ y ‘hacer desaparecer’.

Tabla 101. Sentidos y combinatoria léxica de *disipar*

| Sentido del verbo | CONCEPTO asignado a la CL en REDES | Descriptor de la CL en REDES | Argumentos miembros de la CL en REDES |
|---------------------|------------------------------------|---|--|
| ‘esparcir’ | --- | ciertas masas de gas | <i>viento, nube, aire, humo, gas, niebla</i> |
| | --- | manifestaciones de energía o sus efectos, esp. si se transmiten a través de ondas | <i>luz, sonido, fuerza, imagen, música</i> |
| ‘malgastar’ | (DINERO) | bienes materiales | <i>fortuna, dinero, hacienda, herencia</i> |
| ‘hacer desaparecer’ | INFORMACIÓN, INCERTIDUMBRE | informaciones desconocidas, no contrastadas o no fundamentadas; sentimientos de incertidumbre, inseguridad o recelo | <i>duda, sospecha, incógnita, rumor, interrogante, especulación, suspicacia, incertidumbre, susceptibilidad, prejuicio</i> |
| | PENSAMIENTO | nociones susceptibles de ser pensadas, evocadas o razonadas | <i>opinión, idea, recuerdo, teoría</i> |
| | INCLINACIÓN | susts. de carácter prospectivo que designan lo que se desea o se pretende alcanzar | <i>sueño, ilusión, expectativa, esperanza, utopía, quimera</i> |
| | CONFUSIÓN | situaciones enrevesadas, enredadas o faltas de claridad | <i>equivoco, malentendido, confusión, caos, marasmo</i> |
| | --- | situaciones difíciles, peligrosas, adversas o conflictivas | <i>peligro, problema, dificultad, amenaza, crisis, polémica, enfrentamiento, tensión</i> |
| | TEMOR | inquietud en diversos grados | <i>miedo, temor, pánico, susto, preocupación, angustia</i> |
| | AFLICCIÓN | estados de aflicción, esp. los provocados por la ausencia de voluntad, de ánimo o de interés | <i>tristeza, tedio, aburrimiento, pesimismo</i> |
| | SENTIMIENTO HOSTIL | sentimientos de irritación o de animadversión | <i>ira, resentimiento, hostilidad, crispación</i> |
| | SENTIMIENTO | satisfacción, exaltación y los contrarios a la clase G | <i>entusiasmo, euforia, agradecimiento, calma</i> |

(a) *disipar* ‘esparcir, separar’

El primer sentido de *disipar*, ‘esparcir, separar’, lo vinculamos preliminarmente al evento +SPLIT_00. El PS de este concepto expresa que un Tema no especificado (x1) cambia a un Referente no especificado (x2) con un instrumento, con el resultado de que ese Tema existe separado en sus partes:

Tabla 102. Descripción de +SPLIT_00 en FunGramKB

| | |
|-----------------|--|
| Superordinado: | +CHANGE_00 |
| MT: | (x1)Theme (x2)Referent |
| PS: | +(e1: +CHANGE_00 (x1)Theme (x2)Referent (f1)Instrument (f2: (e2: +EXIST_00 (x3: +PART_00)Theme))Result) |
| Descripción: | separate into parts or portions; "divide the cake into three equal parts"; "The British carved up the Ottoman Empire after World War I |
| Lexicalización: | <i>dividir, separar, partir</i> (español), <i>divide, separate, split, split up</i> (inglés) |

Este sentido ocurre cuando *disipar* se combina con dos clases léxicas, que hemos vinculado con +GAS_00 y +ENERGY_00. La tabla siguiente presenta la comparación entre la primera clase léxica y el concepto +GAS. Se trata de una clase descrita por un concepto superordinado (tipo II); todos los argumentos son lexicalizaciones de conceptos subordinados a +GAS: *viento*[+WIND_00], *nube*, *niebla* [+CLOUD_00], *aire*[+AIR_00], *humo*[+SMOKE_00], *gas*[+GAS_00]:

Tabla 103. Comparación clase léxica de REDES y +GAS_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|--|--|
| Superordinado: | | +SUBSTANCE_00 |
| Concepto: | --- | +GAS_00 |
| PS: | | +(e1: +BE_00 (x1: +GAS_00)Theme (x2: +SUBSTANCE_00)Referent) |
| Descripción: | ciertas masas de gas | a fluid in the gaseous state having neither independent shape nor volume and being able to expand indefinitely |
| Manifestación léxica | <i>viento, nube, aire, humo, gas, niebla</i> | <i>gas</i> (inglés) |

La segunda clase léxica se vincula a +ENERGY_00 es una clase mixta (tipo III). Uno de sus miembros, *fuerza*, puede vincularse a +ENERGY_00, y otro, *luz*, lexicaliza a +LIGHT_01, un subordinado de +ENERGY_00. Los demás son identificados en FunGramKB como atributos físicos: *sonido*[+SOUND_00], *imagen*[+IMAGE_00],

música[+MUSIC_00]; estos argumentos adquieren el sentido de ‘energía’ de una manera figurada:

Tabla 104. Comparación clase léxica de REDES y +ENERGY_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|---|---|
| Superordinado: | --- | +SUBSTANCE_00 |
| Concepto: | --- | +ENERGY_00 |
| PS: | --- | +(e1: +BE_00 (x1: +ENERGY_00)Theme (x2: +SUBSTANCE_00)Referent) (e2: +PRODUCE_00 (x3: +HUMAN_00)Theme (x1)Referent (f1: +COAL_00 ^ +WIND_00 ^ +SUN_00 ^ +FUEL_00)Origin) |
| Descripción: | manifestaciones de energía o sus efectos, esp. si se transmiten a través de ondas | (physics) the capacity of a physical system to do work; the units of energy are joules or ergs; "energy can take a wide variety of forms" |
| Manifestación léxica | <i>luz, sonido, fuerza, imagen, música</i> | <i>energy, power</i> (inglés) |

La combinatoria léxica de *disipar* nos revela algo adicional sobre su sentido: en este caso, la luz o la energía se separan en partes tan minúsculas que incluso pueden desaparecer. No es lo mismo separar en diferentes partes un bizcocho o una nación, que separar energía o luz. Por ello, proponemos el concepto terminal \$DISSIPATE_00, que especifica el Referente como +GAS_00 o +ENERGY_00:

Tabla 105. Descripción de \$DISSIPATE_00 en FunGramKB (concepto propuesto)

| | |
|-----------------|--|
| Superordinado: | +SPLIT_00 |
| MT: | (x1)Theme (x2: +GAS_00^+ENERGY_00)Referent |
| PS: | +((e1: +SPLIT_00 (x1)Theme (x2)Referent (f1: (e2: +EXIST_00 (x3: m +PART_00)Theme))Result)(e3: +BE_01 (x3)Theme (x4: m +SMALL_00)Attribute)) |
| Descripción: | to separate light or energy into minuscule parts or portions, to the point of disappearing: <i>the fog dissipated</i> |
| Lexicalización: | <i>disipar</i> (español), <i>disipar</i> (inglés) |

En la plantilla léxica de *disipar*[\$DISSIPATE_00], aparecerán como colocaciones los argumentos que no se pueden vincular a las entidades +GAS_00 y +ENERGY_00: *sonido, fuerza, imagen, música*.

(b) *disipar* ‘malgastar’

El segundo sentido de *disipar* lo vinculamos con el evento conceptual +WASTE_00, cuya descripción en FunGramKB repetimos aquí:

Tabla 106. Descripción de +WASTE_00 en FunGramKB

| | |
|-----------------|--|
| Superordinado: | +USE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2)Referent |
| PS: | +(e1: +USE_00 (x1)Theme (x2)Referent (f1: +CARELESS_00)Manner) |
| Descripción: | use inefficiently or inappropriately; "waste heat" |
| Lexicalización: | <i>malgastar</i> (español), <i>waste</i> (inglés) |

Este sentido ocurre con una clase léxica que vinculamos con la entidad conceptual +MONEY_00. Esta es una clase mixta (tipo III): los argumentos *fortuna*, *dinero* lexicalizan el concepto +MONEY_00, pero *hacienda* y *herencia* lo adquieren figuradamente. Cabe señalar aquí que el concepto +MONEY_00 tiene lexicalizaciones que pertenecen a diferentes variedades dialectales y diferentes registros: *cuarto*, *guita*, *mosca*, *parné*, *pasta*, *perras*, *plata*, *tela*; el modelo de FunGramKB permite especificar esta información en la entrada de cada palabra en el lexicón.

Tabla 107. Comparación entre clase léxica de REDES y +MONEY_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|---|--|
| Superordinado: | --- | +GROUP_00 |
| Concepto: | (DINERO) | +MONEY_00 |
| PS: | --- | +(e1: +BE_00 (x1: +MONEY_00)Theme (x2: +GROUP_00)Referent)(e2: +COMPRISE_00 (x2)Theme (x3: +COIN_00 +BANKNOTE_00 +CHEQUE_00)Referent)) |
| Descripción: | bienes materiales | the most common medium of exchange; functions as legal tender; "we tried to collect the money he owed us" |
| Manifestación léxica | <i>fortuna</i> , <i>dinero</i> , <i>hacienda</i> , <i>herencia</i> | <i>capital</i> , <i>cuarto</i> , <i>dinero</i> , <i>guita</i> , <i>mosca</i> , <i>parné</i> , <i>pasta</i> , <i>perras</i> , <i>plata</i> , <i>tela</i> (español), <i>capital</i> , <i>dosh</i> , <i>dough</i> , <i>money</i> (inglés) |

El verbo *disipar* demuestra, en su vínculo con +WASTE_00, una preferencia de selección exclusiva por +MONEY_00. Proponemos por lo tanto la creación del subconcepto -WASTE_MONEY, que se presenta en la siguiente tabla. Nótese aquí que entre las lexicalizaciones de -WASTE_MONEY figuran otros verbos *-derrochar*, *dilapidar*,

despilfarrar, desperdiciar– que comparten la preferencia de selección de +MONEY_00, además de *disipar*.

Tabla 108. Descripción de –WASTE MONEY (subconcepto propuesto)

| | |
|-----------------|---|
| Superconcepto: | +WASTE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2: +MONEY_00)Referent |
| Descripción: | use inefficiently or inappropriately; "waste money" |
| Lexicalización: | <i>disipar, derrochar, dilapidar, despilfarrar, desperdiciar</i> (español), <i>dissipate</i> (inglés) |

En la plantilla léxica de *disipar*[-WASTE MONEY] aparecerán como colocaciones los argumentos que no pueden vincularse a +MONEY_00: *hacienda y herencia*.

(c) *disipar* ‘hacer desaparecer’

El tercer sentido de *disipar*, ‘hacer desaparecer’, es el sentido dominante con respecto a la combinatoria del verbo. Ocurre con las nueve clases léxicas cerradas o clases léxicas en propiedad. Lo hemos vinculado con el concepto +FINISH_00, cuyo Tema y Referente no están especificados:

Tabla 109. Descripción de +FINISH_00 en FunGramKB

| | |
|-----------------|---|
| Hiperónimo: | +DO_00 |
| MT: | (x1)Theme (x2)Referent |
| PS: | +(e1: egr pres +DO_00 (x1)Theme (x2)Referent) +(e2: npast +DO_00 (x1)Theme (x2)Referent) |
| Descripción: | come or bring to a finish or an end; "He finished the dishes"; "She completed the requirements for her Master's Degree" |
| Lexicalización: | <i>acabar, concluir, rematar, terminar</i> (inglés), <i>complete, end, finish, stop</i> (español) |

La primera clase léxica de *disipar*[+FINISH_00] pertenece al tipo II. Es una clase descrita por +BELIEF_00, una noción que está superodinada a los conceptos asignados en REDES: INFORMACIÓN, INCERTIDUMBRE. Ninguno de los argumentos tiene presencia todavía en FunGramKB.

Tabla 110. Comparación entre clase léxica de REDES y +BELIEF_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|--|--|
| Superodinado: | --- | +THOUGHT_00 |
| Concepto: | INFORMACIÓN, INCERTIDUMBRE | +BELIEF_00 |
| PS: | --- | +(e1: +BE_00 (x1: +BELIEF_00)Theme (x2: +THOUGHT_00)Referent) +(e2: +KNOW_00 (x3: +HUMAN_00)Theme (x1)Referent) |
| Descripción: | informaciones desconocidas, no contrastadas o no fundamentadas; sentimientos de incertidumbre, inseguridad o recelo | any cognitive content held as true |
| Manifestación léxica | <i>duda, sospecha, incógnita, rumor, interrogante, especulación, suspicacia, incertidumbre, susceptibilidad, prejuicio</i> | <i>creencia</i> (español), <i>belief</i> (inglés) |

La segunda clase léxica es una clase mixta (tipo III), pero bastante estable. Se vincula con la entidad +THOUGHT_00. Los argumentos *idea* y *teoría* lexicalizan a +THOUGHT_00, mientras que *opinión* está vinculado a +OPINION_00, que es subordinado de +THOUGHT_00. Por su parte, *recuerdo* está vinculado a +MEMORY_00, que tiene otra ruta conceptual, pero adquiere este sentido figuradamente.

Tabla 111. Comparación entre clase léxica de REDES y +THOUGHT_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|---|---|
| Superodinado: | --- | #PROPOSITION |
| Concepto: | PENSAMIENTO | +THOUGHT_00 |
| PS: | --- | sp |
| Descripción: | nociones susceptibles de ser pensadas, evocadas o razonadas | any cognitive content |
| Manifestación léxica | <i>opinión, idea, recuerdo, teoría</i> | <i>concepto, reflexión, idea, mente, pensamiento</i> (español), <i>concepto, consideration, idea, meditation, notion</i> (inglés) |

La próxima clase léxica se vincula con +DESIRE_00. Es una clase mixta (tipo III): los argumentos *sueño, ilusión, expectativa, esperanza* y *quimera* pueden lexicalizar a +DESIRE_00, pero *utopía* no.

Tabla 112. Comparación entre clase léxica de REDES y +DESIRE_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|--|--|
| Superodinado: | --- | +FEELING_00 |
| Concepto: | INCLINACIÓN | +DESIRE_00 |
| PS: | --- | +(e1: +BE_00 (x1: +DESIRE_00)Theme (x2: +FEELING_00)Referent) *(e2: +FEEL_00 (x3: (e3: +DESIRE_01 (x4: +HUMAN_00)Theme (x5)Referent))Agent (x4)Theme (x1)Attribute) |
| Descripción: | susts. de carácter prospectivo que designan lo que se desea o se pretende alcanzar | a strong hope or wish |
| Manifestación léxica | <i>sueño, ilusión, expectativa, esperanza utopía, quimera</i> | <i>deseo</i> (español), <i>desire, wish</i> (inglés) |

La próxima clase pertenece al tipo IV: no se puede vincular a ninguna entidad. El descriptor de la clase se acerca bastante a la descripción del concepto +MYSTERY_00, pues comparten la idea de algo que no se entiende, pero las lexicalizaciones son diferentes. Estos argumentos serán entonces representados como colocaciones en la plantilla léxica de *disipar*[+FINISH_00]:

Tabla 113. Comparación entre clase léxica de REDES y +MYSTERY_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB (*no hay vínculo) |
|----------------------|---|--|
| Superodinado: | --- | +STATE_00 |
| Concepto: | CONFUSIÓN | +MYSTERY_00 |
| PS: | --- | +(e1: +BE_00 (x1: +MYSTERY_00)Theme (x2: +STATE_00)Referent) *(e2: n +UNDERSTAND_00 (x3: m +HUMAN_00)Theme (x1)Referent (f1: x1)Scene (f2: (e3: pos n +EXPLAIN_00 (x3)Theme (x1)Referent (x4: +HUMAN_00)Goal))Result) |
| Descripción: | situaciones enrevesadas, enredadas o faltas de claridad | something that baffles understanding and cannot be explained; "how it got out is a mystery" |
| Manifestación léxica | <i>equivoco, caos, malentendido, confusión, marasmo</i> | <i>enigma, misterio</i> (español), <i>enigma, mystery</i> (inglés) |

La próxima clase léxica de REDES se desdobra en dos entidades en FunGramKB: +DANGER_00 y +PROBLEM_00. Cabe señalar que hay cercanía entre los conceptos, porque comparten el superordinado +STATE_00. La clase léxica se describe como ‘situaciones difíciles, peligrosas, adversas o conflictivas’, que incluye las dos nociones: ‘peligro’ y ‘problema’. Los argumentos *peligro*, *amenaza* lexicalizan +DANGER_00; esta se convierte en una clase estable (tipo I):

Tabla 114. Comparación entre clase léxica de REDES y +DANGER_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|--|---|
| Superordinado: | --- | +STATE_00 |
| Concepto: | --- | +DANGER_00 |
| PS: | --- | +(e1: +BE_00 (x1: +DANGER_00)Theme (x2: +STATE_00)Referent) *(e2: pos +DAMAGE_00 (x3)Theme (x4)Referent (f1: x1)Scene) |
| Descripción: | situaciones difíciles, peligrosas, adversas o conflictivas | the condition of being susceptible to harm or injury; "you are in no danger" |
| Manifestación léxica | <i>peligro</i> , <i>amenaza</i> , | <i>riesgo</i> , <i>peligro</i> (español), <i>danger</i> , <i>peril</i> , <i>risk</i> (inglés) |

Los demás argumentos *-problema*, *dificultad*, *amenaza*, *crisis*, *polémica*, *enfrentamiento*, *tensión-* se relacionan con el concepto +PROBLEM_00. Esta es una clase mixta (tipo III): *problema*, *dificultad*, *polémica*, *crisis* lexicalizan a +PROBLEM_00, mientras que *tensión* y *enfrentamiento* adquieren este sentido figuradamente.

Tabla 115. Comparación entre clase léxica de REDES y +PROBLEM_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|--|--|
| Superodinado: | --- | +STATE_00 |
| Concepto: | --- | +PROBLEM_00 |
| PS: | --- | +(e1: +BE_00 (x1: +PROBLEM_00)Theme (x2: +STATE_00)Referent) +(e2: pos n +DO_00 (x3: +HUMAN_00)Theme (x4)Referent (f1: (e3: +BE_01 (x4)Theme (x5: +DIFFICULT_00)Attribute))Reason (f2: x1)Result) |
| Descripción: | situaciones difíciles, peligrosas, adversas o conflictivas | a factor causing trouble in achieving a positive result or tending to produce a negative result; "serious difficulties were encountered in obtaining a pure reagent" |
| Manifestación léxica | <i>problema, tensión, dificultad, crisis, polémica, enfrentamiento</i> | <i>dificultad, problema</i> (español), <i>difficulty, problem, trouble</i> (inglés) |

La próxima clase léxica, que REDES asocia con TEMOR, la hemos vinculado con la entidad +FEAR_00. Es una clase tipo II, porque está descrita por un concepto superodinado: los argumentos *miedo, temor, pánico, susto* lexicalizan el concepto +FEAR_00 y *preocupación y angustia* lexicalizan al concepto +ANXIETY_00, que es subordinado de +FEAR_00.

Tabla 116. Comparación entre clase léxica de REDES y +FEAR_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|--|---|
| Superodinado: | --- | +FEELING_00 |
| Concepto: | TEMOR | +FEAR_00 |
| PS: | --- | +(e1: +BE_00 (x1: +FEAR_00)Theme (x2: +FEELING_00)Referent) *((e2: +FEEL_00 (x3: (e3: +THINK_00 (x4: +HUMAN_00)Theme (x5: (e4: +DO_00 (x6)Theme (x7)Referent))Referent))Agent (x4)Theme (x1)Attribute)(e5: +DISLIKE_00 (x7)Agent (x4)Theme)) |
| Descripción: | inquietud en diversos grados | the feeling you get when you are afraid |
| Manifestación léxica | <i>miedo, temor, pánico, susto, preocupación, angustia</i> | <i>aprensión, miedo, temor</i> (español), <i>fear, fright</i> (inglés) |

La próxima clase léxica, que REDES asocia a AFLICCIÓN, la hemos vinculado a la entidad +SADNESS_00. Se trata de una clase estable (tipo I); todos los argumentos pueden lexicalizar el concepto +SADNESS_00: *tristeza, tedio, aburrimiento, pesimismo*.

Tabla 117. Comparación entre clase léxica de REDES y +SADNESS_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|--|---|
| Superodinado: | --- | +FEELING_00 |
| Concepto: | AFLICCIÓN | +SADNESS_00 |
| PS: | --- | +(e1: +BE_00 (x1: +SADNESS_00)Theme (x2: +FEELING_00)Referent) *(e2: +FEEL_00 (x3: (e3: +DISLIKE_00 (x4)Agent (x5: +HUMAN_00)Theme))Agent (x5)Theme (x1)Attribute) |
| Descripción: | estados de aflicción, esp. los provocados por la ausencia de voluntad, de ánimo o de interés | the state of being sad; "she tired of his perpetual sadness" |
| Manifestación léxica | <i>tristeza, tedio, aburrimiento, pesimismo</i> | (español), (inglés) |

La próxima clase léxica, que REDES asocia a SENTIMIENTO HOSTIL, la hemos vinculado a la entidad +ANGER_00. Esta también es una clase estable (tipo I); todos los argumentos pueden lexicalizar el concepto +ANGER_00: *ira, resentimiento, hostilidad, crispación*:

Tabla 118. Comparación entre clase léxica de REDES y +ANGER_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|---|---|
| Superodinado: | --- | +FEELING_00 |
| Concepto: | SENTIMIENTO HOSTIL | +ANGER_00 |
| PS: | --- | +(e1: +BE_00 (x1: +ANGER_00)Theme (x2: +FEELING_00)Referent) *(e2: +FEEL_00 (x3: (e3: +DISLIKE_00 (x4)Agent (x5: +HUMAN_00)Theme))Agent (x5)Theme (x1)Attribute (f1: (e4: +WANT_00 (x5)Theme (x6: (e5: +SHOUT_00 (x5)Theme (x7)Referent (x8)Goal))Referent))Result (f2: (e6: +WANT_00 (x5)Theme (x9: (e7: +HURT_00 (x5)Theme (x8)Referent))Referent))Result) |
| Descripción: | sentimientos de irritación o de animadversión | a feeling which makes you want to shout at someone or hurt them because they have behaved in an unfair, cruel, offensive etc. way, or because you think that a situation is unfair, unacceptable etc |
| Manifestación léxica | <i>ira, resentimiento, hostilidad, crispación</i> | <i>cabreo, enfado, enojo</i> (español), <i>anger</i> (inglés) |

REDES asocia la última clase léxica de *disipar*[+FINISH_00] a SENTIMIENTO; nosotros la hemos vinculado en FunGramKB a \$ENTHUSIASM_00. Es una clase mixta (tipo III), en donde *entusiasmo* y *euforia* pueden lexicalizar el concepto, pero *agradecimiento* y *calma* asumen el significado figuradamente.

Tabla 119. Comparación entre clase léxica de REDES y \$ENTHUSIASM_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|--|---------------------|
| Superodinado: | --- | +FEELING_00 |
| Concepto: | SENTIMIENTO | \$ENTHUSIASM_00 |
| Descripción: | satisfacción, exaltación y los contrarios a la clase G | --- |
| Manifestación léxica | <i>entusiasmo, euforia, agradecimiento, calma</i> | (español), (inglés) |

Presentamos en la próxima tabla las rutas conceptuales de las clases léxicas con que se combina *disipar*[+FINISH_00], el sentido más dominante de *disipar*, a la luz de su combinatoria. Todas las selecciones de este verbo en este sentido son entidades abstractas, y queda claro que hay una preferencia marcada por los ‘sentimientos’:

Tabla 120. Rutas conceptuales de conceptos vinculados a clases léxicas que se combinan con *disipar*[+FINISH_00]

| Rutas conceptuales -----> -----> -----> -----> -----> | | | | | |
|---|--------------|------------|---------------|----------|--------------|
| #Abstract | +State | +Danger | | | |
| | | +Problem | | | |
| | #Proposition | +Thought | | | |
| | | | +Belief | | |
| | | +Attribute | +Physical_Att | +Feeling | +Fear |
| | | | | | +Sadness |
| | | | | | \$Enthusiasm |
| | | | | | +Anger |
| | | | | | +Desire |

A raíz de esto, proponemos la creación de los siguientes subconceptos. En primer lugar, -FINISH_FEELING, que recoge las preferencias por +FEAR_00, +SADNESS_00, \$ENTHUSIASM_00, +ANGER_00 y +DESIRE_00. No podemos hacer lo mismo con +STATE_00 y +THOUGHT_00, porque aunque son superordinados de otros conceptos seleccionados, son conceptos amplios que contienen muchos otros subordinados que no son seleccionados. Por lo tanto, proponemos los siguientes subconceptos: -FINISH_DANGER, -FINISH_PROBLEM, -FINISH_THOUGHT, -FINISH_BELIEF.

Tabla 121. Descripción de -FINISH_FEELING (subconcepto propuesto)

| | |
|-----------------|--|
| Hiperónimo: | +DO_00 |
| MT: | (x1)Theme (x2: +FEELING_00)Referent |
| Descripción: | bring a feeling to a finish or an end; “her doubts dissipated” |
| Lexicalización: | <i>disipar</i> (español), <i>dissipate</i> (inglés) |

Tabla 122. Descripción de -FINISH_DANGER (subconcepto propuesto)

| | |
|-----------------|---|
| Hiperónimo: | +DO_00 |
| MT: | (x1)Theme (x2: +DANGER_00)Referent |
| Descripción: | bring a danger to a finish or an end; “the risk dissipated” |
| Lexicalización: | <i>disipar</i> (español), <i>dissipate</i> (inglés) |

Tabla 123. Descripción de -FINISH_PROBLEM (subconcepto propuesto)

| | |
|-----------------|--|
| Hiperónimo: | +DO_00 |
| MT: | (x1)Theme (x2: +PROBLEM_00)Referent |
| Descripción: | come or bring a problem to a finish or an end; “the difficulties dissipated” |
| Lexicalización: | <i>disipar</i> (español), <i>dissipate</i> (inglés) |

Tabla 124. Descripción de -FINISH_THOUGHT (subconcepto propuesto)

| | |
|-----------------|--|
| Hiperónimo: | +DO_00 |
| MT: | (x1)Theme (x2: +THOUGHT_00)Referent |
| Descripción: | bring a thought to a finish or an end; “those opinions dissipated” |
| Lexicalización: | <i>disipar</i> (español), <i>dissipate</i> (inglés) |

Tabla 125. Descripción de -FINISH_BELIEF (subconcepto propuesto)

| | |
|-----------------|--|
| Hiperónimo: | +DO_00 |
| MT: | (x1)Theme (x2: +BELIEF_00)Referent |
| Descripción: | bring a belief to a finish or an end; “her faith dissipated” |
| Lexicalización: | <i>disipar</i> (español), <i>dissipate</i> (inglés) |

2.2.1 Resumen *DISIPAR*

La próxima tabla muestra el resumen de los vínculos del verbo *disipar* en FunGramKB. La columna de la izquierda muestra el vínculo con un evento verbal, mientras que la de la derecha muestra las colocaciones que se ubicarán en la plantilla léxica de *disipar* en ese sentido; solo los argumentos que no están representados conceptualmente en una preferencia de selección pasan a figurar como colocaciones.

Tabla 126. Resumen: *DISIPAR* en FunGramKB (vinculación propuesta)

| Evento conceptual | Colocaciones |
|-------------------|---|
| \$DISSIPATE_00 | <i>sonido, imagen, música</i> |
| -WASTE_MONEY | <i>hacienda, herencia</i> |
| +FINISH_00 | <i>equivoco, caos, malentendido, confusión, marasmo</i> |
| -FINISH_FEELING | <i>utopía y quimera; agradecimiento y calma</i> |
| -FINISH_DANGER | --- |
| -FINISH_PROBLEM | <i>tensión, enfrentamiento</i> |
| -FINISH_THOUGHT | <i>opinión, recuerdo</i> |
| -FINISH_BELIEF | --- |

La tabla siguiente resume la tipología de las clases con que se combina *disipar*, y los conceptos asignados en FunGramKB y REDES a cada clase léxica. Recordemos que las clases léxicas tipo I y II son representadas por preferencias de selección conceptuales. Mientras, las clases tipo III tendrán algunos miembros que pertenecen al paradigma semántico y otros que pertenecen al paradigma discursivo o pragmático; estos últimos se incluirán como colocaciones en la plantilla léxica. Los miembros de las clases del tipo IV tendrán que figurar en su totalidad como colocaciones; pertenecen pues al paradigma discursivo o pragmático:

Tabla 127. Tipología de clases léxicas que se combinan con disipar con respecto a FunGramKB

| Tipo de CL | CONCEPTO en FunGramKB | CONCEPTO de CL en REDES | Argumentos CL – paradigma semántico | Argumentos CL – paradigma discursivo o pragmático |
|------------|-----------------------|----------------------------|---|--|
| I | +SADNESS_00 | AFLICCIÓN | <i>tristeza, aburrimiento, pesimismo, tedio,</i> | --- |
| I | +ANGER_00 | SENTIMIENTO HOSTIL | <i>resentimiento, hostilidad, crispación, ira</i> | --- |
| I | +DANGER_00 | --- | <i>amenaza, problema, polémica, enfrentamiento, tensión, peligro, crisis, dificultad,</i> | --- |
| II | +GAS_00 | --- | <i>niebla, viento, aire, humo, gas, nube,</i> | --- |
| II | +BELIEF_00 | INFORMACIÓN, INCERTIDUMBRE | <i>duda, incógnita, interrogante, especulación, suspicacia, incertidumbre, susceptibilidad, prejuicio, sospecha, rumor,</i> | --- |
| II | +FEAR_00 | TEMOR | <i>miedo, reocupación, susto, angustia, temor, pánico,</i> | --- |
| III | +ENERGY_00 | --- | <i>luz, fuerza,</i> | <i>sonido, imagen, música</i> |
| III | +MONEY_00 | --- | <i>fortuna, dinero</i> | <i>hacienda, herencia</i> |
| III | +THOUGHT_00 | PENSAMIENTO | <i>idea, teoría</i> | <i>opinión, recuerdo</i> |
| III | +PROBLEM_00 | --- | <i>problema, crisis, polémica, dificultad,</i> | <i>tensión, enfrentamiento</i> |
| III | +DESIRE_00 | INCLINACIÓN | <i>ilusión, expectativa, esperanza, quimera, sueño,</i> | <i>utopía</i> |
| III | \$ENTHUSIASM_00 | SENTIMIENTO | <i>entusiasmo, euforia,</i> | <i>agradecimiento, calma</i> |
| IV | --- | CONFUSIÓN | --- | <i>equivoco, malentendido, confusión, marasmo, caos,</i> |

2.3 DERROCHAR

Para el verbo *derrochar*, REDES indica que existe un sentido “físico” y otro “figurado”. Nosotros equiparamos el sentido físico a ‘malgastar’ y el figurado a ‘manifestar en abundancia’. La tabla siguiente presenta los datos que nos da REDES:

Tabla 128. Sentidos y combinatoria léxica de *derrochar*

| Sentido del verbo | CONCEPTO asignado a la CL en REDES | Descriptor de la CL en REDES | Argumentos miembros de la CL en REDES |
|----------------------------|------------------------------------|--|---|
| ‘malgastar’ | (DINERO) | dinero u otros bienes | <i>fortuna, dinero, riqueza</i> |
| | (RECURSO) | diversos recursos | <i>munición, agua</i> |
| ‘manifestar en abundancia’ | SENTIMIENTO | sentimientos o sensaciones de complacencia, entusiasmo y afecto, a menudo elevados | <i>alegría, satisfacción, felicidad, euforia, optimismo, entusiasmo, ilusión, pasión, efusión, amor, cariño</i> |
| | CUALIDAD | valor o coraje, a veces extremos | <i>valor, valentía, arrojo, bravura, coraje, heroísmo</i> |
| | CUALIDAD | voluntad, tesón | <i>empeño, voluntad, entrega, ganas</i> |
| | CUALIDAD | serenidad o comedimiento, generalmente ante adversidad | <i>prudencia, modestia, mesura, seguridad, templanza, aplomo, temple</i> |
| | CUALIDAD, FUERZA | fuerza, energía, y rasgos expansivos de personalidad o conducta | <i>fuerza, carácter, vitalidad, temperamento, energía, fortaleza, garra, casta, vigor, poderío, raza, fiereza, empuje</i> |
| | VIRTUD | otras virtudes, capacidades y aptitudes, frec. naturales | <i>destreza, facultad, talento, inteligencia, dotes virtud, cualidad</i> |
| | CUALIDAD | facultades, naturales o adquiridas relacionadas con el conocimiento y buen juicio, o su ejercicio | <i>sensatez, sabiduría, sentido común, coherencia, conocimiento, idea, criterio, saber</i> |
| | --- | facultades mentales relacionadas a la creatividad | <i>imaginación, creatividad, ingenio, fantasía, genio, inventiva</i> |
| | CUALIDAD | cualidades, actitudes y facultades relativas a la capacidad de elegir o a la forma de comportarse o presentarse ante los demás | <i>estilo, elegancia, clase, buen gusto, glamour, lujo, estética, fineza, finura</i> |

| | | | |
|--|-------------|--|--|
| | CUALIDAD | jovialidad, extroversión, desenfado, atractivo personal y cualidades análogos | <i>simpatía, carisma, naturalidad, desparpajo, espontaneidad, gracia, encanto, salero, sonrisa</i> |
| | CUALIDAD | saber hacer, profesionalidad, y rasgos de experiencia | <i>soltura, oficio, tablas, elocuencia, dominio</i> |
| | CUALIDAD | cualidades del carácter, más frec. si son sobre su disposición a los demás | <i>sinceridad, amabilidad, generosidad, caridad, ternura, honradez, humanidad, pundonor, bondad, piedad, nobleza</i> |
| | --- | sentimientos negativos, esp. insatisfacción e ira | <i>rabia, saña, furor, pesimismo, frustración, amargura</i> |
| | SENTIMIENTO | orgullo, soberbia, vanidad | <i>altanería, autoestima, vanidad, presunción,</i> |
| | --- | actitudes o manifestaciones humanísticas, mordaces o burlescas, a veces con intención solapada | <i>humor, ironía, malicia, sarcasmo, socarronería</i> |

(a) derrochar ‘malgastar’

El sentido físico de ‘malgastar’ lo vinculamos preliminarmente con el evento conceptual +WASTE_00:

Tabla 129. Descripción de +WASTE_00 en FunGramKB

| | |
|-------------------------------|--|
| Hiperónimo: | +USE_00 |
| Marco temático (MT) | (x1: +HUMAN_00)Theme (x2)Referent |
| Postulado de significado (PS) | +(e1: +USE_00 (x1)Theme (x2)Referent (f1: +CARELESS_00)Manner) |
| Descripción | use inefficiently or inappropriately; "waste heat" |
| Lexicalización | <i>malgastar</i> (español), <i>waste</i> (inglés) |

Este sentido ocurre cuando *derrochar* entra en combinación con dos clases léxicas. La primera la vinculamos a +MONEY_00. Es una clase estable (tipo I); todos sus miembros pueden lexicalizar el concepto +MONEY_00:

Tabla 130. Comparación entre clase léxica de REDES y +MONEY_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|---------------------------------|---|
| Superodinado: | --- | +GROUP_00 |
| Concepto: | DINERO | +MONEY_00 |
| PS: | --- | +((e1: +BE_00 (x1: +MONEY_00)Theme (x2: +GROUP_00)Referent)(e2: +COMPRISE_00 (x2)Theme (x3: +COIN_00 +BANKNOTE_00 +CHEQUE_00)Referent)) |
| Descripción: | dinero u otros bienes | the most common medium of exchange; functions as legal tender; "we tried to collect the money he owed us" |
| Manifestación léxica | <i>fortuna, dinero, riqueza</i> | <i>capital, cuarto, dinero, guita, mosca, parné, pasta, perras, plata, tela</i> (español), <i>capital, dosh, dough, money</i> (inglés) |

Este sentido también ocurre junto con la clase descrita como ‘diversos recursos’, cuyos miembros son *munición* y *agua*. Esta clase pertenece al tipo IV: no se puede vincular a ningún concepto en FunGramKB, porque la noción de ‘recurso’, que agrupa la clase léxica, no puede tener representación en una ontología, por depender del contexto pragmático.

Ahora bien, con respecto a +WASTE_00, el predicado *derrochar* –al igual que *despilfarrar*, que veremos más adelante– expresa un nivel más extremo del evento de ‘malgastar’. Así lo reflejan los diccionarios (subrayamos los contornos relevantes):

Tabla 131. Información lexicográfica de *derrochar* ‘malgastar en exceso’

| |
|---|
| Hiperónimos: |
| malgastar (x2), emplear, gastar (x2) |
| Contornos: |
| una persona (x3)// el dinero (x3) o hacienda; otras cosas que posee, como el valor, las energías, el humor; otra cosa (x2) // <u>excesivamente</u> , con insensatez o <u>exceso</u> (x2), <u>demasiado</u> , de forma insensata o sin necesidad |
| Colocaciones: |
| <i>el dinero</i> (x2), <i>el agua</i> , <i>el sueldo</i> , <i>energías</i> , <i>recursos</i> , <i>gasolina</i> |
| Sinónimos: |
| <i>despilfarrar</i> , <i>dilapidar</i> , <i>disipar</i> , <i>malgastar</i> , <i>tirar</i> |

Proponemos un cambio al PS de +WASTE_00 que refuerce el evento de +WASTE_00 con el cuantificador relativo *m* (de *much*) (Periñán y Mairal, 2010). El concepto terminal se llamaría \$\$SQUANDER_00, y estaría lexicalizado por *despilfarrar* y *derrochar* en español y *squander* en inglés. La propuesta para un nuevo concepto terminal \$\$SQUANDER_00 luce así:

Tabla 132. Descripción de \$SQUANDER_00 (concepto propuesto)

| | |
|-----------------|--|
| Hiperónimo: | +WASTE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2)Referent |
| PS: | +(e1: WASTE_00 (x1)Theme (x2)Referent (f1: +MUCH_00)Quantity |
| Descripción: | to waste in excess: “squander money” |
| Lexicalización: | <i>despilfarrar, derrochar</i> (español), <i>squander</i> (inglés) |

En este caso, dejaremos el Referente sin especificar, para dar cabida a la noción de ‘recurso’, que no se pudo vincular a una entidad conceptual. Por esta razón, en la plantilla léxica de *derrochar*[+\$SQUANDER_00], ubicaremos como colocaciones todos los argumentos con que se combina: *fortuna, dinero, riqueza; munición, agua*.

(b) *derrochar* ‘manifestar en abundancia’

El sentido dominante de *derrochar*, que ocurre con la mayor parte de su combinatoria léxica, es ‘manifestar en abundancia’ y lo vinculamos al evento \$ABOUND_00. Este concepto tiene típicamente un Tema humano o animal y un Referente que puede ser corpuscular, animal, sustancia o organización. El PS expresa que el Tema tiene al Referente en gran cantidad:

Tabla 133. Descripción de \$ABOUND_00

| | |
|-----------------|--|
| Superordinado: | +HAVE_00 |
| MT: | (x1: +HUMAN_00^+ANIMAL_00)Theme (x2: +CORPUSCULAR_00^+ANIMAL_00^+SUBSTANCE_00^+ORGANIZATION_00) Referent |
| PS: | +(e1: +HAVE_00 (x1)Theme (x2)Referent (f1: +MUCH_00)Quantity) |
| Descripción: | Have in large numbers |
| Lexicalización: | <i>abundar, rebosar</i> (inglés), <i>abound, be rich</i> (español) |

Veamos ahora con qué clases léxicas se combina *derrochar* para activar este sentido. La primera clase léxica se vincula al concepto terminal +PLEASURE_00. Esta clase léxica es una clase mixta (tipo III), porque tiene algunos miembros que lexicalizan sin problema la noción de +PLEASURE_00, como *alegría, satisfacción, felicidad, euforia, efusión*, pero el resto –*optimismo, entusiasmo, ilusión, pasión, amor, cariño*– adquiere ese valor de manera figurada:

Tabla 134. Comparación entre clase léxica de REDES y +PLEASURE_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|-----------------------|---|--|
| Superodinado: | --- | +FEELING_00 |
| Concepto: | SENTIMIENTO | +PLEASURE_00 |
| PS: | --- | +(e1: +BE_00 (x1: +PLEASURE_00)Theme (x2: +FEELING_00)Referent) *(e2: +FEEL_00 (x3: (e3: +LIKE_00 (x4)Agent (x5: +HUMAN_00)Theme))Agent (x5)Theme (x1)Attribute) |
| Descripción: | sentimientos o sensaciones de complacencia, entusiasmo y afecto, a menudo elevados | the feeling of happiness, enjoyment, or satisfaction that you get from an experience |
| Manifestación léxica: | <i>alegría, satisfacción, felicidad, euforia, optimismo, entusiasmo, ilusión, pasión, efusión, amor, cariño</i> | <i>agrado, alegría, contentamiento, contento, deleite, dicha, felicidad, gozo, gusto, júbilo, placer, regocijo, satisfacción (español), bliss, cheer, contentment, delight, enjoyment, felicity, gladness, happiness, joy, jubilation, pleasure, relish, satisfaction, zest (inglés)</i> |

La segunda clase léxica se vincula a +COURAGE_00. Es una de nueve clases léxicas que REDES identifica con el concepto CUALIDAD, pero que en FunGramKB se vinculan con varias entidades conceptuales diferentes. Esta clase es estable (tipo I): todos sus miembros –*valor, valentía, arrojo, bravura, coraje, heroísmo*– pueden lexicalizar el concepto +COURAGE_00 y la descripción del concepto aplica a la clase léxica.

Tabla 135. Comparación entre clase léxica de REDES y +COURAGE_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|---|--|
| Superodinado: | --- | +BEHAVIOUR_00 |
| Concepto: | CUALIDAD | +COURAGE_00 |
| PS: | --- | +(e1: +BE_00 (x1: +COURAGE_00)Theme (x2: +BEHAVIOUR_00)Referent) *((e2: +DO_00 (x3: +HUMAN_00)Theme (x4)Referent (f1: x1)Manner)(e3: +BE_01 (x4)Theme (x5: +DANGEROUS_00)Attribute)) |
| Descripción: | valor o coraje, a veces extremos | the quality of being brave when you are in danger, in pain, in a difficult situation etc |
| Manifestación léxica | <i>valor, valentía, arrojo, bravura, coraje, heroísmo</i> | <i>agalla, arrojo, audacia, braveza, bravura, coraje, desnudo, huevo, intrepidez, valentía, valor (español), ball, boldness, bottle, bravery, courage, daring, grit, gut. intrepidity, mettle, nerve, spunk (inglés)</i> |

La tercera clase léxica, que REDES identifica con CUALIDAD, la hemos vinculado a +EFFORT_00. Esta es una clase estable (tipo I), dado que todos sus miembros podrían lexicalizar a +EFFORT_00:

Tabla 136. Comparación entre clase léxica de REDES y +EFFORT_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|---|--|
| Superodinado: | --- | +OCCURRENCE_00 |
| Concepto: | CUALIDAD | +EFFORT_00 |
| PS: | --- | +(e1: +BE_00 (x1: +EFFORT_00)Theme (x2: +OCCURRENCE_00)Referent) +(e2: +TRY_00 (x3: +HUMAN_00 ^ +ANIMAL_00)Theme (x4)Referent (f1: +MUCH_00)Quantity (f2: x1)Scene) |
| Descripción: | voluntad, tesón | earnest and conscientious activity intended to do or accomplish something: "made an effort to cover all the reading material"; "wished him luck in his endeavor"; "she gave it a good try" |
| Manifestación léxica | <i>empeño, voluntad, entrega, ganas</i> | <i>afán, ahínco, ánimo, animosidad, brío, empeño, esfuerzo, intento, sudor</i> (español), <i>attempt, effort, endeavour, try</i> (inglés) |

Hay dos clases léxicas que REDES también identifica como CUALIDAD, que hemos vinculado a +WISDOM_00. Se trata de dos clases mixtas (tipo III), que incluyen miembros que lexicalizan el concepto +WISDOM_00, como *prudencia, medida, sensatez, sabiduría, sentido común*, pero también otros, que adquieren el sentido de manera figurada, como, *coherencia, conocimiento, idea, criterio, saber*:

Tabla 137. Comparación entre clase léxica de REDES y +WISDOM_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|---|---|
| Superodinado: | --- | +INTELLIGENCE_00 |
| Concepto: | CUALIDAD | +WISDOM_00 |
| PS: | --- | +(e1: +BE_00 (x1: +WISDOM_00)Theme (x2: +INTELLIGENCE_00)Referent) |
| Descripción: | serenidad o comedimiento, generalmente ante adversidad /// facultades, naturales o adquiridas relacionadas con el conocimiento y buen juicio, o su ejercicio | the ability to make sensible decisions about what to do and when to do it |
| Manifestación léxica | <i>prudencia, modestia, medida, seguridad, templanza, aplomo, temple</i> /// <i>sensatez, sabiduría, sentido común, coherencia, conocimiento, idea, criterio, saber</i> | <i>cordura, juicio, prudencia, sensatez, sentido común</i> , (español), <i>common sense, judgment, prudence, sense, wisdom</i> , (inglés) |

La próxima clase léxica la vinculamos con +STRENGHT_00. Se trata nuevamente de una clase mixta (tipo III), en la que solo algunos de los argumentos lexicalizan el concepto, por ejemplo, *fuerza, energía, fortaleza, vigor, poderío, empuje*. Los demás adquieren el sentido de modo figurado: *carácter, vitalidad, temperamento, garra, casta, raza, fiereza*.

Tabla 138. Comparación entre clase léxica de REDES y +STRENGHT_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|---|--|
| Superordinado: | --- | +PHYSICAL_ATT_00 |
| Concepto: | CUALIDAD, FUERZA | +STRENGHT_00 |
| PS: | --- | +(e1: +BE_00 (x1: +STRENGTH_00)Theme (x2: +PHYSICAL_ATT_00)Referent) *(e2: +BE_01 (x3: +HUMAN_00 ^ +ANIMAL_00)Theme (x1)Attribute) |
| Descripción: | fuerza, energía, y rasgos expansivos de personalidad o conducta | the property of being physically or mentally strong; "fatigue sapped his strenght" |
| Manifestación léxica | <i>fuerza, carácter, vitalidad, temperamento, energía, fortaleza, garra, casta, vigor, poderío, raza, fiereza, empuje</i> | <i>fuerza</i> (español), <i>force, strenght</i> (ing) |

Hay dos clases léxicas que hemos vinculado a la entidad +ABILITY_00. REDES identifica una con VIRTUD y otra con CUALIDAD. La que REDES asocia con VIRTUD es una clase mixta (tipo III), donde los miembros *destreza, facultad, talento, dotes* lexicalizan el concepto +ABILITY_00, *inteligencia* corresponde a un concepto subordinado, y los demás pueden adquirir ese sentido en combinación con el predicado. La que REDES asocia con CUALIDAD es una clase que se describe por medio de un superordinado (tipo II). El descriptor ‘saber hacer, profesionalidad, y rasgos de experiencia’ se acerca bastante a la descripción de +ABILITY_00, pero sus miembros son todos subordinados: *soltura, oficio, tablas, elocuencia, dominio*.

Tabla 139. Comparación entre clase léxica de REDES y +ABILITY_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|--|---|
| Superodinado: | --- | +PSYCHOLOGICAL_ATT_00 |
| Concepto: | VIRTUD /// CUALIDAD | +ABILITY_00 |
| PS: | --- | +(e1: +BE_00 (x1: +ABILITY_00)Theme (x2: +PSYCHOLOGICAL_ATT_00)Referent) *(e2: pos +DO_00 (x3: +HUMAN_00 ^ +ANIMAL_00)Theme (x4)Referent (f1: x1)Means) |
| Descripción: | otras virtudes, capacidades y aptitudes, frec. naturales /// saber hacer, profesionalidad, y rasgos de experiencia | possession of the qualities (especially mental qualities) required to do something or get something done; "danger heightened his powers of discrimination" |
| Manifestación léxica | <i>destreza, facultad, talento, inteligencia, dotes, virtud, cualidad /// soltura, oficio, tablas, elocuencia, dominio</i> | <i>aptitud, capacidad, competencia, destreza, don, facultad, talento</i> (español), <i>ability, capability, capacity, competence, faculty, gift, power, talent</i> (inglés) |

La próxima clase léxica que se combina con *derrochar* es estable (tipo I). La vinculamos a la entidad +IMAGINATION_00. Todos los miembros de la clase léxica pueden lexicalizar el concepto, y la descripción del concepto en FunGramKB aplica a la clase léxica:

Tabla 140. Comparación entre clase léxica de REDES y +IMAGINATION_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|--|--|
| Superodinado: | --- | +COGNITIVE_ATT_00 |
| Concepto: | --- | +IMAGINATION_00 |
| PS: | --- | +(e1: +BE_00 (x1: +IMAGINATION_00)Theme (x2: +COGNITIVE_ATT_00)Referent) *(e2: +IMAGINE_00 (x3)Theme (x4)Referent (f1: x1)Means) |
| Descripción: | facultades mentales relacionadas a la creatividad | the ability to form mental images of things or events; "he could still hear her in his imagination" |
| Manifestación léxica | <i>imaginación, creatividad, ingenio, fantasía, genio, inventiva</i> | <i>creatividad, imaginación</i> (español), <i>creativity, imagination</i> (inglés) |

La próxima clase, que REDES asocia a CUALIDAD, es una clase del tipo IV: no se puede vincular satisfactoriamente a ninguna entidad. El concepto CUALIDAD de REDES nos lleva al concepto terminal \$FEATURE_00 de FunGramKB, que es técnicamente un concepto superordinado de los miembros de la clase: *estilo, elegancia, clase, buen gusto, glamour, lujo, estética, fineza, finura*. Sin embargo, los miembros son cualidades muy específicas para un concepto tan amplio como \$FEATURE_00, por lo que este vínculo no resulta útil a la hora de describir la clase. Los miembros de esta clase léxica pasarán al espacio de colocaciones en la plantilla léxica de *derrochar*[+ABOUND_00]:

Tabla 141. Comparación entre clase léxica de REDES y \$FEATURE_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB (*no hay vínculo) |
|----------------------|--|---|
| Superordinado: | --- | +ATTRIBUTE_00 |
| Concepto: | CUALIDAD | \$FEATURE_00 |
| PS: | --- | +(e1: +BE_00 (x1: \$FEATURE_00)Theme (x2: +ATTRIBUTE_00)Referent) *(e2: +BE_01 (x1)Theme (x3: +IMPORTANT_00)Attribute) |
| Descripción: | cualidades, actitudes y facultades relativas a la capacidad de elegir o a la forma de comportarse o presentarse ante los demás | a prominent aspect of something: "the map showed roads and other features"; "generosity is one of his best characteristics" |
| Manifestación léxica | <i>estilo, elegancia, clase, buen gusto, glamour, lujo, estética, fineza, finura</i> | <i>característica, cualidad, propiedad, rasgo</i> (español), <i>characteristic, feature, property, quality</i> (inglés) |

Vinculamos dos clases léxicas que se combinan con *derrochar* al concepto +HUMOUR_00. La primera de ellas está asociada en REDES a CUALIDAD. Se trata de una clase mixta (tipo III), que tiene argumentos que lexicalizan el concepto +HUMOUR_00, como *gracia, encanto, sonrisa*, pero también otros que se adhieren al concepto por sentido figurado, como *simpatía, carisma, naturalidad, desparpajo, espontaneidad*. La segunda clase léxica que vinculamos con +HUMOUR_00 es una clase léxica estable (tipo I); todos sus miembros lexicalizan el concepto +HUMOUR_00: *humor, ironía, malicia, sarcasmo, socarronería*.

Tabla 142. Comparación entre clase léxica de REDES y +HUMOUR_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|--|--|
| Superordinado: | --- | +BEHAVIOUR_00 |
| Concepto: | CUALIDAD /// --- | +HUMOUR_00 |
| PS: | --- | +(e1: +BE_00 (x1: +HUMOUR_00)Theme (x2: +ATTRIBUTE_00)Referent) *((e2: +DO_00 (x3: +HUMAN_00)Theme (x4)Referent (f1: x1)Manner)(e3: +BE_01 (x4)Theme (x5: +FUNNY_00)Attribute)) |
| Descripción: | jovialidad, extroversión, desenfado, atractivo personal y cualidades análogas /// actitudes o manifestaciones humanísticas, mordaces o burlescas, a veces con intención solapada | the quality of being funny; "I fail to see the humor in it" |
| Manifestación léxica | <i>simpatía, carisma, naturalidad, desparpajo, espontaneidad, gracia, encanto, salero, sonrisa /// humor, ironía, malicia, sarcasmo, socarronería</i> | <i>humor</i> (español), <i>humour</i> (inglés) |

La próxima clase es descrita por +PSYCHOLOGICAL_ATT_00, un concepto superordinado a los argumentos de la clase (tipo II). Todos los argumentos son nociones subordinadas a la noción de atributo psicológico. Sin embargo, al igual que ocurre con la clase vinculada a \$FEATURE_00 que vimos arriba, en este caso, el vínculo no es útil para describir la clase: el concepto +PSYCHOLOGICAL_ATT_00 es muy general para una clase léxica con miembros tan específicos como: *sinceridad, amabilidad, generosidad, caridad, ternura, honradez, humanidad, pundonor, bondad, piedad, nobleza*:

Tabla 143. Comparación entre clase léxica de REDES y +PSYCHOLOGICAL_ATT_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|--|---|
| Superodinado: | --- | +ATTRIBUTE_00 |
| Concepto: | CUALIDAD | +PSYCHOLOGICAL_ATT_00 |
| PS: | --- | +(e1: +BE_00 (x1: +PSYCHOLOGICAL_ATT_00)Theme (x2: +ATTRIBUTE_00)Referent) +(e2: +BE_01 (x3: +HUMAN_00 ^ +ANIMAL_00)Theme (x1)Attribute) |
| Descripción: | cualidades del carácter, más frec. si son sobre su disposición a los demás | psychological attribute |
| Manifestación léxica | <i>sinceridad, amabilidad, generosidad, caridad, ternura, honradez, humanidad, pundonor, bondad, piedad, nobleza</i> | --- |

Vinculamos la próxima clase léxica, descrita por REDES como CUALIDAD, al concepto +ANGER_00. Se trata de una clase estable (tipo I), cuyos miembros pueden lexicalizar el concepto, y la descripción del concepto aplica a la clase léxica:

Tabla 144. Comparación entre clase léxica de REDES y +ANGER_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|---|---|
| Superodinado: | --- | +FEELING_00 |
| Concepto: | CUALIDAD | +ANGER_00 |
| PS: | --- | +(e1: +BE_00 (x1: +ANGER_00)Theme (x2: +FEELING_00)Referent) *(e2: +FEEL_00 (x3: (e3: +DISLIKE_00 (x4)Agent (x5: +HUMAN_00)Theme))Agent (x5)Theme (x1)Attribute (f1: (e4: +WANT_00 (x5)Theme (x6: (e5: +SHOUT_00 (x5)Theme (x7)Referent (x8)Goal))Referent))Result (f2: (e6: +WANT_00 (x5)Theme (x9: (e7: +HURT_00 (x5)Theme (x8)Referent))Referent))Result) |
| Descripción: | sentimientos negativos, esp. insatisfacción e ira | a feeling which makes you want to shout at someone or hurt them because they have behaved in an unfair, cruel, offensive etc way, or because you think that a situation is unfair, unacceptable etc |
| Manifestación léxica | <i>rabia, saña, furor, pesimismo, frustración, amargura</i> | <i>cabreo, enfado, enojo</i> (español), <i>anger</i> (inglés) |

La última clase léxica, descrita por REDES como SENTIMIENTO, la hemos vinculado al concepto +PRIDE_00. Se trata nuevamente de una clase estable (tipo I), cuyos miembros pueden lexicalizar el concepto; asimismo, la descripción del concepto aplica a la clase léxica:

Tabla 145. Comparación entre clase léxica de REDES y +PRIDE_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|--|---|
| Superodinado: | --- | +PLEASURE_00 |
| Concepto: | SENTIMIENTO | +PRIDE_00 |
| PS: | --- | +(e1: +BE_00 (x1: +PRIDE_00)Theme (x2: +PLEASURE_00)Referent) *((e2: +FEEL_00 (x3: (e3: past +DO_00 (x4: +HUMAN_00)Theme (x5)Referent))Agent (x4)Theme (x1)Attribute)(e4: +THINK_00 (x4)Theme (x6: (e5: +BE_01 (x5)Theme (x7: m +GOOD_00)Attribute))Referent)) |
| Descripción: | orgullo, soberbia, vanidad | a feeling that you are proud of something that you or someone connected with you has achieved |
| Manifestación léxica | <i>altanería, autoestima, vanidad, presunción,</i> | <i>orgullo</i> (español), <i>pride</i> (inglés) |

Después de considerar todas las clases léxicas con que se combina *derrochar* en su sentido de \$ABOUND_00, presentamos en la siguiente tabla las rutas conceptuales de las entidades seleccionadas. Es evidente que, en su vínculo con \$ABOUND_00, *derrochar* tiene una preferencia por los atributos psicológicos:

Tabla 146. Rutas conceptuales de conceptos vinculados a clases léxicas que se combinan con *derrochar*[\$ABOUND_00]

| Rutas conceptuales -----> -----> -----> -----> -----> | | | | | |
|---|---------------------|-------------|----------------|---------------|---------|
| #Abstract- #Attribute- +Attribute | +Psycholog_ _Att | | | | |
| | | +Ability | | | |
| | | | +Cognitive Att | +Intelligence | +Wisdom |
| | | | | +Imagination | |
| | | +Feeling | +Anger | | |
| | | | +Pride | | |
| | | | +Pleasure | | |
| | | +Behaviour | +Courage | | |
| | | | +Humour | | |
| | +Physical-Att | +Strength | | | |
| #Physical | #Process | +Occurrence | +Effort | | |

A raíz de esta evidencia, proponemos la creación de tres nuevos subconceptos: -ABOUND_PSYCHOLOGICAL_ATT, que abarca la mayoría de las preferencias conceptuales, -ABOUND_STRENGHT y -ABOUND_EFFORT. Las siguientes tablas presentan los subconceptos:

Tabla 147. Descripción de -ABOUND_PSYCHOLOGICAL_ATT (subconcepto propuesto)

| | |
|------------------------|--|
| Superordinado: | +HAVE_00 |
| MT: | (x1: +HUMAN_00^+ANIMAL_00)Theme (x2: PSYCHOLOGICAL_ATT) Referent |
| Descripción: | Have in large numbers |
| Lexicalización: | <i>rebosar, derrochar</i> (español), |

Tabla 148. Descripción de -ABOUND_STRENGHT (subconcepto propuesto)

| | |
|------------------------|---|
| Superordinado: | +HAVE_00 |
| MT: | (x1: +HUMAN_00^+ANIMAL_00)Theme (x2: STRENGHT) Referent |
| Descripción: | Have and display abundant strenght |
| Lexicalización: | <i>rebosar, derrochar</i> (español) |

Tabla 149. Descripción de -ABOUND_EFFORT (subconcepto propuesto)

| | |
|------------------------|---|
| Superordinado: | +HAVE_00 |
| MT: | (x1: +HUMAN_00^+ANIMAL_00)Theme (x2: +EFFORT_00) Referent |
| Descripción: | Have and display a lot of effort |
| Lexicalización: | <i>rebosar, derrochar</i> (español) |

2.3.1 Resumen DERROCHAR

La próxima tabla muestra el resumen de los vínculos del verbo *derrochar* en FunGramKB. La columna de la izquierda muestra el vínculo con los eventos verbales, mientras que la de la derecha muestra las colocaciones que se ubicarán en la plantilla léxica de *derrochar* en cada sentido; solo los argumentos que no están representados conceptualmente en una preferencia de selección pasan a figurar como colocaciones.

Tabla 150. Resumen: *DERROCHAR* en FunGramKB (vinculación propuesta)

| Evento conceptual | Colocaciones |
|----------------------------|---|
| \$SQUANDER_00 | <i>fortuna, dinero, riqueza; munición, agua</i> |
| +ABOUND_00 | <i>estilo, elegancia, clase, buen gusto, glamour, lujo, estética, fineza, finura</i> |
| -ABOUND_ PSYCHOLOGICAL_ATT | <i>optimismo, entusiasmo, ilusión, pasión, amor, cariño; coherencia, conocimiento, idea, criterio, saber; virtud, cualidad, soltura, oficio, tablas, elocuencia, dominio; simpatía, carisma, naturalidad, desparpajo, espontaneidad</i> |
| -ABOUND_STRENGHT | <i>carácter, vitalidad, temperamento, garra, casta, raza, fiereza</i> |
| -ABOUND_EFFORT | --- |

La tabla siguiente resume la tipología de las clases con que se combina *derrochar*, y los conceptos asignados en FunGramKB y REDES a cada clase léxica. Recordemos que las clases léxicas tipo I y II son representadas por entidades que asumen el rol de preferencias de selección conceptuales. Las clases tipo III tienen algunos miembros que no están representados por una entidad conceptual y que, por lo tanto, necesitarán figurar como colocaciones en la plantilla léxica del predicado; estos se ubican en la quinta columna de la tabla. Los miembros de las clases del tipo IV tendrán que figurar en su totalidad como colocaciones; estos se ubican también en la quinta columna:

Tabla 151. Tipología de clases léxicas que se combinan con *derrochar* con respecto a FunGramKB

| Tipo de clase | CONCEPTO en FunGramKB | CONCEPTO en REDES | Argumentos CL – paradigma semántico | Argumentos CL – paradigma discursivo y pragmático |
|---------------|-----------------------|---------------------|--|--|
| I | +MONEY_00 | (DINERO) | <i>fortuna, dinero, riqueza</i> | --- |
| I | +COURAGE_00 | CUALIDAD | <i>valor, coraje, valentía, arrojo, bravura, heroísmo</i> | --- |
| I | +EFFORT_00 | CUALIDAD | <i>voluntad, entrega, ganas, empeño</i> | --- |
| I | +IMAGINATION_00 | --- | <i>ingenio, imaginación, fantasía, creatividad, genio, inventiva</i> | --- |
| I | +ANGER_00 | CUALIDAD | <i>rabia, saña, furor, pesimismo, frustración, amargura</i> | --- |
| I | +PRIDE_00 | SENTIMIENTO | <i>altanería, autoestima, vanidad, presunción,</i> | --- |
| II | +PSYCHOLOGICAL_ATT | CUALIDAD | <i>sinceridad, amabilidad, generosidad, caridad, ternura, honradez, humanidad, pundonor, bondad, piedad, nobleza</i> | --- |
| III | +PLEASURE_00 | SENTIMIENTO | <i>alegría, euforia, efusión, satisfacción, felicidad</i> | <i>optimismo, entusiasmo, ilusión, pasión, amor, cariño</i> |
| III | +WISDOM_00 | CUALIDAD | <i>prudencia, modestia, mesura, seguridad, templanza, aplomo, temple; sensatez, sabiduría, sentido común</i> | <i>coherencia, conocimiento, idea, criterio, saber</i> |
| III | +STRENGTH_00 | CUALIDAD, FUERZA | <i>fuerza, energía, fortaleza, vigor, poderío, empuje</i> | <i>carácter, casta, vitalidad, raza, temperamento, garra, fiereza</i> |
| III /// II | +ABILITY_00 | VIRTUD /// CUALIDAD | <i>destreza, facultad, talento, dote; soltura, oficio, tablas, dominio, elocuencia</i> | <i>inteligencia, virtud, cualidad,</i> |
| III | +HUMOUR_00 | | <i>gracia, encanto, salero, sonrisa; ironía, malicia, sarcasmo, socarronería, humor</i> | <i>simpatía, carisma, naturalidad, desparpajo, espontaneidad</i> |
| IV | --- | (RECURSO) | --- | <i>munición, agua</i> |
| IV | --- | CUALIDAD | --- | <i>estilo, elegancia, clase, buen gusto, glamour, lujo, estética, fineza, finura</i> |

2.4 MALOGRAR

REDES no distingue ningún sentido para *malograr*. Nosotros proponemos dos: ‘malgastar’ y ‘no llegar al término deseado, fracasar’. La tabla siguiente presenta los datos que provee REDES:

Tabla152. Sentidos y combinatoria léxica de *malograr*

| Sentido del verbo en REDES | CONCEPTO asignado a la CL en REDES | Descriptor de la CL en REDES | Argumentos miembros de la CL en REDES |
|--|------------------------------------|--|--|
| ‘malgastar’ | FUTURO | <i>futuro</i> y coyuntura favorable | <i>futuro, oportunidad, posibilidad</i> |
| | FACULTAD | facultades físicas o mentales | <i>talento, capacidad, inteligencia, voz</i> |
| ‘no llegar al término deseado, fracasar’ | --- | sucesos o acontecimientos | <i>viaje, estreno, operación, reforma</i> |
| | --- | otras nociones como resultado de un proyecto | <i>el libro, la película</i> |
| | PROYECTO | susts de naturaleza prospectiva que designan lo que se pretende conseguir, los medios para conseguirlo o la confianza que se deposita en ellos | <i>aspiración, objetivo, expectativa, esperanza, planteamiento</i> |
| | ESFUERZO | trabajo o afán, y diversas capacidades que se destinan a la consecución de un objetivo | <i>esfuerzo, empeño, energía</i> |
| | CURSO | curso o trayectoria | <i>carrera, vida, trayectoria</i> |
| | ÉXITO | resultado, generalmente favorable | <i>victoria, resultado</i> |

(a) *malograr* ‘malgastar’

El sentido de ‘malgastar’ lo hemos vinculado al evento conceptual +WASTE_00, cuya descripción en FunGramKB repetimos en la tabla siguiente:

Tabla 153. Descripción de +WASTE_00 en FunGramKB

| | |
|-----------------|--|
| Hiperónimo: | +USE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2)Referent |
| PS: | +(e1: +USE_00 (x1)Theme (x2)Referent (f1: +CARELESS_00)Manner) |
| Descripción: | use inefficiently or inappropriately; "waste heat" |
| Lexicalización: | <i>malgastar</i> (inglés), <i>waste</i> (español) |

El verbo *malograr* adquiere este sentido en combinación con dos clases léxicas. La primera es una clase mixta (tipo III), que vinculamos a la entidad +FUTURE_00. El sustantivo *futuro* lexicaliza este concepto, mientras que *oportunidad* y *posibilidad* advienen a él de manera figurada:

Tabla 154. Comparación entre clase léxica de REDES y +FUTURE_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|------------------------|---|---|
| Superodinado: | --- | +PERIOD_00 |
| Concepto: | FUTURO | +FUTURE_00 |
| PS: | --- | +(e1: +BE_00 (x1: +FUTURE_00)Theme (x2: +PERIOD_00)Referent) +(e2: fut +DO_00 (x3)Theme (x4)Referent (f1: x1)Time) |
| Descripción semántica: | <i>futuro</i> y coyuntura favorable | the time yet to come |
| Manifestación léxica | <i>futuro, oportunidad, posibilidad</i> | <i>futuro, mañana, porvenir</i> (español), <i>future, tomorrow</i> (inglés) |

La segunda clase léxica con que se combina *malograr*[+WASTE_00] la vinculamos con la entidad +ABILITY_00. Es una clase mixta (tipo III): los miembros *talento* y *capacidad* lexicalizan a +ABILITY_00, mientras que *inteligencia* lexicaliza a +INTELLIGENCE, que es subordinado de +ABILITY_00. Por su parte, *voz* lexicaliza, en su sentido recto, a +VOICE_00, que es un atributo físico, con otra ruta conceptual. En *voz* tenemos un ejemplo de una extensión semántica normal –de la palabra *voz* a la noción de ‘habilidad’, en una oración como “una *voz* brillante y armoniosa, pero *malograda* por falta de cuidados”–, que, es sin embargo, muy difícil de codificar en una ontología conceptual.

Tabla 155. Comparación de clase léxica de REDES y +ABILITY_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|------------------------|--|---|
| Superodinado: | | +PSYCHOLOGICAL_ATT_00 |
| Concepto: | FACULTAD | +ABILITY_00 |
| PS: | | +(e1: +BE_00 (x1: +ABILITY_00)Theme (x2: +PSYCHOLOGICAL_ATT_00) Referent) *(e2: pos +DO_00 (x3: +HUMAN_00 ^ +ANIMAL_00)Theme (x4)Referent (f1: x1)Means) |
| Descripción semántica: | facultades físicas o mentales | possession of the qualities (especially mental qualities) required to do something or get something done; "danger heightened his powers of discrimination" |
| Manifestación léxica | <i>talento, capacidad, inteligencia, voz</i> | <i>aptitud, capacidad, competencia, destreza, don, facultad, talento</i> (español), <i>ability, capability, capacity, competence, faculty, gift, power, talent</i> (inglés) |

A partir de estas dos selecciones léxicas, proponemos dos subconceptos: -WASTE_FUTURE y -WASTE_ABILITY.

Tabla 156. Descripción de -WASTE_FUTURE en FunGramKB (subconcepto propuesto)

| | |
|-----------------|--|
| Hiperónimo: | +USE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2: +FUTURE_00)Referent |
| Descripción: | use future time inefficiently or inappropriately; "waste future" |
| Lexicalización: | <i>malograr</i> (español) |

Tabla 157. Descripción de -WASTE_ABILITY en FunGramKB (subconcepto propuesto)

| | |
|-----------------|---|
| Hiperónimo: | +USE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2: +ABILITY_00)Referent |
| Descripción: | use tabilities inefficiently or inappropriately; "waste your talents" |
| Lexicalización: | <i>malograr</i> (español) |

En la plantilla léxica de *malograr*[-WASTE_FUTURE] se añaden como colocaciones *oportunidad* y *posibilidad*, porque no pueden ser representadas por +FUTURE_00, y en la de -WASTE_ABILITY se añade *voz* como colocación.

(b) *malograr* ‘no llegar al término deseado, fracasar’

El segundo sentido del verbo *malograr* es ‘no llegar al término deseado, fracasar’. Este sentido lo vinculamos con el evento conceptual +FAIL_00, cuya descripción en FunGramKB reproducimos a continuación:

Tabla 158. Descripción de +FAIL_00 en FunGramKB

| | |
|-----------------|---|
| Superordinado: | +TRY_00 |
| MT: | (x1: +HUMAN_00 ^ +ANIMAL_00)Theme (x2)Referent |
| PS: | +(e1: +TRY_00 (x1)Theme (x2)Referent (f1: (e2: n +DO_00 (x1)Theme (x2)Referent))Result) |
| Descripción | be unsuccessful; "Where do today's public schools fail?"; "The attempt to rescue the hostages failed miserably" |
| Lexicalización: | <i>fallar fracasar, malograr</i> (español), <i>fail</i> (inglés) |

Este es el sentido dominante para *malograr*; ocurre con la mayoría de las clases léxicas con que se combina este verbo. La primera clase léxica es del tipo IV: no se puede vincular satisfactoriamente con una entidad en FunGramKB. REDES la describe como

‘sucesos o acontecimientos’. La entidad más cercana es +OCCURRENCE_00, pero resulta demasiado amplia, especialmente considerando sus hipónimos, que incluyen nociones negativas que no se podrían *malograr*, como +DAMAGE_00, +VIOLENCE_00, +DECREASE_00.

Tabla 159. Comparación entre clase léxica de REDES y +OCCURRENCE_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB (*no hay vínculo) |
|------------------------|---|---|
| Superodinado: | --- | #PROCESS |
| Concepto: | --- | +OCCURRENCE_00 |
| PS: | --- | +(e1: +DO_00 (x1)Theme (x2)Referent) |
| Descripción semántica: | sucesos o acontecimientos | something that happens |
| Manifestación léxica: | <i>viaje, estreno, operación, reforma</i> | <i>acción, acontecimiento, actividad, acto, evento, incidente, lance, proceso, suceso</i> (español), <i>act, action, activity, event, happening, occurrence, process</i> (inglés) |

La próxima clase léxica con que se combina *malograr*[+FAIL_00] es también del tipo IV: no se puede vincular a entidad alguna. REDES la describe como ‘otras nociones como resultado de un proyecto’, por ejemplo: *el libro, la película*. Estos dos argumentos de *malograr* tienen en sus sentidos rectos vínculos a entidades en la ontología: +BOOK_00 y +MOVIE_00, respectivamente, pero asumen el sentido de ‘resultado’ en combinación con *malograr*. Este es un ejemplo de cómo un aspecto del significado –más que un significado nuevo– se activa según el contexto combinatorio.

La próxima clase léxica se vincula a la entidad +PLAN_00, a raíz de su asociación con el concepto PROYECTO y de su descripción semántica en REDES (‘lo que se pretende conseguir, los medios para conseguirlo’). Sin embargo, es una clase léxica mixta (tipo III), porque ninguna de las unidades léxicas lexicalizan este concepto; de hecho, los argumentos –*aspiración, objetivo, expectativa, esperanza, planteamiento*– se corresponden mejor con el concepto +DESIRE_00. Sin embargo, es más natural que se *malogre* un ‘plan’ a que se *malogre* un ‘sentimiento’, por lo que mantenemos el vínculo con +PLAN_00:

Tabla 160. Comparación entre clase léxica de REDES y +PLAN_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|------------------------|--|--|
| Superordinado: | | +INFORMATION_00 |
| Concepto: | PROYECTO | +PLAN_00 |
| PS: | | +(e1: +BE_00 (x1: +PLAN_00)Theme (x2: +INFORMATION_00)Referent) *(e2: +DO_00 (x3)Theme (x4)Referent (f1: x1)Manner) |
| Descripción semántica: | susts de naturaleza prospectiva que designan lo que se pretende conseguir, los medios para conseguirlo o la confianza que se deposita en ellos | a series of steps to be carried out or goals to be accomplished; "they drew up a six-step plan"; "they discussed plans for a new bond issue" |
| Manifestación léxica | <i>aspiración, objetivo, expectativa, esperanza, planteamiento</i> | <i>estrategia, plan, programa, proyecto (español), plan, programme, project, strategy (inglés)</i> |

La próxima clase léxica es una clase mixta (tipo III), pero bastante estable. Se vincula con la entidad +EFFORT_00, que esfuerzo y empeño lexicalizan. El otro argumento – *energía*– se vincula ya a +ENERGY_00 y a +STRENGTH_00, pero puede adquirir el sentido de ‘esfuerzo’ de manera figurada.

Tabla 161. Comparación entre clase léxica de REDES y +EFFORT_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|------------------------|--|--|
| Superordinado: | | +OCCURRENCE_00 |
| Concepto: | ESFUERZO | +EFFORT_00 |
| PS: | | +(e1: +BE_00 (x1: +EFFORT_00)Theme (x2: +OCCURRENCE_00)Referent) +(e2: +TRY_00 (x3: +HUMAN_00 ^ +ANIMAL_00)Theme (x4)Referent (f1: +MUCH_00)Quantity (f2: x1)Scene) |
| Descripción semántica: | trabajo o afán, y diversas capacidades que se destinan a la consecución de un objetivo | earnest and conscientious activity intended to do or accomplish something: "made an effort to cover all the reading material"; "wished him luck in his endeavor"; "she gave it a good try" |
| Manifestación léxica | <i>esfuerzo, empeño, energía</i> | <i>afán, ahínco, ánimo, animosidad, brío, empeño, esfuerzo, intento, sudor (español), attempt, effort, endeavour, try (inglés)</i> |

La próxima clase pertenece al tipo IV: no se puede vincular con ninguna entidad. Su descriptor es ‘curso o trayectoria’, se asocia con el concepto CURSO en REDES y sus miembros son *carrera, vida, trayectoria*. El problema aquí es que la noción de ‘curso’ se puede asignar a prácticamente cualquier palabra que contenga en su significado aspectos de duración de tiempo. Nuevamente, se trata de un aspecto de sentido –y no de un sentido nuevo– que se activa en combinación con el verbo *malograr*.

La última clase con que se combina *malograr*[+FAIL_00] es una clase mixta (tipo III), pero bastante estable. El descriptor semántico hace claro que se trata de un resultado exitoso, lo cual está presente en la descripción semántica de +VICTORY_00, que es subordinado de +OCCURRENCE_00, a diferencia, por ejemplo, de +SUCCESS_00, que está ubicado bajo +STATE_00. El argumento *victoria* lexicaliza a +VICTORY_00, mientras que *resultado* adquiere ese valor en combinación con *malograr* (“un sensacionalismo que en más de una ocasión ha *malogrado* el *resultado* final de buena parte de sus películas”).

Tabla 162. Comparación entre clase léxica de REDES y +VICTORY_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|------------------------|-----------------------------------|--|
| Superodinado: | --- | +OCCURRENCE_00 |
| Concepto: | ÉXITO | +VICTORY_00 |
| PS: | --- | +(e1: +BE_00 (x1: +VICTORY_00)Theme (x2: +OCCURRENCE_00)Referent) +(e2: +WIN_00 (x3: +HUMAN_00 ^ +TEAM_00)Theme (x4)Referent (f1: x1)Scene) |
| Descripción semántica: | resultado, generalmente favorable | a successful ending of a struggle or contest; "the general always gets credit for his army's victory" |
| Manifestación léxica | <i>victoria, resultado</i> | <i>triumfo, victoria</i> (español), <i>triumph, victory, win</i> (inglés) |

La tabla siguiente muestra las rutas conceptuales de las clases léxicas con que se combina *malograr* en su sentido dominante: +FAIL_00:

Tabla 163. Rutas conceptuales de conceptos vinculados a clases léxicas que se combinan con *malograr*[FAIL+ 00]

| Rutas conceptuales -----> -----> -----, -----> | | | | |
|--|--------------|-------------|--------------|-------|
| #Abstract | #Proposition | +Thought | +Information | +Plan |
| #Physical | #Process | +Occurrence | +Victory | |
| | | | +Effort | |

A raíz de la evidencia de combinatoria léxica de *malograr*[+FAIL_00], se pueden proponer los siguientes subconceptos: -FAIL_PLAN, -FAIL_EFFORT y -FAIL_VICTORY. Aunque +VICTORY_00 y +EFFORT_00 son subordinados de +OCCURRENCE_00, no usamos esta entidad (+OCCURRENCE_00) para crear un subconcepto que los abarque a ambos, porque incluye demasiados otros conceptos subordinados:

Tabla 164. Descripción de -FAIL_PLAN (subconcepto propuesto)

| | |
|-----------------|---|
| Superordinado: | +TRY_00 |
| MT: | (x1: +HUMAN_00)Theme (x2: +PLAN_00)Referent |
| Descripción | be unsuccessful at a plan; “the plan failed” |
| Lexicalización: | <i>malograr</i> (español), <i>fail</i> (inglés) |

Tabla 165. Descripción de -FAIL_EFFORT (subconcepto propuesto)

| | |
|-----------------|--|
| Superordinado: | +TRY_00 |
| MT: | (x1: +HUMAN_00)Theme (x2: +EFFORT_00)Referent |
| Descripción | be unsuccessful with an effort; “our attempt failed” |
| Lexicalización: | <i>malograr</i> (español), <i>fail</i> , <i>to be ruined</i> , <i>come to nothing</i> (inglés) |

Tabla 166. Descripción de -FAIL_VICTORY (subconcepto propuesto)

| | |
|-----------------|--|
| Superordinado: | +TRY_00 |
| MT: | (x1: +HUMAN_00)Theme (x2: +VICTORY_00)Referent |
| Descripción | be unsuccessful with a victory; “our triumph was ruined” |
| Lexicalización: | <i>malograr</i> (español), <i>to be ruined</i> , <i>come to nothing</i> (inglés) |

La plantilla léxica de *malograr*[-FAIL_PLAN] tiene las colocaciones *aspiración*, *objetivo*, *expectativa*, *esperanza*, *planteamiento*, que no se pueden vincular a +PLAN_00; la de *malograr*[-FAIL_EFFORT] tiene la colocación *energía*; y *malograr*[-FAIL_VICTORY] tiene la colocación *resultado*. Las demás clases léxicas no se pudieron vincular a entidades, por lo que sus miembros pasan a ser colocaciones en la entrada léxica de *malograr*[+FAIL]: *viaje*, *estreno*, *operación*, *reforma*; *el libro*, *la película*; *carrera*, *vida*, *trayectoria*.

2.4.1 Resumen MALOGRAR

La próxima tabla resume el vínculo conceptual de *malograr* con diferentes eventos, en la primera columna, y las colocaciones que se ubican la plantilla léxica de *malograr* en cada uno de los vínculos. Recordemos que las colocaciones :

Tabla 167. Resumen: MALOGRAR en FunGramKB (vinculación propuesta)

| Evento conceptual | Colocaciones |
|-------------------|--|
| -WASTE_FUTURE | <i>oportunidad, posibilidad</i> |
| -WASTE_ABILITY | <i>voz</i> |
| +FAIL_00 | <i>viaje, estreno, operación, reforma; el libro, la película; carrera, vida, trayectoria</i> |
| -FAIL_PLAN | <i>aspiración, objetivo, expectativa, esperanza, planteamiento</i> |
| -FAIL_EFFORT | <i>energía</i> |
| -FAIL_VICTORY | <i>resultado</i> |

La próxima tabla muestra la tipología de las clases léxicas con que se combina *malograr*. Llama la atención que, con este verbo, no se combinaron clases estables (tipo I) o clases descritas por superordinados (tipo II); todas fueron mixtas (tipo III) o clases que no se pudieron vincular (tipo IV). Por otro lado, queda claro que el verbo *malograr*, en el sentido de +FAIL_00, activa en todos los argumentos con que se combina el aspecto de sentido de ‘resultado final de un proceso, potencialmente positivo’:

Tabla 168. Tipología de clases léxicas que se combinan con *malograr* con respecto a FunGramKB

| Tipo de clase | CONCEPTO en FunGramKB | CONCEPTO en REDES | Argumentos CL - Paradigma semántico | Argumentos CL - Paradigma discursivo o pragmático |
|---------------|-----------------------|-------------------|---|--|
| III | +FUTURE_00 | FUTURO | <i>futuro</i> | <i>oportunidad, posibilidad</i> |
| III | +ABILITY_00 | FACULTAD | <i>talento, capacidad, inteligencia</i> | <i>voz</i> |
| III | +PLAN_00 | PROYECTO | --- | <i>aspiración, objetivo, expectativa, esperanza, planteamiento</i> |
| III | +EFFORT_00 | ESFUERZO | <i>esfuerzo, empeño,</i> | <i>energía</i> |
| III | +VICTORY_00 | ÉXITO | <i>victoria</i> | <i>resultado</i> |
| IV | --- | --- | --- | <i>viaje, estreno, operación, reforma</i> |
| IV | --- | --- | --- | <i>el libro, la película</i> |
| IV | --- | CURSO | --- | <i>carrera, vida, trayectoria</i> |

2.5 MALGASTAR

Malgastar es uno de cinco verbos monosémicos que comparten el vínculo único con +WASTE_00. El verbo *malgastar* significa, a grandes rasgos, ‘desperdiciar, usar mal o

descuidadamente'. Lo hemos vinculado con el evento conceptual +WASTE_00, cuya descripción en FunGramKB repetimos en la tabla siguiente:

Tabla 169. Descripción de +WASTE_00 en FunGramKB

| | |
|-----------------|--|
| Hiperónimo: | +USE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2)Referent |
| PS: | +(e1: +USE_00 (x1)Theme (x2)Referent (f1: +CARELESS_00)Manner) |
| Descripción: | use inefficiently or inappropriately; "waste heat" |
| Lexicalización: | <i>malgastar</i> (inglés), <i>waste</i> (español) |

Aunque *malgastar* solo tiene un sentido, su combinatoria sistemática incluye ocho clases léxicas: dos clases abiertas y seis cerradas. La tabla siguiente muestra las clases léxicas de REDES. Estas restricciones no afectan su significado conceptual; el vínculo con +WASTE_00, cuyo Referente no está especificado semánticamente, se mantiene en combinación con todas las clases léxicas.

Tabla 170. Combinatoria léxica de *malgastar*, en REDES

| CONCEPTO asignado a la CL en REDES | Descriptor de la CL en REDES | Argumentos miembros de la CL en REDES |
|------------------------------------|---|--|
| RECURSO | magnitudes, bienes materiales, recursos | <i>dinero, agua, corcho, petróleo, gasolina, papel</i> |
| TIEMPO | unidades temporales, incluidas las que se refieren a la existencia | <i>vida, juventud, día</i> |
| FUERZA | fuerza o energía, y sus efectos, literal y figurado | <i>fuerza, energía, esfuerzo, salud</i> |
| CAPACIDAD, FACULTAD | facultades, aptitudes o cualidades humanas, esp. intelectivas; estados de reconocimiento personal | <i>talento, capacidad, don, cualidad, generosidad, valor, credibilidad, prestigio, experiencia</i> |
| POSIBILIDAD | oportunidad u ocasión, ventaja o circunstancia favorable, frec. deportivo | <i>oportunidad, ocasión, balón, ventaja, punto, opción, privilegio, posibilidad</i> |
| ÉXITO | resultado feliz de algo | <i>éxito, victoria, triunfo</i> |
| INFORMACIÓN | unidades, materias y recursos informativos, expresivos o artísticos | <i>saliva, línea, palabra, página, óleo, tinta, cinta</i> |
| ARMA | munición, acciones en las que se usa o formas en que se almacena | <i>disparo, munición, cartuchera, bala</i> |

A continuación examinamos los posibles vínculos de estas clases léxicas con entidades en FunGramKB. La primera clase léxica pertenece al tipo IV: no se puede vincular

a ninguna entidad satisfactoriamente. Hemos visto este caso anteriormente: la noción de RECURSO de REDES no tiene equivalente en una ontología conceptual, por tratarse de una valoración pragmática que podría aplicar a una multiplicidad de cosas. El único concepto que podría dar cabida a los miembros de esta clase, como concepto superordinado, es +SUBSTANCE_00, pero estaría respondiendo al significado recto de las palabras. La noción de ‘recurso’ la asignan los hablantes a estas sustancias, pues en el plano pragmático, lo son:

Tabla 171. Descripción de clase léxica que se combina con *malgastar*

| | |
|-------------------------------|--|
| Tipo de información | REDES |
| Concepto: | RECURSO |
| Descripción semántica: | magnitudes, bienes materiales, recursos |
| Manifestación léxica | <i>dinero, corcho, agua, petróleo, gasolina, papel</i> |

La segunda clase léxica con que se combina *malgastar* la vinculamos con +PERIOD_00. Esta es una clase descrita por un concepto superordinado (tipo II), pues algunos de sus miembros lexicalizan conceptos subordinados: *día*[+DAY_00], *juventud*[\$ADOLESCENCE]. El concepto es adecuado, aunque más amplio que la clase de REDES; también se pierde la especificación que hace el descriptor: ‘incluidas las que se refieren a la existencia’.

Tabla 172. Comparación entre clase léxica de REDES y +PERIOD_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|-------------------------------|--|---|
| Superordinado: | --- | +TIME_00 |
| Concepto: | TIEMPO | +PERIOD_00 |
| PS: | --- | +(e1: +BE_00 (x1: +PERIOD_00)Theme (x2: +TIME_00)Referent) |
| Descripción semántica: | unidades temporales, incluidas las que se refieren a la existencia | a particular length of time with a beginning and an end |
| Manifestación léxica | <i>día, juventud, vida</i> | <i>edad, época, era, intervalo, periodo, tiempo</i> (español), <i>age, epoch, era, interval, period, span, stretch, term, time</i> (inglés) |

Hemos vinculado la tercera clase con que se combina *malgastar*[+WASTE_00] a la entidad +STRENGTH_00. Se trata de una clase mixta (tipo III), en la que algunos argumentos lexicalizan al concepto y otros adquieren ese sentido de manera figurada en combinación con *malgastar*. De los miembros de la clase léxica, solo se pueden añadir como lexicalizaciones a +STRENGTH_00 *fuerza* (que ya aparece) y *energía*, que se puede vincular

tanto a +STRENGHT_00 como a +ENERGY_00. Por su parte, *esfuerzo* y *salud* se usan en sentido figurado y no podrían ubicarse aquí; ambos tienen entradas conceptuales en FunGramKB bajo sus sentidos rectos: +EFFORT_00 y +HEALTH_00, respectivamente. Estos son, sin duda, conceptos que también pueden ser seleccionados por *malgastar*[+WASTE_00], pero lo que interesa es capturar la generalización de REDES de que se puede *malgastar* todo lo que se conciba como ‘fuerza’, literal o figuradamente:

Tabla 173. Comparación entre clase léxica de REDES y +STRENGHT_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|------------------------|---|--|
| Superordinado: | --- | +PHYSICAL_ATT_00 |
| Concepto: | FUERZA | +STRENGHT_00 |
| PS: | --- | +(e1: +BE_00 (x1: +STRENGTH_00)Theme (x2: +PHYSICAL_ATT_00)Referent) *(e2: +BE_01 (x3: +HUMAN_00 ^ +ANIMAL_00)Theme (x1)Attribute) |
| Descripción semántica: | fuerza o energía, y sus efectos, literal y figurado | the property of being physically or mentally strong; "fatigue sapped his strenght" |
| Manifestación léxica | <i>fuerza, energía, esfuerzo, salud</i> | <i>fuerza</i> (español), <i>force, strenght</i> (ing) |

Vinculamos la cuarta clase léxica con que se combina *malgastar*[+WASTE_00] con la entidad +ABILITY_00. Esta es una clase mixta (tipo III), que tiene miembros que lexicalizan el concepto (*talento, capacidad, don*) y otros que advienen a él de manera figurada (*cualidad, valor, generosidad, credibilidad, prestigio, experiencia*).

Tabla 174. Comparación entre clase léxica de REDES y +ABILITY_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|------------------------|--|---|
| Superordinado: | --- | +PSYCHOLOGICAL_ATT_00 |
| Concepto: | CAPACIDAD, FACULTAD | +ABILITY_00 |
| PS: | --- | +(e1: +BE_00 (x1: +ABILITY_00)Theme (x2: +PSYCHOLOGICAL_ATT_00) Referent) *(e2: pos +DO_00 (x3: +HUMAN_00 ^ +ANIMAL_00)Theme (x4)Referent (f1: x1)Means) |
| Descripción semántica: | facultades, aptitudes o cualidades humanas, esp. intelectivas; estados de reconocimiento personal | possession of the qualities (especially mental qualities) required to do something or get something done; "danger heightened his powers of discrimination" |
| Manifestación léxica: | <i>talento, capacidad, don, cualidad, valor, generosidad, credibilidad, prestigio, experiencia</i> | <i>aptitud, capacidad, competencia, destreza, don, facultad, talento</i> (español), <i>ability, capability, capacity, competence, faculty, gift, power, talent</i> (inglés) |

La quinta clase léxica con que se combina *malgastar*[+WASTE_00] pertenece al tipo IV: no se puede vincular a entidad alguna. Esta clase léxica tiene el descriptor ‘oportunidad u ocasión, ventaja o circunstancia favorable, frecuentemente deportiva’ y se asocia al concepto POSIBILIDAD en REDES. Incluye argumentos como *oportunidad, ocasión, opción, posibilidad*, que pueden pertenecer a esa clase semántica por sus rasgos inherentes, pero también con otros como *balón, ventaja, punto*, que no pueden ser asimilados a ese paradigma por criterios semánticos. Estos últimos se convierten en un tipo de ‘oportunidad u ocasión’ al combinarse con *malgastar*. El paradigma semántico se amplía para incluir un paradigma pragmático o discursivo es considerada una “operación atributiva implícita” (Bosque, 2004: CXXXIII). El hablante, o el idioma, les asigna a palabras como *balón, ventaja, punto* el rasgo que define el paradigma semántico. Ni la ontología de FunGramKB –ni prácticamente ningún diccionario– puede dar cabida a este tipo de extensión.

La próxima clase se vincula a la entidad +VICTORY_00, y es una clase estable (tipo I). Este caso demuestra que REDES y FunGramKB a veces poseen un nivel equivalente de granularidad. A partir del descriptor: ‘resultado feliz de algo’ y de los argumentos: *éxito, victoria, triunfo*, surgen en FunGramKB dos posibles conceptos: +SUCCESS_00 (‘when you achieve what you want or intend’) y +VICTORY (‘a successful ending of a struggle or contest’). Al mirar la clase léxica cuidadosamente, comprobamos que no se trata del éxito como estado, sino del éxito como resultado, lo que se refleja en los superordinados STATE_00 y +OCCURRENCE, de +SUCCESS_00 y +VICTORY, respectivamente:

Tabla 175. Comparación entre clase léxica de REDES y +VICTORY_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|------------------------|---------------------------------|--|
| Superordinado: | --- | +OCCURRENCE_00 |
| Concepto: | ÉXITO | +VICTORY_00 |
| PS: | --- | +(e1: +BE_00 (x1: +VICTORY_00)Theme (x2: +OCCURRENCE_00)Referent) +(e2: +WIN_00 (x3: +HUMAN_00 ^ +TEAM_00)Theme (x4)Referent (f1: x1)Scene) |
| Descripción semántica: | resultado feliz de algo | a successful ending of a struggle or contest; "the general always gets credit for his army's victory" |
| Manifestación léxica: | <i>éxito, victoria, triunfo</i> | <i>triunfo, victoria</i> (español), <i>triumph, victory, win</i> (inglés) |

La próxima clase léxica con que se combina *malgastar*[+WASTE_00] la vinculamos con +INFORMATION_OBJECT_00. Esta es una clase mixta (tipo III): algunos miembros, como *palabra*[+WORD_00] y *línea*[LINE_00], son subordinados de +INFORMATION_OBJECT_00 (a través de +SPEECH_00 y +MARK_00, respectivamente). Pero otros

adquieren la valoración de +INFORMATION_OBJECT_00 de manera figurada, en este caso a través de procesos de metonimia: *saliva, página, línea, óleo, tinta, cinta*.

Tabla 176. Comparación entre clase léxica de REDES y +INFORMATION_OBJECT_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|------------------------|---|--|
| Superodinado: | --- | #SELF_CONNECTED_OBJECT |
| Concepto: | INFORMACIÓN | +INFORMATION_OBJECT_00 |
| PS: | --- | +(e1: +COMPRISE_00 (x1: +INFORMATION_OBJECT_00)Theme (x2: +INFORMATION_00)Referent) |
| Descripción semántica: | unidades, materias y recursos informativos, expresivos o artísticos | Something that provides information |
| Manifestación léxica | <i>saliva, línea, palabra, página, óleo, tinta, cinta</i> | <i>indicio, señal, signo, símbolo</i> (español), <i>sign, signal, symbol, token</i> (inglés) |

La última clase léxica con que se combina *malgastar*[+WASTE_00] es una clase mixta (tipo III). Dos de los argumentos *-bala, munición-* pueden lexicalizar a +BULLET_00, pero *disparo* y *cartuchera* no. En este caso, *cartuchera* y *disparo* pasan a significar ‘munición’ por un proceso de metonimia. Cabe señalar que la clase léxica de REDES incluye tres referentes diferentes pero relacionados: ‘munición’ (*bala, munición*), ‘acciones en las que se usa’ (*disparo*) y ‘sitios en que se almacena’ (*cartuchera*), aunque los agrupa bajo el concepto ARMA:

Tabla 177. Comparación entre clase léxica de REDES y +BULLET_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|------------------------|--|------------------------------|
| Superodinado: | --- | +PROJECTILE_00 |
| Concepto: | ARMA | +BULLET_00 |
| PS: | --- | --- |
| Descripción semántica: | munición, acciones en las que se usa o formas en que se almacena | --- |
| Manifestación léxica | <i>bala, disparo, munición, cartuchera</i> | <i>bullet, slug</i> (inglés) |

La próxima tabla muestra las rutas conceptuales de las clases léxicas con que se combina *malgastar*[+WASTE_00]. No hay muchas muchas rutas similares; más bien, hay bastante diversidad en la selección de *malgastar*:

Tabla 178. Rutas conceptuales de conceptos vinculados a clases léxicas que se combinan con *malgastar*[WASTE+_00]

| Rutas conceptuales -----> -----> -----> -----> -----> | | | | | |
|--|---------------------------------|-------------|-----------------------|------------------|----------------|
| #Abstract | #Quantity | #Time | +Time | +Period | |
| #Abstract | #Attribute | +Attribute | +Physical_Att | +Strength | |
| | | | +Psychological_Att | +Ability | |
| #Physical | #Process | +Occurrence | +Victory | | |
| #Physical - #Object - #Self_ Connected_ Object | +Information _Object | | | | |
| | +Artificial_ Object | +Substance | +Solid --> +Weapon | +Projectile | +Bullet |

Las preferencias de selección son diversas y no permiten ser agrupadas en un número reducido de subconceptos. Proponemos pues la creación de los siguientes subconceptos de +WASTE_00, basados en las diferentes preferencias de selección constatadas: -WASTE_PERIOD, -WASTE_STRENGTH, -WASTE_ABILITY, WASTE_VICTORY, WASTE_INFORMATION_OBJECT, WASTE_BULLET, que se describen a continuación

Tabla 179. Descripción de -WASTE_PERIOD (subconcepto propuesto)

| | |
|-----------------|--|
| Hiperónimo: | +USE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2:+PERIOD_00)Referent |
| Descripción: | use time inefficiently or inappropriately; "waste an hour" |
| Lexicalización: | <i>malgastar</i> (español), <i>waste</i> (inglés) |

Tabla 180. Descripción de -WASTE_STRENGTH (subconcepto propuesto)

| | |
|-----------------|--|
| Hiperónimo: | +USE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2:+STRENGTH_00)Referent |
| Descripción: | use strenght inefficiently or inappropriately; "waste your energy" |
| Lexicalización: | <i>malgastar</i> (español), <i>waste</i> (inglés) |

Tabla 181. Descripción de -WASTE_ABILITY (subconcepto propuesto)

| | |
|-----------------|---|
| Hiperónimo: | +USE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2:+ABILITY_00)Referent |
| Descripción: | use abilities inefficiently or inappropriately; "waste ones talent" |
| Lexicalización: | <i>malgastar</i> (español), <i>waste</i> (inglés) |

Tabla 182. Descripción de -WASTE_VICTORY (subconcepto propuesto)

| | |
|-----------------|---|
| Hiperónimo: | +USE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2:+ABILITY_00)Referent |
| Descripción: | use achievements inefficiently or inappropriately; "waste your accomplishments" |
| Lexicalización: | <i>malgastar</i> (español), <i>waste</i> (inglés) |

Tabla 183. Descripción de -WASTE_INFORMATION_OBJECT (subconcepto propuesto)

| | |
|-----------------|--|
| Hiperónimo: | +USE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2:+INFORMATION_OBJECT_00) Referent |
| Descripción: | use information inefficiently or inappropriately; "waste the data" |
| Lexicalización: | <i>malgastar</i> (español), <i>waste</i> (inglés) |

Tabla 184. Descripción de -WASTE_BULLET (subconcepto propuesto)

| | |
|-----------------|---|
| Hiperónimo: | +USE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2:+BULLET_00) Referent |
| Descripción: | use ammunitions inefficiently or inappropriately; "waste bullets" |
| Lexicalización: | <i>malgastar</i> (español), <i>waste</i> (inglés) |

2.5.1 Resumen MALGASTAR

La próxima tabla muestra el resumen de los vínculos del verbo *malgastar* en FunGramKB. La columna de la izquierda muestra el vínculo con un evento verbal, mientras que la de la derecha muestra las colocaciones que se ubicarán en la plantilla léxica de *malgastar* en ese sentido; recordemos que solo los argumentos que no están representados conceptualmente en una preferencia de selección pasan a figurar como colocaciones.

Tabla 185. Resumen: MALGASTAR en FunGramKB (vinculación propuesta)

| Evento conceptual | Colocaciones |
|------------------------------|--|
| +WASTE_00 | <i>dinero, corcho, agua, petróleo, gasolina, papel; oportunidad, ocasión, opción, posibilidad, balón, ventaja, punto</i> |
| -WASTE_PERIOD | --- |
| -WASTE_STRENGTH | <i>esfuerzo, salud</i> |
| -WASTE_ABILITY | <i>cualidad, valor, generosidad, credibilidad, prestigio, experiencia</i> |
| -WASTE_VICTORY_00 | --- |
| -WASTE_INFORMATION_OBJECT_00 | <i>saliva, óleo, tinta, cinta</i> |
| -WASTE_BULLET | <i>disparo, cartuchera</i> |

En la tabla siguiente, se muestra la vinculación con una entidad de FunGramKB de cada clase léxica con que se combina *malgastar*[+WASTE_00], y su clasificación por tipos.

Tabla 186. Tipología de clases léxicas que se combinan con *malgastar* con respecto a FunGramKB

| Tipo de clase | CONCEPTO en FunGramKB | CONCEPTO en REDES | Argumentos CL – paradigma semántico | Argumentos CL – paradigma discursivo pragmático |
|---------------|------------------------|---------------------|-------------------------------------|---|
| I | +VICTORY_00 | ÉXITO | <i>éxito, victoria, triunfo</i> | --- |
| II | +PERIOD_00 | TIEMPO | <i>vida, juventud, día</i> | --- |
| III | +STRENGHT_00 | FUERZA | <i>fuerza, energía</i> | <i>esfuerzo, salud</i> |
| III | +ABILITY_00 | CAPACIDAD, FACULTAD | <i>talento, capacidad, don</i> | <i>cualidad, valor, generosidad, credibilidad, prestigio, experiencia</i> |
| III | +INFORMATION_OBJECT_00 | INFORMACIÓN | <i>palabra, página, línea</i> | <i>saliva, óleo, tinta, cinta</i> |
| III | +BULLET_00 | ARMA | <i>munición, bala</i> | <i>disparo, cartuchera</i> |
| IV | --- | RECURSO | --- | <i>dinero, agua, corcho, petróleo, gasolina, papel</i> |
| IV | --- | POSIBILIDAD | --- | <i>oportunidad, ocasión, balón, ventaja, punto, opción, privilegio, posibilidad</i> |

En el caso de *malgastar* –al igual que en el de otros predicados monosémicos– se hace muy claro que la extensión de un concepto no necesariamente influye o cambia la intensidad del mismo. Podríamos decir también que la intensidad del concepto no incluye su extensión, pero esto no es del todo cierto. Si bien no se puede hacer una generalización a nivel conceptual sobre las clases léxicas que elige *malgastar* –porque no comparten rutas conceptuales–, sí es evidente que solo se puede *malgastar* algo que se considere un ‘recurso’ o ‘algo valioso’. También es evidente el fenómeno de coerción semántica (Pustejovsky, 1995) o atribución implícita (Bosque, 2004b), en donde los argumentos adquieren propiedades semánticas dadas en combinación con el predicado. Aunque algunos de los argumentos tienen valoraciones positivas inherentes (*talento, generosidad, prestigio, ventaja, éxito, victoria, triunfo*), otros adquieren esa valoración (*saliva, óleo, tinta, cinta*) en combinación con *malgastar*.

2.6 DILAPIDAR

REDES no identifica ningún sentido para el verbo *dilapidar*. Es un verbo monosémico que significa ‘malgastar’; lo vinculamos con +WASTE_00:

Tabla187. Descripción de +WASTE_00 en FunGramKB

| | |
|-----------------|--|
| Hiperónimo: | +USE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2)Referent |
| PS: | +(e1: +USE_00 (x1)Theme (x2)Referent (f1: +CARELESS_00)Manner) |
| Descripción: | use inefficiently or inappropriately; "waste heat" |
| Lexicalización: | <i>malgastar</i> (inglés), <i>waste</i> (español) |

La tabla siguiente muestra la combinatoria léxica de *dilapidar*, identificada por REDES:

Tabla 188. Combinatoria léxica de *dilapidar*, en REDES

| CONCEPTO asignado a la CL en REDES | Descriptor de la CL en REDES | Argumentos miembros de la CL en REDES |
|------------------------------------|--|---|
| (DINERO) | bienes materiales | <i>capital, dinero</i> |
| (RECURSO) | recursos naturales | <i>agua, oxígeno</i> |
| (RECURSO) | personas o grupos, cuando se asimilan a lo anterior | <i>jugador, equipo</i> |
| POSIBILIDAD | situación favorable para ejecutar o conseguir algo, frec. deportivo | <i>oportunidad, situación, posibilidad, opción, baza, ventaja, renta, margen</i> |
| ESFUERZO | esfuerzo o intento porfiado de alcanzar alguna cosa, y otros resultados deseables | <i>esfuerzo, intento, victoria, logro, energía, resistencia, trabajo</i> |
| TIEMPO | sustantivos temporales | <i>tiempo, historia, año, hora, instante, momento, futuro, carrera, vida</i> |
| --- | fama, prestigio y méritos y atributos notables antes los demás | <i>fama, prestigio, crédito, credibilidad, imagen, gloria, grandeza, seña de identidad, liderazgo</i> |
| FACULTAD, PROMINENCIA | facultades del individuo, relacionadas con la creación, conocimiento o capacidad de ejecución | <i>talento, capacidad, inteligencia, sapiencia, idea, experiencia, intuición</i> |
| ASPIRACIÓN | deseos, anhelos, aspiraciones | <i>aspiración, sueño, esperanza, ilusión</i> |
| SENTIMIENTO | <i>sentimiento</i> , y susts. de afecto, solidaridad y otros sentimientos y actitudes nobles asociado a la estima hacia otras personas o cosas | <i>sentimiento, confianza, cariño, solidaridad, respaldo, apoyo, concordia, fe</i> |

Veamos ahora la vinculación de cada clase léxica con entidades de la ontología de FunGramKB. La primera clase léxica se vincula con la entidad +MONEY_00. Es una clase estable (tipo I). Todos los argumentos de la clase léxica pueden lexicalizar la entidad, y la entidad conceptual describe la clase léxica:

Tabla189. Comparación entre clase léxica de REDES y +MONEY_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|------------------------|---|
| Superodinado: | --- | +GROUP_00 |
| Concepto: | (DINERO) | +MONEY_00 |
| PS: | --- | +((e1: +BE_00 (x1: +MONEY_00)Theme (x2: +GROUP_00)Referent)(e2: +COMPRISE_00 (x2)Theme (x3: +COIN_00 +BANKNOTE_00 +CHEQUE_00)Referent)) |
| Descripción: | bienes materiales | the most common medium of exchange; functions as legal tender; "we tried to collect the money he owed us" |
| Manifestación léxica | <i>capital, dinero</i> | <i>capital, cuarto, dinero, guita, mosca, parné, pasta, perras, plata, tela</i> (español), <i>capital, dosh, dough, money</i> (inglés) |

Las próximas dos clases léxicas no se pueden vincular a ninguna entidad. Se describen en REDES como ‘recursos naturales’ (*agua, oxígeno*) y ‘personas o grupos, cuando se asimilan a lo anterior’ (*jugador, equipo*). El problema aquí es la noción de REDES de RECURSO, que hemos discutido antes. No se puede vincular a ninguna entidad en la ontología de FunGramKB, porque es una noción pragmática que depende del contexto situacional, no de los rasgos inherentes de las palabras.

La próxima clase léxica tampoco se puede vincular con la ontología. Se trata de la noción de POSIBILIDAD, que hemos discutido antes. REDES describe esta clase léxica como ‘situación favorable para ejecutar o conseguir algo, frecuentemente deportiva’ y tiene los miembros *ocasión, oportunidad, situación, posibilidad, opción, baza, ventaja, renta, margen*. Esta es una noción que no se puede codificar en la ontología de FunGramKB, por ser una noción pragmática, dependiente del contexto situacional.

Vinculamos la próxima clase léxica con que se combina *dilapidar*[+WASTE_00] con la entidad conceptual +EFFORT_00. Se trata de una clase mixta (tipo III): hay tres argumentos –*esfuerzo, intento, trabajo, energía*– que lexicalizan el concepto, mientras que los demás –*victoria, logro, resistencia*– adquieren el sentido de manera figurada.

Tabla 190. Comparación entre clase léxica de REDES y +EFFORT_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|---|--|
| Superodinado: | --- | +OCCURRENCE_00 |
| Concepto: | ESFUERZO | +EFFORT_00 |
| PS: | --- | +(e1: +BE_00 (x1: +EFFORT_00)Theme (x2: +OCCURRENCE_00)Referent) +(e2: +TRY_00 (x3: +HUMAN_00 ^ +ANIMAL_00)Theme (x4)Referent (f1: +MUCH_00)Quantity (f2: x1)Scene) |
| Descripción: | esfuerzo o intento porfiado de alcanzar alguna cosa, y otros resultados deseables | earnest and conscientious activity intended to do or accomplish something: "made an effort to cover all the reading material"; "wished him luck in his endeavor"; "she gave it a good try" |
| Manifestación léxica | <i>esfuerzo, victoria, energía, trabajo</i> , <i>intento, logro, resistencia,</i> | <i>afán, ahínco, ánimo, animosidad, brío, empeño, esfuerzo, intento, sudor</i> (español), <i>attempt, effort, endeavour, try</i> (inglés) |

La próxima clase léxica la vinculamos con +PERIOD_00. Se trata de una clase del tipo II, que se describe con un concepto superordinado. Todos los argumentos son subordinados de +PERIOD_00: *tiempo*[+TIME_00], *historia*[+PAST_00], *año*[+YEAR_00], *hora*[+HOUR], *instante*[+MOMENT_00], *momento*[+MOMENT_00], *futuro*[+FUTURE_00]. Solo hay dos argumentos -*carrera, vida*- que adquieren el sentido de ‘periodo de tiempo’ de manera figurada, en combinación con *dilapidar*.

Tabla 191. Comparación entre clase léxica de REDES y +PERIOD_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|-----------------------|--|---|
| Superodinado: | --- | +TIME_00 |
| Concepto: | TIEMPO | +PERIOD_00 |
| PS: | --- | +(e1: +BE_00 (x1: +PERIOD_00)Theme (x2: +TIME_00)Referent) |
| Descripción: | sustantivos temporales | a particular length of time with a beginning and an end |
| Manifestación léxica: | <i>tiempo, historia, año, hora, instante, momento, futuro, carrera, vida</i> | <i>edad, época, era, intervalo, periodo, tiempo</i> (español), <i>age, epoch, era, interval, period, span, stretch, term, time</i> (inglés) |

Vinculamos la próxima clase léxica a la entidad +FAME_00. Es una clase léxica mixta (tipo III). La clase tiene miembros que lexicalizan el concepto +FAME_00: *fama*,

prestigio, y otros que adquieren ese sentido en combinación con *dilapidar*: *crédito*, *credibilidad*, *imagen*, *gloria*, *grandeza*, *seña de identidad*, *liderazgo*.

Tabla 192. Comparación entre clase léxica de REDES y +FAME_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|------------------------------|---|---|
| Superodinado: | --- | +STATE_00 |
| Concepto: | --- | +FAME_00 |
| PS: | --- | +(e1: +BE_00 (x1: +FAME_00)Theme (x2: +STATE_00)Referent) +(e2: +BE_01 (x3)Theme (x4: m +POPULAR_00)Attribute (f1: x1)Scene) |
| Descripción: | fama, prestigio y méritos y atributos notables antes los demás | the state or quality of being widely honored and acclaimed |
| Manifestación léxica: | <i>fama, prestigio, crédito, credibilidad, imagen, gloria, grandeza, seña de identidad, liderazgo</i> | <i>celebridad, fama, popularidad, renombre</i> (español), <i>celebrity, fame, popularity, renown</i> (inglés) |

Vinculamos la próxima clase léxica a la entidad conceptual +ABILITY_00. Es una clase mixta (tipo III). Los argumentos *talento* y *capacidad* son lexicalizaciones de +ABILITY_00, mientras que *inteligencia*, *sapiencia* son subordinados; por otro lado, *idea*, *experiencia*, *intuición* adquieren el sentido de ‘habilidad’ en combinación con *dilapidar*.

Tabla 193. Comparación entre clase léxica de REDES y +ABILITY_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|------------------------------|---|---|
| Superodinado: | --- | +PSYCHOLOGICAL_ATT_00 |
| Concepto: | FACULTAD, PROMINENCIA | +ABILITY_00 |
| PS: | --- | +(e1: +BE_00 (x1: +ABILITY_00)Theme (x2: +PSYCHOLOGICAL_ATT_00)Referent) *(e2: pos +DO_00 (x3: +HUMAN_00 ^ +ANIMAL_00)Theme (x4)Referent (f1: x1)Means) |
| Descripción: | facultades del individuo, relacionadas con la creación, conocimiento o capacidad de ejecución | possession of the qualities (especially mental qualities) required to do something or get something done; "danger heightened his powers of discrimination" |
| Manifestación léxica: | <i>talento, capacidad, inteligencia, sapiencia, idea, experiencia, intuición</i> | <i>aptitud, capacidad, competencia, destreza, don, facultad, talento</i> (español), <i>ability, capability, capacity, competence, faculty, gift, power, talent</i> (inglés) |

La próxima clase léxica se vincula con +DESIRE_00. Es una clase léxica estable (tipo I), cuyos miembros pueden lexicalizar el concepto: *aspiración, sueño, esperanza, ilusión*:

Tabla 194. Comparación entre clase léxica de REDES y +DESIRE_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|--|--|
| Superodinado: | --- | +FEELING_00 |
| Concepto: | ASPIRACIÓN, INCLINACIÓN | +DESIRE_00 |
| PS: | --- | +(e1: +BE_00 (x1: +DESIRE_00)Theme (x2: +FEELING_00)Referent) *(e2: +FEEL_00 (x3: (e3: +DESIRE_01 (x4: +HUMAN_00)Theme (x5)Referent))Agent (x4)Theme (x1)Attribute) |
| Descripción: | deseos, anhelos, aspiraciones | a strong hope or wish |
| Manifestación léxica | <i>aspiración, sueño, esperanza, ilusión</i> | <i>deseo</i> (español), <i>desire, wish</i> (inglés) |

Vinculamos la próxima clase a la entidad conceptual +FEELING_00. Se trata de una clase mixta (tipo III). *Sentimiento* es una lexicalización de +FEELING_00, mientras que *confianza, cariño* son subordinados. Por otro lado, *solidaridad, respaldo, apoyo, concordia, fe* adquieren el sentido de ‘sentimiento’ en combinación con *dilapidar*:

Tabla 195. Comparación entre clase léxica de REDES y +FEELING_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|-----------------------|---|--|
| Superodinado: | --- | +PSYCHOLOGICAL_ATT_00 |
| Concepto: | SENTIMIENTO | +FEELING_00 |
| PS: | --- | +(e1: +BE_00 (x1: +FEELING_00)Theme (x2: +PSYCHOLOGICAL_ATT_00)Referent) +(e2: +FEEL_00 (x3)Agent (x4: +HUMAN_00)Theme (x1)Attribute) |
| Descripción: | <i>sentimiento</i> , y susts. de afecto, <i>solidaridad</i> y otros sentimientos y actitudes nobles asociado a la estima hacia otras personas o cosas | an emotion that you feel; emotional attribute |
| Manifestación léxica: | <i>sentimiento, confianza, cariño, solidaridad, respaldo, apoyo, concordia, fe</i> | <i>ánimo, emoción, sentimiento</i> (español), <i>emotion, feeling, humour, mood, sentiment, spirit, temper</i> (inglés) |

La siguiente tabla muestra las rutas conceptuales de las clases léxicas con las que se combina el verbo *dilapidar*[+WASTE_00]. La selección es diversa, pero en su mayoría abstracta, y contiene varios atributos psicológicos:

Tabla 196. Rutas conceptuales de conceptos vinculados a clases léxicas que se combinan con *dilapidar*[+WASTE]

| Rutas conceptuales -----> -----> -----> -----> | | | | | |
|--|------------|-------------|--------------------|----------|---------|
| #Abstract | #Attribute | +Attribute | +Psychological-Att | +Ability | |
| | | | | +Feeling | |
| | | | | +Feeling | +Desire |
| | #Quantity | #Time | +Time | +Period | |
| | +State | +Fame | | | |
| #Physical | #Object | #Collection | +Group | +Money | |
| | #Process | +Occurrence | +Effort | | |

A pesar de que *dilapidar* elige varios atributos psicológicos pertenecientes a los subordinados +ABILITY_00 y +FEELING_00, no incluye el tercer subordinado de +PSYCHOLOGICAL_ATT, que es +BEHAVIOUR_00. Por eso, optamos por no formular un subconcepto -WASTE_PSYCHOLOGICAL_ATT. En su lugar, formulamos: -WASTE_ABILITY y -WASTE_FEELING, por un lado, y -WASTE_PERIOD, -WASTE_FAME, -WASTE_MONEY, -WASTE_EFFORT, por otro. Los subconceptos se describen a continuación:

Tabla 197. Descripción de -WASTE_ABILITY (subconcepto propuesto)

| | |
|-----------------|--|
| Hiperónimo: | +USE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2:+ABILITY_00)Referent |
| Descripción: | use abilities inefficiently or inappropriately; "waste your talents" |
| Lexicalización: | <i>dilapidar</i> (español), <i>waste</i> (inglés) |

Tabla 198. Descripción de -WASTE_FEELING (subconcepto propuesto)

| | |
|-----------------|--|
| Hiperónimo: | +USE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2:+FEELING_00)Referent |
| Descripción: | use emotions inefficiently or inappropriately; "waste your love" |
| Lexicalización: | <i>dilapidar</i> (español), <i>waste</i> (inglés) |

Tabla 199. Descripción de -WASTE_PERIOD (subconcepto propuesto)

| | |
|-----------------|---|
| Hiperónimo: | +USE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2:+PERIOD_00)Referent |
| Descripción: | use times inefficiently or inappropriately; "waste an hour" |
| Lexicalización: | <i>dilapidar</i> (español), <i>waste</i> (inglés) |

Tabla 200. Descripción de -WASTE_FAME (subconcepto propuesto)

| | |
|-----------------|---|
| Hiperónimo: | +USE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2:+FAME_00)Referent |
| Descripción: | use fame inefficiently or inappropriately; "waste the recognition you have" |
| Lexicalización: | <i>dilapidar</i> (español), <i>waste</i> (inglés) |

Tabla 201. Descripción de -WASTE_MONEY (subconcepto propuesto)

| | |
|-----------------|---|
| Hiperónimo: | +USE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2:+MONEY_00)Referent |
| Descripción: | use money inefficiently or inappropriately; "waste an hour" |
| Lexicalización: | <i>dilapidar</i> (español), <i>waste</i> (inglés) |

Tabla 202. Descripción de -WASTE EFFORT (subconcepto propuesto)

| | |
|-----------------|---|
| Hiperónimo: | +USE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2:+EFFORT_00)Referent |
| Descripción: | use an effort inefficiently or inappropriately; "a wasted effort" |
| Lexicalización: | <i>dilapidar</i> (español), <i>waste</i> (inglés) |

2.6.1 Resumen DILAPIDAR

La tabla siguiente muestra todos los vínculos de *dilapidar* en FunGramKB, tanto con eventos conceptuales en la ontología como con colocaciones en la plantilla léxica:

Tabla 203. Resumen: DILAPIDAR en FunGramKB (vinculación propuesta)

| Evento conceptual | Colocaciones |
|-------------------|---|
| +WASTE_00 | <i>agua, oxígeno; jugador, equipo; ocasión, oportunidad, situación, posibilidad, opción, baza, ventaja, renta, margen</i> |
| -WASTE_ABILITY | <i>idea, experiencia, intuición</i> |
| -WASTE_FEELING | <i>solidaridad, respaldo, apoyo, concordia, fe</i> |
| -WASTE_PERIOD | --- |
| -WASTE_FAME | <i>crédito, credibilidad, imagen, gloria, grandeza, seña de identidad, liderazgo</i> |
| -WASTE_MONEY | --- |
| -WASTE EFFORT | <i>victoria, logro, resistencia</i> |

La próxima tabla presenta la tipología de las clases léxicas con que se combina *dilapidar*[+WASTE_00], y la distribución de los argumentos entre el paradigma semántico y el paradigma discursivo o pragmático:

Tabla 204. Tipología de clases léxicas que se combinan con *dilapidar* con respecto a FunGramKB

| Tipo de clase | CONCEPTO en FunGramKB | CONCEPTO en REDES | Argumentos CL – paradigma semántico | Argumentos CL – paradigma discursivo o pragmático |
|---------------|-----------------------|-----------------------|--|---|
| I | +MONEY_00 | (DINERO) | <i>capital, dinero</i> | --- |
| I | +DESIRE_00 | ASPIRACIÓN | <i>aspiración, sueño, esperanza, ilusión</i> | --- |
| II | +PERIOD_00 | TIEMPO | <i>tiempo, historia, año, hora, instante, momento, futuro, carrera, vida</i> | --- |
| III | +EFFORT_00 | ESFUERZO | <i>esfuerzo, intento, energía, trabajo</i> | <i>victoria, logro, resistencia</i> |
| III | +FAME_00 | --- | <i>fama, prestigio,</i> | <i>crédito, credibilidad, imagen, gloria, grandeza, seña de identidad, liderazgo</i> |
| III | +ABILITY_00 | FACULTAD, PROMINENCIA | <i>talento, capacidad, inteligencia, sapiencia</i> | <i>idea, experiencia, intuición</i> |
| III | +FEELING_00 | SENTIMIENTO | <i>sentimiento, confianza, cariño</i> | <i>solidaridad, respaldo, apoyo, concordia, fe</i> |
| IV | --- | (RECURSO) | --- | <i>agua, oxígeno; jugador, equipo</i> |
| IV | --- | POSIBILIDAD | --- | <i>ocasión, oportunidad, situación, posibilidad, opción, baza, ventaja, renta, margen</i> |

2.7 DESAPROVECHAR

El verbo *desaprovechar* significa, a grandes rasgos, ‘dejar pasar algo que redundaría en ventaja, perder lo que se había adelantado’. Al igual que *malgastar* y *dilapidar*, en FunGramKB se vincula solo con +WASTE_00:

Tabla 205. Descripción de +WASTE_00 en FunGramKB

| | |
|-----------------|---|
| Hiperónimo: | +USE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2)Referent |
| PS: | +(e1: +USE_00 (x1)Theme (x2)Referent (f1: CARELESS_00)Manner) |
| Descripción: | use inefficiently or inappropriately; "waste heat" |
| Lexicalización: | <i>malgastar</i> (inglés), <i>waste</i> (español) |

Tomamos la combinatoria sistemática de *desaprovechar* del diccionario PRÁCTICO⁹⁰. Sus restricciones de selección léxica no afectan su significado conceptual; el vínculo con +WASTE_00, se mantiene en combinación con todas las clases léxicas.

Tabla 206. Combinatoria léxica de *desaprovechar*, en PRÁCTICO

| (Concepto) | Argumentos |
|-----------------------|---|
| (CAPACIDAD, FACULTAD) | <i>talento, capacidad, facultad, facilidad, conocimiento, cualidad, disposición, carrera, potencial</i> |
| (FUERZA) | <i>fuerza, energía, impulso, esfuerzo</i> |
| (TIEMPO) | <i>tiempo, momento</i> |
| (POSIBILIDAD) | <i>ocasión, oportunidad, opción, ventaja, racha, visita, presencia, viaje</i> |
| (IDEA) | <i>información, idea, iniciativa, sugerencia</i> |
| (RECURSO) | <i>recurso</i> |

La primera clase léxica de *desaprovechar*[+WASTE_00] la vinculamos con la entidad +ABILITY_00. Es una clase mixta (tipo III): los argumentos *talento, capacidad, facultad, facilidad, potencial* lexicalizan el concepto +ABILITY_00, pero *conocimiento, cualidad, disposición, carrera* adquieren ese valor de manera figurada en combinación con *desaprovechar*.

⁹⁰ PRÁCTICO no ofrece descriptores para las clases léxicas, ni las asocia con conceptos. Sin embargo, siguiendo el modelo de REDES, les hemos asignado a las clases léxicas un concepto del índice conceptual de REDES, que se demarca entre paréntesis.

Tabla 207. Comparación entre clase léxica de REDES y +ABILITY_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|-------------------------------|---|--|
| Superodinado: | --- | +PSYCHOLOGICAL_ATT_00 |
| Concepto: | (CAPACIDAD, FACULTAD) | +ABILITY_00 |
| PS: | --- | +(e1: +BE_00 (x1: +ABILITY_00)Theme (x2: +PSYCHOLOGICAL_ATT_00)Referent) *(e2: pos +DO_00 (x3: +HUMAN_00 ^ +ANIMAL_00)Theme (x4)Referent (f1: x1)Means) |
| Descripción semántica: | --- | possession of the qualities (especially mental qualities) required to do something or get something done; "danger heightened his powers of discrimination" |
| Manifestación léxica: | <i>talento, capacidad, facultad, facilidad, conocimiento, cualidad, disposición, carrera, potencial</i> | <i>aptitud, capacidad, competencia, destreza, don, facultad, talento (español), ability, capability, capacity, competence, faculty, gift, power, talent (inglés)</i> |

Vinculamos la segunda clase léxica a +STRENGTH_00. Se trata de una clase léxica estable (tipo I), donde todos los argumentos pueden lexicalizar el concepto:

Tabla 208. Comparación entre clase léxica de REDES y +STRENGHT_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|-------------------------------|---|--|
| Superodinado: | --- | +PHYSICAL_ATT_00 |
| Concepto: | (FUERZA) | +STRENGHT_00 |
| PS: | --- | +(e1: +BE_00 (x1: +STRENGTH_00)Theme (x2: +PHYSICAL_ATT_00)Referent) *(e2: +BE_01 (x3: +HUMAN_00 ^ +ANIMAL_00)Theme (x1)Attribute) |
| Descripción semántica: | --- | the property of being physically or mentally strong; "fatigue sapped his strenght" |
| Manifestación léxica: | <i>fuerza, energía, impulso, esfuerzo</i> | <i>fuerza (español), force, strenght (ing)</i> |

La próxima clase léxica se vincula a +TIME_00. Es una clase estable (tipo I), con miembros que pueden lexicalizar la entidad +TIME_00:

Tabla 209. Comparación entre clase léxica de REDES y +TIME_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|------------------------|------------------------|--|
| Superordinado: | --- | #TIME_00 |
| Concepto: | (TIEMPO) | +TIME_00 |
| PS: | --- | +(e1: +BE_00 (x1: +TIME_00)Theme (x2: +QUANTITY_00)Referent) |
| Descripción semántica: | --- | the thing that is measured in minutes, hours, days, years etc using clocks |
| Manifestación léxica | <i>tiempo, momento</i> | <i>tiempo</i> (español), <i>time</i> (ing) |

La próxima clase léxica pertenece al tipo IV: no se puede vincular a una entidad de FunGramKB. Se trata de una clase que ya hemos visto: se correspondería con el concepto POSIBILIDAD en REDES, pues sus miembros son *ocasión, oportunidad, opción, ventaja, racha, clima, visita, presencia*. Sin embargo, la noción de ‘posibilidad’ no tiene representación en la ontología. Por lo tanto, todos sus miembros serán ubicados en como colocaciones en la plantilla léxica de *desaprovechar*.

La próxima clase léxica con que se combina *desaprovechar*[+WASTE_00] la vinculamos con la entidad +INFORMATION_00. Esta es una clase mixta (tipo III): *información* lexicaliza el concepto +INFORMATION_00, mientras *que idea, iniciativa, sugerencia* adquieren el valor al combinarse con *desaprovechar*.

Tabla 210. Comparación entre clase léxica de REDES y +INFORMATION_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|------------------------|--|--|
| Superordinado: | --- | +THOUGHT_00 |
| Concepto: | (INFORMACIÓN) | +INFORMATION_00 |
| PS: | --- | +(e1: +BE_00 (x1: +INFORMATION_00)Theme (x2: +THOUGHT_00)Referent) *((e2: +KNOW_00 (x3)Theme (x1)Referent)(e3: +BE_01 (x4)Theme (x1)Attribute)) |
| Descripción semántica: | --- | facts or details that tell you something about a situation, person, event etc |
| Manifestación léxica | <i>información, idea, iniciativa, sugerencia</i> | <i>dato, información</i> (español), <i>data, information</i> (inglés) |

El último argumento, *recurso*, no se pudo vincular a ningún concepto. Se correspondería con el concepto RECURSO en REDES, que, como hemos comentado antes, no tiene equivalente en una ontología como FunGramKB. Por lo tanto, pertenece al tipo IV.

La próxima tabla muestra las rutas conceptuales de las clases léxicas con que se combina *desaprovechar*[+WASTE_00]. No comparten mucho, salvo que son todas entidades abstractas:

Tabla 211. Rutas conceptuales de conceptos vinculados a clases léxicas que se combinan con *desaprovechar*[WASTE+ 00]

| Rutas conceptuales -----> -----> -----> -----> -----> | | | | | |
|---|--------------|------------|--------------------|-----------|--|
| #Abstract | #Proposition | +Thought | +Information | | |
| #Abstract | #Quantity | #Time | +Time | | |
| #Abstract | #Attribute | +Attribute | +Physical_Att | +Strength | |
| | | | +Psychological_Att | +Ability | |

A partir de estas selecciones léxicas, proponemos la creación de los siguientes subconceptos de +WASTE_00: -WASTE_TIME, -WASTE_STRENGTH, -WASTE_ABILITY, -WASTE_INFORMATION, que se describen a continuación.

Tabla 212. Descripción de -WASTE_TIME (subconcepto propuesto)

| | |
|-----------------|--|
| Hiperónimo: | +USE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2:+TIME_00)Referent |
| Descripción: | use time inefficiently or inappropriately; "waste an hour" |
| Lexicalización: | <i>desaprovechar</i> (español), <i>waste</i> (inglés) |

Tabla 213. Descripción de -WASTE_STRENGTH (subconcepto propuesto)

| | |
|-----------------|--|
| Hiperónimo: | +USE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2:+STRENGTH_00)Referent |
| Descripción: | use strenght inefficiently or inappropriately; "waste your energy" |
| Lexicalización: | <i>desaprovechar</i> (español), <i>waste</i> (inglés) |

Tabla 214. Descripción de -WASTE_ABILITY (subconcepto propuesto)

| | |
|-----------------|---|
| Hiperónimo: | +USE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2:+ABILITY_00)Referent |
| Descripción: | use abilities inefficiently or inappropriately; "waste ones talent" |
| Lexicalización: | <i>desaprovechar</i> (español), <i>waste</i> (inglés) |

Tabla 215. Descripción de -WASTE_INFORMATION (subconcepto propuesto)

| | |
|-----------------|---|
| Hiperónimo: | +USE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2:+INFORMATION_00)Referent |
| Descripción: | use achievements inefficiently or inappropriately; "waste your accomplishments" |
| Lexicalización: | <i>desaprovechar</i> (español), <i>waste</i> (inglés) |

2.7.1 Resumen DESAPROVECHAR

La próxima tabla muestra el resumen de los vínculos del verbo *desaprovechar* en FunGramKB. La columna de la izquierda muestra el vínculo con un evento verbal, mientras que la de la derecha muestra las colocaciones que se ubicarán en la plantilla léxica de *desaprovechar* en ese sentido. Solo los argumentos que no están representados conceptualmente en una preferencia de selección se consideran colocaciones.

Tabla 216. Resumen: DESAPROVECHAR en FunGramKB (vinculación propuesta)

| Evento conceptual | Colocaciones |
|--------------------|---|
| +WASTE_00 | <i>oportunidad, opción, ventaja, racha, visita, presencia, viaje; recurso</i> |
| -WASTE_TIME | --- |
| -WASTE_STRENGTH | --- |
| -WASTE_ABILITY | <i>conocimiento, calidad, disposición, carrera</i> |
| -WASTE_INFORMATION | <i>idea, iniciativa, sugerencia</i> |

En la tabla siguiente, se muestra la vinculación con una entidad de FunGramKB de cada clase léxica con que se combina *desaprovechar*[+WASTE_00], y su clasificación por tipos. Asimismo, se distribuyen los argumentos entre los que pertenecen al paradigma semántico de las clases y los que pertenecen a paradigmas discursivos o pragmáticos.

Tabla 217. Tipología de clases léxicas que se combinan con *desaprovechar* con respecto a FunGramKB

| Tipo de clase | CONCEPTO en FunGramKB | CONCEPTO en REDES | Argumentos CL – Paradigma semántico | Argumentos CL – Paradigma discursivo pragmático |
|---------------|-----------------------|-----------------------|---|---|
| I | +STRENGTH_00 | (FUERZA) | <i>fuerza, energía, impulso, esfuerzo</i> | --- |
| I | +TIME_00 | (TIEMPO) | <i>tiempo, momento</i> | --- |
| III | +ABILITY_00 | (CAPACIDAD, FACULTAD) | <i>talento, capacidad, facultad, facilidad, potencial</i> | <i>conocimiento, cualidad, disposición, carrera</i> |
| III | +INFORMATION_00 | (IDEA) | <i>información</i> | <i>idea, iniciativa, sugerencia</i> |
| IV | --- | (POSIBILIDAD) | --- | <i>ocasión, oportunidad, opción, ventaja, racha, visita, presencia, viaje</i> |
| IV | --- | (RECURSO) | --- | <i>recurso</i> |

En el caso de *desaprovechar*, no se puede hacer una generalización a nivel conceptual sobre las clases léxicas seleccionadas –porque no comparten rutas conceptuales–, pero sí es evidente que solo se puede *desaprovechar* algo que supone una posible ventaja o provecho. Aunque algunos de los argumentos tienen valoraciones positivas inherentes (*talento, capacidad, facultad, facilidad, potencial, ventaja*), otros adquieren esa valoración en combinación con *desaprovechar* (*tiempo, momento, equipo, persona*).

2.8 DESPILFARRAR

Presentamos la combinatoria del verbo *despilfarrar*, según el diccionario PRÁCTICO, en la tabla siguiente:

Tabla 218. Combinatoria léxica de *despilfarrar*, en PRÁCTICO

| CONCEPTO de la CL (usando índice de REDES) | Argumentos miembros de la CL en REDES |
|--|---|
| (DINERO) | <i>dinero, fortuna, millones, ayuda, crédito, ganancia, presupuesto</i> |
| (RECURSO) | <i>energía, agua, reservas</i> |
| (ALIMENTO) | <i>comida, alimento</i> |
| (POSIBILIDAD) | <i>ocasión, oportunidad, posibilidad</i> |
| (CAPACIDAD) | <i>ingenio</i> |
| (ARMA) | <i>bala, munición</i> |

Despilfarrar es un verbo monosémico, cuyo único sentido es ‘malgastar’, que preliminarmente vinculamos en FunGramKB a +WASTE_00:

Tabla 219. Descripción de +WASTE_00 en FunGramKB

| | |
|-----------------|--|
| Superordinado: | +USE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2)Referent |
| PS: | +(e1: +USE_00 (x1)Theme (x2)Referent (f1: +CARELESS_00)Manner) |
| Descripción: | use inefficiently or inappropriately; "waste heat" |
| Lexicalización: | <i>malgastar</i> (español), <i>waste</i> (inglés) |

Como mencionamos antes sobre *derrochar* [+WASTE_00], el predicado expresa un nivel más extremo del evento de ‘malgastar’. Así lo reflejan los diccionarios (subrayamos los contornos relevantes):

Tabla 220. Información lexicográfica de *despilfarrar* ‘malgastar en exceso’

| |
|---|
| Hiperónimos: |
| gastar (x4), consumir, derrochar, malgastar |
| Contornos: |
| una persona // el dinero (4), una cosa (x2), el caudal // sin necesidad (x2), <u>de forma insensata, profusamente</u> , con insensatez, <u>con exceso o sin necesidad</u> , <u>excesiva</u> o indebidamente, <u>en mucha más cantidad de lo necesario o prudente</u> // <u>en gastos desarreglados</u> , en alguna ocasión, en cosas innecesarias |
| Colocaciones: |
| <i>fortuna, todo lo que gana, mucho dinero // en juegos y diversiones, en tonterías</i> |
| Sinónimos: |
| <i>dilapidar</i> (x2), <i>derrochar</i> , <i>disipar</i> , <i>malgastar</i> , <i>tirar</i> |

De ahí que propongamos un cambio al PS de +WASTE_00 para reforzar el evento de +WASTE_00 con el cuantificador relativo *m* (de *much*) (Periñán y Mairal, 2010). El concepto terminal se llamaría \$\$SQUANDER_00, y estaría lexicalizado por *despilfarrar* y *derrochar* en español y *squander* en inglés:

Tabla 221. Descripción de \$\$SQUANDER_00 (concepto terminal propuesto)

| | |
|-----------------|--|
| Hiperónimo: | +WASTE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2)Referent |
| PS: | +(e1:+WASTE_00 (x1)Theme (x2)Referent (f1: +MUCH_00)Quantity) |
| Descripción: | to waste in excess: “squander money” |
| Lexicalización: | <i>despilfarrar</i> , <i>derrochar</i> (español), <i>squander</i> (inglés) |

En cuanto a la combinatoria de *despilfarrar*, hemos vinculado la primera clase léxica a +MONEY_00. Es una clase mixta (tipo III): sus miembros incluyen *fortuna* y *dinero*, que pueden lexicalizar el concepto; *crédito*, *ganancia* y *presupuesto*, que podrían ser

subordinados de +MONEY_00; y *millones* y *ayuda*, que adquieren el sentido de ‘dinero’ de manera figurada:

Tabla 222. Comparación entre clase léxica de REDES y +MONEY_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|---|--|
| Superodinado: | --- | +GROUP_00 |
| Concepto: | (DINERO) | +MONEY_00 |
| PS: | --- | +(e1: +BE_00 (x1: +MONEY_00)Theme (x2: +GROUP_00)Referent)(e2: +COMPRISE_00 (x2)Theme (x3: +COIN_00 +BANKNOTE_00 +CHEQUE_00)Referent)) |
| Descripción: | --- | the most common medium of exchange; functions as legal tender; "we tried to collect the money he owed us" |
| Manifestación léxica | <i>dinero, fortuna, millones, ayuda, crédito, ganancia, presupuesto</i> | <i>capital, cuarto, dinero, guita, mosca, parné, pasta, perras, plata, tela</i> (español), <i>capital, dosh, dough, money</i> (inglés) |

Con la segunda clase, tenemos una dificultad que hemos enfrentado antes. Sus miembros *–energía, agua, reservas–* se identificarían en REDES con la noción de RECURSO, que no tiene equivalencia en FunGramKB por tratarse de un concepto que denota una finalidad práctica que puede asignarse a diferentes entidades según el contexto pragmático. Esta clase tiene el tipo IV: no es codificable en FunGramKB.

La próxima clase es estable (tipo I): la equivalencia entre la clase léxica, cuyos miembros son *comida* y *alimento*, y la entidad conceptual es sólida.

Tabla 223. Comparación entre clase léxica de REDES y +FOOD_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|-------------------------|---|
| Superodinado: | --- | +LIQUID_00 ^ +SOLID_00 |
| Concepto: | (ALIMENTO) | +FOOD_00 |
| PS: | --- | +(e1: +BE_00 (x1: +FOOD_00)Theme (x2: +LIQUID_00 ^ +SOLID_00)Referent) +(e2: +INGEST_00 (x3: +HUMAN_00 ^ +ANIMAL_00)Agent (x1)Theme (f1: (e3: +BE_01 (x3)Theme (x4: +HUNGRY_00)Attribute))Scene) |
| Descripción: | --- | any substance that can be metabolized by an organism to give energy and build tissue |
| Manifestación léxica | <i>comida, alimento</i> | <i>comida</i> (español), <i>food, foodstuff, nutrient</i> (inglés) |

La próxima clase apunta a una noción problemática: POSIBILIDAD. Sus miembros son: *ocasión, oportunidad, posibilidad*. Esta noción no tiene equivalencia en FunGramKB, porque puede aplicar a muchas palabras según el contexto pragmático. Se trata de una clase del tipo IV, que no se puede vincular a ninguna entidad.

La próxima clase la vinculamos a la entidad +ABILITY_00, que *talento* lexicaliza. *Ingenio*, por su parte, refleja un concepto subordinado a +ABILITY_00, vinculado a +INTELLIGENCE_00. Es una clase del tipo II, descrita por superordinado.

Tabla 224. Comparación entre clase léxica de REDES y +ABILITY_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|-----------------------|-------------------------|---|
| Superordinado: | --- | +PSYCHOLOGICAL_ATT_00 |
| Concepto: | (CAPACIDAD) | +ABILITY_00 |
| PS: | --- | +(e1: +BE_00 (x1: +ABILITY_00)Theme (x2: +PSYCHOLOGICAL_ATT_00)Referent) *(e2: pos +DO_00 (x3: +HUMAN_00 ^ +ANIMAL_00)Theme (x4)Referent (f1: x1)Means) |
| Descripción: | --- | possession of the qualities (especially mental qualities) required to do something or get something done; "danger heightened his powers of discrimination" |
| Manifestación léxica: | <i>talento, ingenio</i> | <i>aptitud, capacidad, competencia, destreza, don, facultad, talento</i> (español), <i>ability, capability, capacity, competence, faculty, gift, power, talent</i> (inglés) |

La última clase con que se combina *despilfarrar* la vinculamos con +BULLET_00. Esta es una clase estable (tipo I), pues todos los miembros pueden lexicalizar la entidad.

Tabla 225. Comparación entre clase léxica de REDES y +BULLET_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|-----------------------|-----------------------|---|
| Superordinado: | --- | +PROJECTILE_00 |
| Concepto: | (ARMA) | +BULLET_00 |
| PS: | --- | --- |
| Descripción: | --- | --- |
| Manifestación léxica: | <i>bala, munición</i> | --- (español), <i>bullet, slug</i> (inglés) |

La tabla siguiente muestra las rutas conceptuales de las clases léxicas seleccionadas por *despilfarrar*[$\$SQUANDER_00$]. La mayoría son entidades físicas, pero, aparte de eso, tienen rutas muy diversas:

Tabla 226. Rutas conceptuales de conceptos vinculados a clases léxicas que se combinan con *despilfarrar*[\$\$SQUANDER_00]

| Rutas conceptuales -----> -----> -----> -----> -----> | | | | | |
|---|------------------------|--------------------------------------|--------------------|------------|---------|
| #Abstract | #Attribute | +Attribute | +Psychological-Att | +Ability | |
| #Physical --> #Object | #Collection | +Group | +Money | | |
| | #Self_Connected_Object | +Artificial_Object --> +Substance | +Solid | +Food | |
| | | | +Weapon | Projectile | +Bullet |

A partir de estas selecciones léxicas, proponemos la creación de los siguientes subconceptos de \$\$SQUANDER_00: -SQUANDER_ABILITY, -SQUANDER_MONEY, -SQUANDER_FOOD, -SQUANDER_BULLET, que se describen a continuación:

Tabla 227. Descripción de -SQUANDER_ABILITY (subconcepto propuesto)

| | |
|-----------------|---|
| Hiperónimo: | +WASTE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2:+ABILITY_00)Referent |
| Descripción: | waste abilities dramatically; "squander your talents" |
| Lexicalización: | <i>dilapidar</i> (español), <i>waste</i> (inglés) |

Tabla 228. Descripción de -SQUANDER_MONEY (subconcepto propuesto)

| | |
|-----------------|---|
| Hiperónimo: | +WASTE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2:+MONEY_00)Referent |
| Descripción: | waste money dramatically; "squander a fortune" |
| Lexicalización: | <i>dilapidar</i> (español), <i>waste</i> (inglés) |

Tabla 229. Descripción de -SQUANDER_FOOD (subconcepto propuesto)

| | |
|-----------------|---|
| Hiperónimo: | +WASTE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2:+FOOD_00)Referent |
| Descripción: | waste food dramatically; "squander your food" |
| Lexicalización: | <i>dilapidar</i> (español), <i>waste</i> (inglés) |

Tabla 230. Descripción de -SQUANDER_BULLET (subconcepto propuesto)

| | |
|-----------------|--|
| Hiperónimo: | +WASTE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2:+BULLET_00)Referent |
| Descripción: | waste ammunitions dramatically; "squander the bullets" |
| Lexicalización: | <i>dilapidar</i> (español), <i>waste</i> (inglés) |

2.8.1 Resumen *DESPILFARRAR*

La tabla siguiente muestra los vínculos de *despilfarrar* en FunGramKB, tanto con eventos conceptuales en la ontología, como las colocaciones que se ubican en la plantilla léxica:

Tabla 231. Resumen: *DESPILFARRAR* en FunGramKB (vinculación propuesta)

| Evento conceptual | Colocaciones |
|-------------------|---|
| \$SQUANDER_00 | <i>energía, agua, reservas; ocasión, oportunidad, posibilidad</i> |
| -SQUANDER_ABILITY | --- |
| -SQUANDER_MONEY | <i>millones, ayuda</i> |
| -SQUANDER_FOOD | --- |
| -SQUANDER_BULLET | --- |

En la tabla siguiente, se muestra la vinculación con una entidad de FunGramKB de cada clase léxica con que se combina *despilfarrar*[+WASTE_00], y su clasificación por tipos. Además, se muestran los argumentos de las clases léxicas que pertenecen a un paradigma semántico y los que pertenecen a paradigmas discursivos o pragmáticos:

Tabla 232. Tipología de clases léxicas que se combinan con *despilfarrar*

| Tipo de clase | CONCEPTO en FunGramKB | CONCEPTO en REDES | Argumentos CL – paradigma semántico | Argumentos CL – paradigma discursivo o pragmático |
|---------------|-----------------------|-------------------|--|---|
| I | +FOOD_00 | (ALIMENTO) | <i>comida, alimento</i> | --- |
| I | +BULLET | (ARMA) | <i>bala, munición</i> | --- |
| II | +ABILITY_00 | (CAPACIDAD) | <i>ingenio</i> | --- |
| III | +MONEY_00 | (DINERO) | <i>dinero, fortuna, crédito, ganancia, presupuesto</i> | <i>millones, ayuda</i> |
| IV | --- | (RECURSO) | --- | <i>energía, agua, reservas</i> |
| IV | --- | (POSIBILIDAD) | --- | <i>ocasión, oportunidad, posibilidad</i> |

2.9 *DESPERDICAR*

La combinatoria del verbo *desperdiciar* viene de PRÁCTICO, y se presenta en la próxima tabla:

Tabla 233. Combinatoria léxica de *desperdiciar* en PRÁCTICO

| CONCEPTO asignado a la CL (partiendo de los conceptos de REDES) | Argumentos miembros de la CL en REDES |
|---|---|
| (POSIBILIDAD) | <i>ocasión, oportunidad, ventaja, penalti, voto</i> |
| (TIEMPO) | <i>tiempo, día, semana</i> |
| (TIEMPO) | <i>porvenir, vacaciones, juventud, vida</i> |
| (CAPACIDAD) | <i>talento, capacidad</i> |
| (DINERO) | <i>dinero, riqueza</i> |
| (FUERZA) | <i>fuerzas</i> |
| (ARMA) | <i>bala, munición</i> |
| (RECURSO) | <i>sangre, luz, saliva, petróleo, agua, aceite, energía</i> |
| (ALIMENTO) | <i>comida, alimento</i> |

La primera clase léxica pertenece al tipo IV: no se puede vincular a una entidad en FunGramKB. Se trata de la noción de POSIBILIDAD, con los argumentos *ocasión, oportunidad, ventaja, penalti, voto*. No hay entidad que se corresponda con esta noción, especialmente cuando los miembros son palabras como *ventaja, penalti, voto*, que adquieren este sentido en un contexto situacional pragmático.

Hay dos clases léxicas que vinculamos a +PERIOD_00. La primera tiene miembros que incluyen una lexicalización del concepto (*tiempo*) y dos lexicalizaciones de conceptos subordinados (*día, semana*); pertenece al tipo II. La segunda incluye un miembro que lexicaliza una noción subordinada de +PERIOD_00: *juventud*[+ADOLESCENCE_00], pero los demás son sustantivos que activan el sentido de ‘periodo de tiempo’ en combinación con *desperdiciar*; pertenece al tipo III.

Tabla 234. Comparación entre clase léxica de REDES y +PERIOD_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|--|--|
| Superodinado: | --- | +TIME_00 |
| Concepto: | (TIEMPO) /// (TIEMPO) | +PERIOD_00 |
| PS: | --- | +(e1: +BE_00 (x1: +PERIOD_00)Theme (x2: +TIME_00)Referent) |
| Descripción: | --- | a particular length of time with a beginning and an end |
| Manifestación léxica | <i>tiempo, día, semana // porvenir, vacaciones, juventud, vida</i> | <i>edad, época, era, intervalo, periodo, tiempo (español), age, epoch, era, interval, period, span, stretch, term, time (inglés)</i> |

La próxima clase léxica es una clase estable (tipo I). Los dos argumentos *-talento, capacidad-* lexicalizan la entidad +ABILITY_00:

Tabla 235. Comparación entre clase léxica de REDES y +ABILITY_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|---------------------------|---|
| Superodinado: | --- | +PSYCHOLOGICAL_ATT_00 |
| Concepto: | (CAPACIDAD) | +ABILITY_00 |
| PS: | --- | +(e1: +BE_00 (x1: +ABILITY_00)Theme (x2: +PSYCHOLOGICAL_ATT_00) Referent) *(e2: pos +DO_00 (x3: +HUMAN_00 ^ +ANIMAL_00)Theme (x4)Referent (f1: x1)Means) |
| Descripción: | --- | possession of the qualities (especially mental qualities) required to do something or get something done; "danger heightened his powers of discrimination" |
| Manifestación léxica | <i>talento, capacidad</i> | <i>aptitud, capacidad, competencia, destreza, don, facultad, talento</i> (español), <i>ability, capability, capacity, competence, faculty, gift, power, talent</i> (inglés) |

La próxima clase léxica también es una clase estable (tipo I). La vinculamos con la entidad +MONEY_00, que está lexicalizada por los dos argumentos: *dinero, riqueza*.

Tabla 236. Comparación entre clase léxica de REDES y +MONEY_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|-----------------------|------------------------|--|
| Superodinado: | --- | +GROUP_00 |
| Concepto: | (DINERO) | +MONEY_00 |
| PS: | --- | +(e1: +BE_00 (x1: +MONEY_00)Theme (x2: +GROUP_00)Referent)(e2: +COMPRISE_00 (x2)Theme (x3: +COIN_00 +BANKNOTE_00 +CHEQUE_00)Referent)) |
| Descripción: | --- | the most common medium of exchange; functions as legal tender; "we tried to collect the money he owed us" |
| Manifestación léxica: | <i>dinero, riqueza</i> | <i>capital, cuarto, dinero, guita, mosca, parné, pasta, perras, plata, tela</i> (español), <i>capital, dosh, dough, money</i> (inglés) |

La próxima clase léxica que se combina con *desperdiciar*[+WASTE_00] la vinculamos con +STRENGHT_00. Tiene un solo miembro: *fuerzas*, que lexicaliza el concepto +STRENGHT_00:

Tabla 237. Comparación entre clase léxica de REDES y +STRENGHT_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|----------------|---|
| Superodinado: | --- | +PHYSICAL_ATT_00 |
| Concepto: | (FUERZA) | +STRENGHT_00 |
| PS: | --- | +(e1: +BE_00 (x1: +STRENGTH_00)Theme (x2: +PHYSICAL_ATT_00)Referent) *(e2: +BE_01 (x3: +HUMAN_00 ^ +ANIMAL_00)Theme (x1)Attribute) |
| Descripción: | --- | the property of being physically or mentally strong; "fatigue sapped his strenght" |
| Manifestación léxica | <i>fuerzas</i> | <i>fuerza</i> (español), <i>force, strenght</i> (inglés) |

La próxima clase léxica es estable (tipo I). La vinculamos a +BULLET_00, porque sus dos argumentos lexicalizan esta entidad:

Tabla 238. Comparación entre clase léxica de REDES y +BULLET_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|-----------------------|------------------|
| Superodinado: | --- | FUNGRAMKB |
| Concepto: | (ARMA) | +PROJECTILE_00 |
| PS: | --- | +BULLET_00 |
| Descripción: | --- | --- |
| Manifestación léxica | <i>bala, munición</i> | --- |

La próxima clase léxica con que se combina *desperdiciar*[+WASTE_00] pertenece al tipo IV: no se puede vincular a ninguna entidad. Sus miembros incluyen *sangre, luz, saliva, petróleo, agua, aceite, energía*, y REDES los vincularía al concepto RECURSO. Como hemos discutido antes, la noción de RECURSO no tiene equivalente en la ontología de FunGramKB, por ser asignada por criterios pragmáticos. Según sus rasgos inherentes, estas palabras estarían agrupadas bajo +SUBSTANCE_00, pero se perdería la noción de ‘recurso’.

La última clase léxica con que se combina *desperdiciar*[+WASTE_00] es una clase estable (tipo I). La vinculamos con +FOOD_00. Los dos argumentos *–comida, alimento–* lexicalizan perfectamente el concepto.

Tabla 239. Comparación entre clase léxica de REDES y +FOOD_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|------------------------------|-------------------------|---|
| Superordinado: | --- | +LIQUID_00 ^ +SOLID_00 |
| Concepto: | (ALIMENTO) | +FOOD_00 |
| PS: | --- | +(e1: +BE_00 (x1: +FOOD_00)Theme (x2: +LIQUID_00 ^ +SOLID_00)Referent) +(e2: +INGEST_00 (x3: +HUMAN_00 ^ +ANIMAL_00)Agent (x1)Theme (f1: (e3: +BE_01 (x3)Theme (x4: +HUNGRY_00)Attribute))Scene) |
| Descripción: | --- | any substance that can be metabolized by an organism to give energy and build tissue |
| Manifestación léxica: | <i>comida, alimento</i> | <i>comida</i> (español), <i>food, foodstuff, nutrient</i> (inglés) |

La próxima tabla muestra las rutas conceptuales de las clases léxicas con que se combina *desperdiciar*[+WASTE_00]. Las entidades seleccionadas se dividen equitativamente entre entidades abstractas y físicas, pero aparte de eso, tienen rutas distintas:

Tabla 240. Rutas conceptuales de conceptos vinculados a clases léxicas que se combinan con *desperdiciar*[+WASTE_00]

| Rutas conceptuales -----> -----> -----> -----> -----> | | | | | |
|---|------------------------|--------------------------------------|--------------------|------------|---------|
| #Abstract | #Attribute | +Attribute | +Psychological-Att | +Ability | |
| | | | +Physical_Att | +Strength | |
| | #Quantity | #Time | +Time | +Period | |
| #Physical --> #Object | #Self_Connected_Object | +Artificial_Object --> +Substance | +Solid | +Food | |
| | | | +Weapon | Projectile | +Bullet |
| | #Collection | +Group | +Money | | |

A partir de esta selección léxica, proponemos la creación de los siguientes subconceptos de +WASTE_00: -WASTE_ABILITY, -WASTE_STRENGTH, -WASTE_PERIOD, -WASTE_FOOD, -WASTE_BULLET y -WASTE_MONEY, que se describen a continuación.

Tabla 241. Descripción de -WASTE_ABILITY (subconcepto propuesto)

| | |
|-----------------|---|
| Hiperónimo: | +USE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2:+ABILITY_00)Referent |
| Descripción: | use abilities inefficiently or inappropriately; "waste ones talent" |
| Lexicalización: | <i>desperdiciar</i> (español), <i>waste</i> (inglés) |

Tabla 242. Descripción de -WASTE_STRENGTH (subconcepto propuesto)

| | |
|-----------------|--|
| Hiperónimo: | +USE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2:+STRENGTH_00)Referent |
| Descripción: | use strenght inefficiently or inappropriately; "waste your energy" |
| Lexicalización: | <i>desperdiciar</i> (español), <i>waste</i> (inglés) |

Tabla 243. Descripción de -WASTE_PERIOD (subconcepto propuesto)

| | |
|-----------------|---|
| Hiperónimo: | +USE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2:+PERIOD_00)Referent |
| Descripción: | use times inefficiently or inappropriately; "waste an hour" |
| Lexicalización: | <i>desperdiciar</i> (español), <i>waste</i> (inglés) |

Tabla 244. Descripción de -WASTE_FOOD (subconcepto propuesto)

| | |
|-----------------|--|
| Hiperónimo: | +USE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2:+FOOD_00)Referent |
| Descripción: | use food inefficiently or inappropriately; "waste your sandwich" |
| Lexicalización: | <i>desperdiciar</i> (español), <i>waste</i> (inglés) |

Tabla 245. Descripción de -WASTE_BULLET (subconcepto propuesto)

| | |
|-----------------|--|
| Hiperónimo: | +USE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2:+BULLET_00)Referent |
| Descripción: | use ammunitions inefficiently or inappropriately; "waste your bullets" |
| Lexicalización: | <i>desperdiciar</i> (español), <i>waste</i> (inglés) |

Tabla 246. Descripción de -WASTE_MONEY (subconcepto propuesto)

| | |
|-----------------|---|
| Hiperónimo: | +USE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2:+MONEY_00)Referent |
| Descripción: | use money inefficiently or inappropriately; "waste an hour" |
| Lexicalización: | <i>desperdiciar</i> (español), <i>waste</i> (inglés) |

2.9.1 Resumen DESPERDICAR

La tabla siguiente muestra los vínculos de *desperdicar* en FunGramKB con eventos conceptuales en la ontología. También se identifican las colocaciones que se ubican en las plantillas léxicas de *desperdicar* en su vínculo con diferentes conceptos:

Tabla 247. Resumen: DESPERDICAR en FunGramKB (vinculación propuesta)

| Evento conceptual | Colocaciones |
|-------------------|---|
| +WASTE_00 | <i>ocasión, oportunidad, ventaja, penalti, voto; sangre, luz, saliva, petróleo, agua, aceite, energía</i> |
| -WASTE_ABILITY | --- |
| -WASTE_STRENGTH | --- |
| -WASTE_PERIOD | <i>porvenir, vacaciones, vida</i> |
| -WASTE_FOOD | --- |
| -WASTE_BULLET | --- |
| -WASTE_MONEY | --- |

En la tabla siguiente, se muestra la vinculación de cada clase léxica con que se combina *desperdicar*[+WASTE_00] con una entidad de FunGramKB, y su clasificación por tipos. Además, se muestran los argumentos miembros de las clases léxicas que pertenecen al paradigma semántico y los que pertenecen a paradigmas discursivos o pragmáticos:

Tabla 248. Tipología de clases léxicas que se combinan con desperdicar

| Tipo de clase | CONCEPTO en FunGramKB | en REDES | en REDES | Argumentos CL – paradigma semántico | Argumentos CL – paradigma discursivo o pragmático |
|---------------|-----------------------|--------------------------|----------|--------------------------------------|---|
| I | +ABILITY_00 | (CAPACIDAD) | | <i>talento, capacidad</i> | --- |
| I | +MONEY_00 | (DINERO) | | <i>dinero, riqueza</i> | --- |
| I | +STRENGHT_00 | (FUERZA) | | <i>fuerzas</i> | |
| I | +BULLET_00 | (ARMA) | | <i>bala, munición</i> | --- |
| I | +FOOD_00 | (ALIMENTO) | | <i>comida, alimento</i> | --- |
| II / III | +PERIOD_00 | (TIEMPO) /// (TIEMPO) | | <i>tiempo, día, semana; juventud</i> | <i>porvenir, vacaciones, vida</i> |
| IV | --- | (POSIBILIDAD) | | --- | <i>ocasión, voto, oportunidad, ventaja, penalti,</i> |
| IV | --- | (RECURSO) | | --- | <i>sangre, luz, saliva, petróleo, agua, aceite, energía</i> |

3. Análisis

En esta sección, analizamos los vínculos conceptuales de los nueve verbos, tanto con respecto a los significados del verbo como a sus preferencias de selección. En particular, nos enfocamos en el sentido compartido por los nueve verbos, ‘malgastar’ o +WASTE_00, para establecer patrones comunes y diferencias en las preferencias de selección a partir del concepto.

El vincular los predicados a conceptos de la ontología de FunGramKB nos permite considerar su monosemia o polisemia en términos conceptuales dentro del marco de la ontología. En el caso de palabras monosémicas, el predicado se vincula a un solo evento, y en el de las palabras polisémicas, a varios. La tabla siguiente muestra la vinculación de los sentidos de los nueve predicados a conceptos de FunGramKB:

Tabla 249. Vinculación de los sentidos de los predicados a conceptos de FunGramKB

| <i>predicado / CONCEPTO</i> | +WASTE | +FAIL | \$ABOUND | +SPLIT | +FINISH | +LOSE |
|-----------------------------|------------------|-------------------|----------|--------|---------|-------|
| <i>malgastar</i> | + | | | | | |
| <i>desaprovechar</i> | + | | | | | |
| <i>dilapidar</i> | + | | | | | |
| <i>derrochar</i> | +\$SQUANDER * | | + | | | |
| <i>malograr</i> | + | + | | | | |
| <i>disipar</i> | + | | | + | + | |
| <i>perder</i> | + | +FAIL +\$MISS* | | | | + |
| <i>despilfarrar</i> | +\$SQUANDER * | | | | | |
| <i>desperdiciar</i> | + | | | | | |

*Conceptos terminales propuestos, ubicados bajo su concepto superordinado

En la tabla vemos claramente el vínculo compartido por los nueve predicados en su vínculo con +WASTE_00, cuyo PS lee: ‘una entidad típicamente humana (x1) usa algo (x2) de manera ineficiente o inapropiada’. Cabe recordar aquí que el vincular predicados a un mismo concepto no implica que sean sinónimos, sino que todos comparten un significado fundamental. Los verbos *derrochar* y *despilfarrar* se vinculan a un concepto terminal nuevo, que hemos llamado \$SQUANDER_00, que le añade a +WASTE_00 un parámetro de mayor intensidad.

La tabla nos permite observar también un posible patrón compartido de alternancia de sentidos o polisemia regular. Los verbos *perder* y *malograr* comparten la alternancia entre los sentidos ‘desperdiciar’ (*perder el tiempo / malograr su futuro*) y ‘fracasar’ (*perder el campeonato / malograr el proyecto*), que hemos vinculado a s conceptos +WASTE_00 y +FAIL_00, respectivamente. Aunque son solo dos predicados, y, por lo tanto, no constituyen un ejemplo válido de polisemia regular (que requiere un mínimo de tres), vale la pena examinarlos más de cerca. Al codificar conceptualmente las preferencias de selección de estos dos predicados, a través de la creación de nuevos subconceptos, se precisa que la alternancia no es exacta en el nivel de subconcepto. Ambos se vinculan con subconceptos de +WASTE_00 que están relacionados: -WASTE_PERIOD y -WASTE_FUTURE, porque +PERIOD_00 es el concepto superordinado de +FUTURE_00. Sin embargo, los subconceptos de +FAIL_00 con que se vinculan no comparten ninguna preferencia de selección.

Tabla 250. Polisemia regular entre *PERDER* y *MALOGRAR*

| <i>PERDER</i> | | <i>MALOGRAR</i> | |
|---------------|-------------------|-----------------|----------------|
| Concepto | Subconcepto | Concepto | Subconcepto |
| +WASTE_00 | -WASTE_PERIOD | +WASTE_00 | -WASTE_FUTURE |
| | | | -WASTE_ABILITY |
| +FAIL_00 | -FAIL_TRIAL | +FAIL_00 | -FAIL_PLAN |
| | -FAIL_BATTLE | | -FAIL Effort |
| | -FAIL_COMPETITION | | -FAIL_VICTORY |

No obstante las diferencias en las preferencias de selección de los subconceptos, todos mantienen el PS del concepto bajo el cual se ubican, por lo cual es válido hablar de un ejemplo de polisemia regular, si bien faltaría un tercer ejemplo.

3.1 Preferencias de selección de +WASTE_00

Si nos enfocamos exclusivamente en el vínculo con +WASTE_00 de los verbos, podemos comparar las preferencias de selección conceptuales de cada verbo. Los nueve verbos eligen un total de 46 clases léxicas en su vínculo con +WASTE_00. En su vínculo con +WASTE_00, los verbos monosémicos se combinan con más clases léxicas que los verbos polisémicos. La tabla siguiente muestra la cantidad de clases léxicas con que se combina cada verbo en su vínculo con +WASTE_00:

Tabla 251. Distribución de clases léxicas entre verbos monosémicos y polisémicos en su vínculo con +WASTE_00

| Verbos monosémicos | | Verbos polisémicos | |
|----------------------|-----------------------------------|--------------------|-----------------------------------|
| Verbo | Número de clases léxicas en REDES | Verbo | Número de clases léxicas en REDES |
| <i>dilapidar</i> | 10 | <i>disipar</i> | 1 |
| <i>desperdiciar</i> | 9 | <i>perder</i> | 2 |
| <i>malgastar</i> | 8 | <i>derrochar</i> | 2 |
| <i>desaprovechar</i> | 6 | <i>malograr</i> | 2 |
| <i>despilfarrar</i> | 6 | | |

Estas 46 clases léxicas diferentes se traducen en 12 entidades conceptuales en FunGramKB, lo que demuestra que el inventario de clases léxicas se puede reducir. Cuando examinamos las entidades conceptuales seleccionadas por los predicados en su vínculo con +WASTE_00 en orden de frecuencia, se hace patente que existen patrones de selección léxico semántica entre ellos. En la tabla siguiente, la columna de la izquierda indica la entidad elegida por los verbos. La columna del medio muestra el subconcepto que se debe crear a partir de la restricción de selección léxica documentada. En la columna de la derecha, presentamos los verbos que seleccionan esa entidad y que lexicalizan el subconcepto. La tabla sugiere que +WASTE_00 tiene, como preferencias de selección regulares, a las entidades: +ABILITY_00, +TIME_00, +MONEY_00, +STRENGTH_00 y +BULLET_00. Estas preferencias se manifiestan en por lo menos tres verbos. También tiene otras preferencias de selección que se manifiestan en un solo verbo: +FEELING_00/ +DESIRE_00, +EFFORT_00, +FAME_00, +VICTORY_00, +INFORMATION_OBJECT_00, +INFORMATION_00. Estas preferencias de selección justifican la creación de los subconceptos que se presentan en la columna del medio.

Tabla 252. Selección de entidades por predicados vinculados con +WASTE_00, y subconceptos propuestos

| Concepto FunGramKB elegido | Subconcepto propuesto | Predicados que lo eligen |
|--|-----------------------------------|--|
| +ABILITY_00 | -WASTE_ABILITY | <i>malgastar, desaprovechar, dilapidar, malograr, despilfarrar, desperdiciar</i> |
| +TIME_00/ +PERIOD_00/ +FUTURE_00 ⁹¹ | -WASTE_TIME | <i>malgastar, desaprovechar, dilapidar, perder, desperdiciar, malograr</i> |
| +MONEY_00 | -WASTE_MONEY | <i>dilapidar, derrochar, disipar, despilfarrar, desperdiciar</i> |
| +STRENGTH_00 | -WASTE_STRENGHT | <i>malgastar, desaprovechar, desperdiciar</i> |
| +BULLET_00 | -WASTE_BULLET | <i>malgastar, despilfarrar, desperdiciar</i> |
| +FOOD_00 | -WASTE_FOOD | <i>despilfarrar, desperdiciar</i> |
| +FEELING_00/ +DESIRE_00 | -WASTE_FEELING | <i>dilapidar</i> |
| +EFFORT_00 | -WASTE_EFFORT | <i>dilapidar</i> |
| +FAME_00 | -WASTE_FAME | <i>dilapidar</i> |
| +VICTORY_00 | -WASTE_VICTORY | <i>malgastar</i> |
| +INFORMATION _OBJECT_00 | -WASTE_ INFORMATION _OBJECT | <i>malgastar</i> |
| +INFORMATION_00 | -WASTE_ INFORMATION | <i>desaprovechar</i> |

La aportación de esta tabla para REDES es la codificación de las restricciones de selección léxica de nueve predicados, que sumaban 46 clases léxicas diferentes, en solo 12 subconceptos. El cúmulo de clases léxicas con que se combinan los verbos en el sentido de +WASTE_00 se ha reducido significativamente mediante la vinculación con las entidades conceptuales. La orientación de concepto a predicado nos permite también ver los verbos que comparten preferencias de selección. Por ejemplo, con el subconcepto -WASTE-ABILITY se vinculan los verbos *malgastar, desaprovechar, dilapidar, malograr, despilfarrar, desperdiciar*. El recurso del subconcepto nos demuestra además que la intensión de un concepto no tiene que cambiar, aunque su extensión sí lo haga.

⁹¹ En casos como este, donde las entidades se relacionan como superordinado y subordinado de forma inmediata, las agrupamos juntas.

Para no perder ningún argumento de las clases léxicas, en la plantilla léxica de cada verbo se da cuenta de las colocaciones que no pueden ser representadas conceptualmente por la preferencia de selección. La tabla siguiente muestra las diferentes colocaciones que se ubican en la plantilla léxica de cada verbo que se vincula con -WASTE_ABILITY. Estos son los argumentos miembros de la clase léxica que no lexicalizan la entidad +ABILITY (aunque sí adquieren ese sentido de manera figurada o pragmática):

Tabla 253. Colocaciones en FunGramKB de verbos vinculados con -WASTE_ABILITY

| Predicado | Colocaciones |
|----------------------|---|
| <i>malgastar</i> | <i>cualidad, valor, generosidad, credibilidad, prestigio, experiencia</i> |
| <i>desaprovechar</i> | <i>conocimiento, disposición, carrera</i> |
| <i>dilapidar</i> | <i>idea, experiencia, intuición</i> |
| <i>malograr</i> | <i>voz</i> |
| <i>despilfarrar</i> | --- |
| <i>desperdiciar</i> | --- |

FunGramKB, por su parte, gana mucha más precisión en la representación de las preferencias de selección de los conceptos. Los datos de restricciones de selección de REDES salen del estudio de corpus de uso del español contemporáneo, y se traducen en preferencias de selección al pasar por el filtro de la ontología. Para efectos de una posible aplicación de PLN, la tabla nos informa de que para vincularse con +WASTE_00, *disipar* se tiene que combinar con +MONEY_00 y *perder* con +PERIOD_00 exclusivamente. Sobre los cinco predicados monosémicos –*malgastar, desaprovechar, dilapidar, despilfarrar y desperdiciar*–, se podría argumentar que no es necesario codificar sus preferencias de selección en una base de conocimiento con aplicaciones de WSD para el PLN. Sin embargo, aunque para la WSD tal vez no sea esencial conocer las preferencias de selección de los verbos monosémicos, los datos de REDES hacen constar que existen preferencias de selección determinadas para cada predicado. Sin duda será más natural decir *despilfarrar comida* que *dilapidar comida*, o *dilapidar sueños* que *despilfarrar sueños*.

3.1.1 Clases léxicas y entidades

En esta sección discutiremos las entidades conceptuales de FunGramKB con las que se vincularon un mínimo de dos clases léxicas de REDES. Se trata de entidades conceptuales que constituyen preferencias de selección de +WASTE_00. Estas entidades son:

+ABILITY_00, +TIME_00 (y subordinados), +MONEY_00, +STRENGTH_00, +BULLET_00 y +FOOD_00.

Seis verbos vinculados con +WASTE_00 (*malgastar, desaprovechar, dilapidar, malograr, despilfarrar, desperdiciar*) eligieron clases que vinculamos con +ABILITY_00, cuya descripción es:

Tabla 254. Descripción de +ABILITY_00 en FunGramKB

| | |
|-----------------|---|
| Hiperónimo: | +PSYCHOLOGICAL_ATTRIBUTE_00 |
| PS: | +(e1: +BE_00 (x1: +ABILITY_00)Theme (x2: +PSYCHOLOGICAL_ATT_00) Referent) *(e2: pos +DO_00 (x3: +HUMAN_00 ^ +ANIMAL_00)Theme (x4)Referent (f1: x1)Means) |
| Descripción: | possession of the qualities (especially mental qualities) required to do something or get something done; "danger heightened his powers of discrimination" |
| Lexicalización: | <i>aptitud, capacidad, competencia, destreza, don, facultad, talento</i> (español); <i>ability, capability, capacity, competence, faculty, gift, power, talent</i> (inglés) |

El que haya seis clases léxicas en nuestra muestra que se vinculan con +ABILITY_00 no quiere decir que las clases léxicas sean idénticas, pero sugiere que tal vez podamos reducir el inventario de clases léxicas. Para evaluar las clases colectivamente, comparamos los miembros, descriptores y conceptos asignados por REDES a cada clase léxica. En la próxima tabla, mostramos las seis clases léxicas que se vinculan con +ABILITY_00. Dividimos los miembros de cada clase léxica en el paradigma de base semántica y el de base pragmática o discursiva:

Tabla 255. Clases de REDES (y PRÁCTICO) vinculadas a +ABILITY_00

| Predicado | Descriptor - y CONCEPTO en REDES | Paradigma de base semántica | Paradigma de base pragmática o discursiva |
|----------------------|---|--|---|
| <i>malgastar</i> | facultades, aptitudes o cualidades humanas, esp. intelectivas; estados de reconocimiento personal- FACULTAD, CAPACIDAD | <i>talento, capacidad, don</i> | <i>cualidad, generosidad, valor, credibilidad, prestigio, experiencia</i> |
| <i>dilapidar</i> | facultades del individuo, relacionadas con la creación, conocimiento o capacidad de ejecución -FACULTAD, PROMINENCIA | <i>talento, capacidad</i> - <i>inteligencia, sapiencia</i> (subordinados) ⁹² | <i>idea, experiencia, intuición</i> |
| <i>malograr</i> | facultades físicas o mentales - FACULTAD | <i>talento, capacidad</i> - <i>inteligencia</i> (subordinado) | <i>voz</i> |
| <i>desaprovechar</i> | (FACULTAD, CAPACIDAD) ⁹³ | <i>talento, capacidad, facultad, facilidad</i> | <i>conocimiento, cualidad, disposición, carrera, potencial</i> |
| <i>despilfarrar</i> | (FACULTAD, CAPACIDAD) | <i>talento, -ingenio</i> (subordinado) | |
| <i>desperdiciar</i> | (FACULTAD, CAPACIDAD) | <i>talento, capacidad</i> | |

Es indiscutible que estas seis clases léxicas apuntan al concepto +ABILITY_00. Esto se consigna en el descriptor y concepto de cada una en REDES. Además, todas las clases contienen como pilar central argumentos que son lexicalizaciones ya vinculadas a +ABILITY_00 (*capacidad, don, facultad, talento*) en FunGramKB. Se trata de sustantivos cuyo significado inherente coincide con el concepto +ABILITY_00. Los demás argumentos advienen a las clases léxicas por procesos de extensión semántica. Ellos presentan los rasgos únicos de cada clase léxica, y, por lo tanto, del verbo que la selecciona. Por ejemplo, *malograr* es el único verbo vinculado a -WASTE_ABILITY cuyo descriptor explicita ‘facultades físicas’; de ahí, que la clase léxica con que se combina incluya *voz*. Por su parte, *dilapidar* se distingue por preferir las ‘facultades relacionadas con la creación o capacidad de

⁹² Marcamos como “subordinados” las lexicalizaciones de conceptos subordinados al concepto identificado.

⁹³ Marcamos entre paréntesis los conceptos asignados por la autora, a partir de los conceptos de REDES, en los casos en que REDES o PRÁCTICO no los asigna.

ejecución’, como *idea, experiencia, intuición*. Y *malgastar* especifica las facultades ‘intelectivas’ y los ‘estados de reconocimiento personal’; de ahí que se combine con *cualidad, generosidad, valor, credibilidad, prestigio, experiencia*. En los paradigmas discursivos o pragmáticos que no se pueden representar conceptualmente radica tal vez la idiosincrasia de cada verbo, o la idiosincrasia del español.

La entidad +TIME_00 y algunas entidades subordinadas fueron seleccionadas por seis verbos en su vínculo con +WASTE_00. Entre estas clases léxicas, hay un vínculo claro con la noción de ‘tiempo’, que se ve en los descriptores y en argumentos como *tiempo, momento*. En este grupo, sin embargo, destacan las nociones subordinadas a ‘tiempo’, como: *año, hora, futuro, día, mes, semana*. También llama la atención de la utilización de *vida* en tres clases léxicas distintas con el sentido de ‘tiempo’. Esto puede indicar que un aspecto que se activa regularmente en el sentido de *vida* es la noción de ‘duración de tiempo’, especialmente en el contexto de ‘malgastarla’.

Tabla 256. Clases léxicas de REDES (y PRÁCTICO) vinculadas a +TIME_00/+PERIOD_00/+FUTURE_00

| Predicado | Descriptor – y CONCEPTO en REDES | Paradigma de base semántica | Paradigma de base pragmática o discursiva |
|----------------------|---|---|---|
| <i>malgastar</i> | unidades temporales, incluidas las que se refieren a la existencia - TIEMPO | <i>día, juventud,</i> | <i>vida</i> |
| <i>dilapidar</i> | sustantivos temporales - TIEMPO | - <i>tiempo, momento,</i> <i>instante</i> - <i>historia, año, hora,</i> <i>futuro</i> (subordinados) | <i>carrera, vida</i> |
| <i>malograr</i> | <i>futuro</i> y coyuntura favorable - FUTURO | <i>futuro</i> | <i>oportunidad,</i> <i>posibilidad</i> |
| <i>desaprovechar</i> | (TIEMPO) | <i>tiempo, momento</i> | --- |
| <i>perder</i> | sustantivos temporales - TIEMPO | - <i>tiempo</i> - <i>día, mes, año</i> (subordinados) | --- |
| <i>desperdiciar</i> | (TIEMPO) | - <i>tiempo,</i> - <i>día, semana, juventud</i> (subordinados) | <i>porvenir,</i> <i>vacaciones, vida</i> |

La próxima entidad seleccionada es +MONEY_00, que es elegida por cinco verbos: *dilapidar, derrochar, disipar, despilfarrar, desperdiciar*. De las cinco clases, tres fueron clases estables, por lo que el vínculo entre estas clases y la entidad es bastante sólido.

Tabla 257. Clases de REDES (y PRÁCTICO) vinculadas a +MONEY_00

| Predicado | Descriptor – y CONCEPTO en REDES | Paradigma de base semántica | Paradigma de base pragmática discursiva |
|---------------------|-------------------------------------|---|---|
| <i>dilapidar</i> | bienes materiales - DINERO | <i>capital, dinero</i> | --- |
| <i>derrochar</i> | dinero u otros bienes - DINERO | <i>fortuna, dinero, riqueza</i> | --- |
| <i>disipar</i> | bienes materiales - DINERO | <i>fortuna, dinero,</i> | <i>hacienda, herencia</i> |
| <i>despilfarrar</i> | (DINERO) | - <i>dinero, fortuna</i> - <i>crédito, ganancia, presupuesto</i> (subordinados) | <i>millones, ayuda</i> |
| <i>desperdiciar</i> | (DINERO) | <i>dinero, riqueza</i> | --- |

Tres clases léxicas se vincularon a la entidad +STRENGTH_00. Todas comparten la lexicalización fundamental de +STRENGTH_00: *fuerza*. Al tener esta evidencia desplegada así, surge la pregunta de si *esfuerzo* debería vincularse también a +STRENGTH_00. Uno de los sentidos de *esfuerzo* es la aplicación de la *fuerza* hacia un objetivo –la relación con *fuerza* también es evidente etimológicamente–. Sin embargo, en FunGramKB no es tan fácil hacer constar la relación, porque *esfuerzo* se vincula con +EFFORT_00, que es subordinado de +PROCESS_00, mientras que +STRENGTH_00 es subordinado de +PHYSICAL_ATTRIBUTE_00. Por ahora, pues, optamos por dejar a *esfuerzo* entre las colocaciones de *malgastar* y *desaprovechar*.

Tabla 258. Clases de REDES (y PRÁCTICO) vinculadas a +STRENGTH_00

| Predicado | Descriptor – y CONCEPTO en REDES | Paradigma de base semántica | Paradigma de base pragmática discursiva |
|----------------------|--|--------------------------------------|---|
| <i>malgastar</i> | fuerza o energía, y sus efectos, literal y figurado (FUERZA) | <i>fuerza, energía,</i> | <i>esfuerzo, salud</i> |
| <i>desaprovechar</i> | (FUERZA) | <i>fuerza, energía, impulso,</i> | <i>esfuerzo</i> |
| <i>desperdiciar</i> | (FUERZA) | <i>fuerzas</i> | --- |

Tres clases léxicas se vincularon a la entidad +BULLET_00. Todas comparten dos lexicalizaciones de +BULLET_00: *bala, munición*. Lo específico de esta selección llama la atención, pero puede reflejar el hecho de que las municiones son consideradas recursos susceptibles de ser malgastados.

Tabla 259. Clases de REDES (y PRÁCTICO) vinculadas a +BULLET_00

| Predicado | Descriptor – y CONCEPTO en REDES | Paradigma de base semántica | Paradigma de base pragmática discursiva |
|---------------------|---|--------------------------------|---|
| <i>malgastar</i> | munición, acciones en las que se usa o formas en que se almacena (ARMA) | <i>munición, bala</i> | <i>disparo, cartuchera,</i> |
| <i>despilfarrar</i> | (ARMA) | <i>bala, munición</i> | --- |
| <i>desperdiciar</i> | (ARMA) | <i>bala, munición</i> | --- |

Finalmente, dos clases léxicas se vincularon a la entidad +FOOD_00. Ambas comparten dos lexicalizaciones de +FOOD_00: *comida, alimento*. La noción del alimento como recurso que puede ser malgastado no presenta ninguna sorpresa; por el contrario, sí llama la atención que solo dos verbos presenten esa preferencia de selección.

Tabla 260. Clases de REDES (y PRÁCTICO) vinculadas a +FOOD_00

| Predicado | Descriptor – y CONCEPTO en REDES | Paradigma de base semántica | Paradigma de base pragmática discursiva |
|---------------------|-------------------------------------|--------------------------------|---|
| <i>despilfarrar</i> | (ALIMENTO) | <i>comida, alimento</i> | --- |
| <i>desperdiciar</i> | (ALIMENTO) | <i>comida, alimento</i> | --- |

Hubo dos nociones que fueron seleccionadas repetidamente por los verbos en su vínculo con +WASTE_00, que no se pudieron vincular satisfactoriamente a ninguna entidad en FunGramKB (todas se clasificaron como tipo IV). La primera noción es ‘recurso’, que estuvo presente en seis clases léxicas elegidas por seis verbos. Como hemos comentado antes, esta noción es imposible de vincular con una entidad conceptual porque la mayoría de las palabras a las que se atribuye la noción de ‘recurso’ tienen sentidos rectos que estarían vinculados con otras entidades en la ontología. Sin embargo, la relación entre la noción de ‘malgastar’ y la noción de ‘recurso’ es fundamental: lo que se combine con +WASTE_00 automáticamente se convierte en un recurso. La tabla siguiente muestra las clases léxicas de REDES y PRÁCTICO que apuntan a esa noción:

Tabla 261. Clases de REDES (y PRÁCTICO) vinculadas a la noción de ‘recurso’

| Predicado | Descriptor - CONCEPTO en REDES | Miembros clase léxica |
|----------------------|---|---|
| <i>derrochar</i> | diversos recursos - RECURSO | <i>munición, agua</i> |
| <i>malgastar</i> | magnitudes, bienes materiales, recursos - RECURSO | <i>dinero, agua, corcho, petróleo, gasolina, papel</i> |
| <i>dilapidar</i> | recursos naturales; personas o grupos, cuando se asimilan a lo anterior - RECURSO | <i>agua, oxígeno; jugador, equipo</i> |
| <i>desaprovechar</i> | (RECURSO) | <i>recurso</i> |
| <i>despilfarrar</i> | (RECURSO) | <i>energía, agua, reservas</i> |
| <i>desperdiciar</i> | (RECURSO) | <i>sangre, luz, saliva, petróleo, agua, aceite, energía</i> |

Hay una segunda noción que expresan seis clases léxicas elegidas por seis verbos: ‘posibilidad’. Este caso es diferente del anterior, dado que hay tres palabras que se repiten en las clases léxicas y que lexicalizan la noción de ‘posibilidad’ por rasgos inherentes a su significado: *oportunidad, posibilidad, ocasión*. Estas deberían tener representación en la ontología. Las otras palabras miembros de las clases léxicas adquieren el valor de ‘posibilidad’ por procesos discursivos o pragmáticos, por ejemplo, *ventaja, renta, margen*. Evidentemente, hay una relación estrecha también entre la noción de ‘malgastar’ y la noción de ‘posibilidad’.

Tabla 262. Clases de REDES (y PRÁCTICO) vinculadas a la noción de ‘posibilidad’

| Predicado | Descriptor y (CONCEPTO) en REDES | Miembros clase léxica |
|----------------------|---|---|
| <i>perder</i> | circunstancias favorables | <i>oportunidad, posibilidad, ocasión</i> |
| <i>malgastar</i> | oportunidad u ocasión, ventaja o circunstancia favorable, frec. deportivo (POSIBILIDAD) | <i>oportunidad, ocasión, balón, ventaja, punto, opción, privilegio, posibilidad</i> |
| <i>dilapidar</i> | situación favorable para ejecutar o conseguir algo, frec. deportivo (POSIBILIDAD) | <i>oportunidad, situación, posibilidad, opción, baza, ventaja, renta, margen</i> |
| <i>desaprovechar</i> | (POSIBILIDAD) | <i>oportunidad, opción, ventaja, racha, visita, presencia, viaje</i> |
| <i>despilfarrar</i> | (POSIBILIDAD) | <i>oportunidad, posibilidad, ocasión</i> |
| <i>desperdiciar</i> | (POSIBILIDAD) | <i>oportunidad, ventaja, penalti, voto</i> |

Todos los resultados aquí analizados serán sintetizados en el capítulo IX, dedicado a las conclusiones de este trabajo.

CAPÍTULO VIII. RESULTADOS Y ANÁLISIS: POLISEMIA REGULAR Y COMBINATORIA LÉXICA

Introducción

En este capítulo, continuamos la presentación y análisis de los resultados. Nos enfocamos aquí en la representación conceptual de la polisemia regular en relación con la combinatoria léxica, en la base de conocimiento FunGramKB. Analizamos en detalle un grupo de cinco verbos relacionados semánticamente: *tejer*, *tramar*, *urdir*, *cocinar(se)* y *maquinar*, que se combinan con 32 clases léxicas en REDES, y que evidencian varios casos de polisemia regular.

1. TEJER, TRAMAR, URDIR, COCINAR(SE), MAQUINAR

Para explorar la representación de la polisemia regular en el marco de FunGramKB, seleccionamos un grupo de predicados semánticamente relacionados en los que percibimos una posible alternancia de sentidos compartida: *tejer*, *tramar*, *urdir*, *cocinar(se)*, *maquinar*. Estos cinco verbos se combinan con un total de 32 clases léxicas diferentes. Seguimos la misma metodología con estos verbos que con los nueve verbos vistos anteriormente.

1.1 TEJER

REDES distingue en *tejer* el sentido físico del sentido figurado, pero no les da contenido semántico. Nosotros definimos el sentido físico como ‘formar una tela u otra cosa con hilos u otro material’. Con respecto al sentido figurado, *tejer* se comporta en cierta manera como un verbo liviano que adquiere diferentes sentidos léxicos verbales según su combinación con complementos nominales. Su significado básico es ‘crear’, pero con las diferentes combinaciones asume los sentidos de ‘conspirar’, ‘escribir’, ‘explicar’, ‘sentir’. La tabla siguiente muestra la combinatoria de *tejer* que presenta REDES, correlacionada con los sentidos que le asignamos:

Tabla 263. Sentidos y combinatoria léxica de *tejer*, en REDES

| Sentido del verbo | CONCEPTO asignado a la CL en REDES | Descriptor de la CL en REDES | Argumentos miembros de la CL en REDES |
|---|------------------------------------|---|--|
| ‘formar una tela u otra cosa con hilos u otro material’ | --- | telas y otros materiales | <i>hilo, punto, mimbre, algodón, lana, seda, tela</i> |
| | --- | objetos formados con ellos [telas y otros materiales] | <i>cesta, estera, canasto, jersey, camisa, sombrero</i> |
| ‘conspirar’ | PROYECTO | plan o proyecto, a veces oculto y destinado a algún fin ilícito; también algunos que designan otras formas de maquinación | <i>estrategia, plan, proyecto, trampa</i> |
| ‘escribir’ | (TEXTO) | obras de creación, generalmente literarias, pero también de otro tipo | <i>textos, discurso, conversación, novela, narración, relato, comedia, espectáculo, canción</i> |
| | (ARGUMENTO) | argumento o trama, u otros contenidos articulados que se les asimilan | <i>intriga, trama, historia, argumento, guión</i> |
| ‘explicar’ | RAZONAMIENTO, PENSAMIENTO | diversas nociones, generalmente dotadas de cierta articulación, susceptibles de ser razonadas o elaboradas | <i>explicación, hipótesis, filosofía</i> |
| ‘desear’ | INCLINACIÓN | sustantivos de carácter prospectivo que designan lo que se anhela o lo que ha de venir | <i>esperanza, ilusión, deseo, sueño, destino</i> |
| ‘sentir’ | SENTIMIENTO | diversas sensaciones o sentimientos | <i>amargura, temor, hastío, ira, paz</i> |
| ‘crear’ | CONFUSIÓN | cosas de contextura o composición compleja o enrevesada, por extensión metafórica; también nociones caracterizadas generalmente por su complejidad o su confusión | <i>jeroglífico, enredo, urdimbre, maraña</i> |
| | CONFUSIÓN | red o entramado, en sentido físico y figurado | <i>red, entramado, telaraña, sistema</i> |
| | INCERTIDUMBRE | nociones caracterizadas por la inseguridad o la falta de certeza | <i>sospecha, suposición, especulación, conjetura, incertidumbre, interrogante, rumor, presunción</i> |
| | RELACIÓN, ACCIÓN CONCERTADA | diversas relaciones de asociación o vínculo, así como otros que las implican o las caracterizan | <i>relación, vínculo, unidad, consenso, coalición, pacto, asociación, fraternidad, confianza</i> |
| | ADVERSIDAD, CONFLICTO | estados de dificultad, inconveniencia o adversidad | <i>crisis, escándalo, tragedia, tensión</i> |

(a) tejer ‘formar una tela u otra cosa con hilos u otro material’

El primer sentido de *tejer* es ‘formar una tela u otra cosa con hilos u otro material’. Lo vinculamos con el evento +WEAVE_00, cuyo MT establece las siguientes preferencias de selección: el Tema es típicamente ‘humano’ y el Referente es típicamente ‘tela’ o ‘canasta’. El PS expresa que el Tema crea el Referente con un Instrumento, que típicamente es una ‘mano’ o una ‘máquina’, usando el ‘hilo’ como Medio, para un Beneficiario ‘humano’.

Tabla 264. Descripción de +WEAVE_00 en FunGramKB

| | |
|------------------------|---|
| Superordinado: | +CREATE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2: +CLOTH_00 ^ +BASKET_00)Referent |
| PS: | +(e1: +CREATE_00 (x1)Theme (x2)Referent (f1: +HAND_00 +MACHINE_00)Instrument (f2: +THREAD_00 ^ +STRING_00)Means (f3: +HUMAN_00)Beneficiary) |
| Descripción: | create a piece of cloth by interlacing strands of fabric, such as wool or cotton |
| Lexicalización: | <i>crochet, knit, weave</i> (inglés), <i>hilar, tejer, tricotar</i> (español) |

Este sentido de *tejer* ocurre con dos clases léxicas. La primera se vincula con +CLOTH_00; la próxima tabla presenta la comparación entre la clase léxica y la entidad conceptual. La clase léxica es del tipo II, porque aunque *tela* lexicaliza a +CLOTH_00, los otros miembros son subordinados: *algodón*[+COTTON_00], *lana*[+WOOL_00], *seda*[+SILK_00] o pueden serlo: *hilo, punto, mimbre*. Como esta preferencia de selección ya está representada en la descripción conceptual de +WEAVE_00, no hay que formular nuevos conceptos terminales o subconceptos.

Tabla 265. Comparación entre clase léxica de REDES y +CLOTH_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------------|---|--|
| Superordinado: | --- | +SOLID_00 |
| Concepto: | --- | +CLOTH_00 |
| PS: | --- | +(e1: +BE_00 (x1: +CLOTH_00)Theme (x2: +ARTIFICIAL_OBJECT_00 & +CORPUSCULAR_00 & +SOLID_00)Referent) |
| Descripción: | telas y otros materiales | something made by weaving or felting or knitting or crocheting natural or synthetic fibers |
| Manifestación léxica | <i>hilo, punto, mimbre, algodón, lana, seda, tela</i> | <i>cloth, fabric, material, textile</i> (inglés) |

La segunda clase léxica con que se combina *tejer*[+WEAVE_00] es del tipo IV: no se puede vincular a ninguna entidad de FunGramKB. El descriptor lee ‘objetos formados con telas y otros materiales’ y los miembros son: *cesta, estera, canasto, jersey, camisa, sombrero*. Dada la amplitud del descriptor y la diversidad de los miembros, esta clase no se puede vincular a una entidad, y sus miembros deberán ser representados como colocaciones de *tejer*[+WEAVE_00].

(b) *tejer* ‘conspirar’

El segundo sentido que asume *tejer* es ‘conspirar’. Lo vinculamos con el evento \$CONSPIRE_00, cuyo MT establece las preferencias de selección ‘humano’ para el Tema, y ‘pensamiento’ para el Referente. El PS establece que el Tema piensa el Referente, de una Manera ‘secreta’ y con Compañía ‘humana’; el Tema piensa en realizar un Referente no especificado con la Compañía; ese Referente no es ‘legal’.

Tabla 266. Descripción de \$CONSPIRE_00 en FunGramKB

| | |
|------------------------|---|
| Superordinado: | +THINK_00 |
| MT: | (x1: +HUMAN_00)Theme (x2:+THOUGHT_00)Referent |
| PS: | +(e1: +THINK_00 (x1)Theme (x2)Referent (f1: +SECRET_00)Manner (f2: +HUMAN_00)Company (f3: ((e2: +DO_00 (x1)Theme (x3)Referent (f2)Company) (e3: n +BE_01 (x3)Theme (x4: +LEGAL_00)Attribute))))Purpose) |
| Descripción: | to secretly plan with someone else_to do something illegal |
| Lexicalización: | <i>conspire, plot</i> (inglés), <i>conchabar, confabular, conspirar, maquinari, tramar</i> (español) |

Este sentido se activa en *tejer* cuando se combina con una clase léxica que vinculamos con +PLAN_00. Esta es una clase léxica del tipo III, porque tiene tres miembros que lexicalizan a la entidad (*estrategia, plan, proyecto*) y uno que adquiere ese sentido en combinación con *tejer* (*trampa*). Como +PLAN_00 es subordinado de +THOUGHT_00, no hay que cambiar las preferencias de selección ni crear un subconcepto nuevo; solo se añade *trampa* en la sección de colocaciones en la plantilla léxica de *tejer*[+CONSPIRE_00].

Tabla 267. Comparación entre clase léxica de REDES y +PLAN_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|-----------------------|---|--|
| Superordinado: | --- | +INFORMATION_00 |
| Concepto: | PROYECTO | +PLAN_00 |
| PS: | --- | +(e1: +BE_00 (x1: +PLAN_00)Theme (x2: +INFORMATION_00)Referent) *(e2: +DO_00 (x3)Theme (x4)Referent (f1: x1)Manner) |
| Descripción: | plan o proyecto, a veces oculto y destinado a algún fin ilícito; también algunos que designan otras formas de maquinación | a series of steps to be carried out or goals to be accomplished; "they drew up a six-step plan"; "they discussed plans for a new bond issue" |
| Manifestación léxica: | <i>estrategia, plan, proyecto, trampa</i> | <i>estrategia, plan, programa, proyecto (español), plan, programme, project, strategy (inglés)</i> |

(c) tejer ‘escribir’

El tercer sentido de *tejer* es ‘escribir’. Este sentido revela a *tejer* como un verbo liviano: inicialmente lo habíamos vinculado a ‘crear’ en combinación con clases léxicas que designan ‘textos’ e ‘historias’, pero luego reconocimos que ‘crear un texto’ es igual a ‘escribir’. Por lo tanto, lo vinculamos en FunGramKB con el evento +WRITE_00, que tiene como preferencia de selección +HUMAN_00 en el Tema, mientras que el Referente no está especificado. El PS de +WRITE_00 expresa que el Tema crea el Referente, usando como Instrumento que típicamente es un ‘lápiz’, ‘bolígrafo’ o ‘máquina’, en una Ubicación ‘superficie’ dada:

Tabla 268. Descripción de +WRITE_00 en FunGramKB

| | |
|-----------------|--|
| Superordinado: | +CREATE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2)Referent |
| PS: | +(e1: +CREATE_00 (x1)Theme (x2)Referent (f1: +PENCIL_00 ^ +PEN_00 ^ +MACHINE_00)Instrument (f2: +SURFACE_00)Location (f3: +ON_00)Position) |
| Descripción: | to produce writing or images on paper |
| Lexicalización: | <i>anotar, apuntar, autografiar, caligrafiar, componer, copiar, dactilografiar, dibujar, escribir, estampar (español) annotate, caricature, chalk, compose, copy, crayon, describe, imprint, jot down, note down, paint, pencil, print, record, run off, sign, stamp, take down, trace, type, write, write down (inglés)</i> |

Veamos las clases léxicas con que se combina *tejer* para adquirir este sentido. La primera clase léxica la vinculamos con +WRITING_00. Esta es una clase mixta, del tipo III: *texto* lexicaliza a la entidad, mientras que *novela, narración, relato, comedia, canción* son subordinados. Por otro lado, *discurso, conversación, espectáculo* adquieren el sentido en la combinación.

Tabla 269. Comparación entre clase léxica de REDES y +WRITING_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|-----------------------|---|---|
| Superordinado: | --- | +INFORMATION_OBJECT_00 |
| Concepto: | (TEXTO) | +WRITING_00 |
| PS: | --- | +(e1: +BE_00 (x1: +WRITING_00)Theme (x2: +INFORMATION_OBJECT_00)Referent) +(e2: +COMPRISE_00 (x1)Theme (x3: +LETTER_01 +NUMBER_00)Referent) *(e3: pos +READ_00 (x4: +HUMAN_00)Theme (x1)Referent) |
| Descripción: | obras de creación, generalmente literarias, pero también de otro tipo | reading matter; anything expressed in letters of the alphabet |
| Manifestación léxica: | <i>textos, discurso, conversación, novela, narración, relato, comedia, espectáculo, canción</i> | <i>escrito, texto</i> (español), <i>text, writing</i> (inglés) |

Si representáramos esta clase léxica como una preferencia de selección de +WRITE_00, tendríamos un subconcepto redundante llamado -WRITE_WRITING. Por otro lado, como el Referente de +WRITE_00 no está especificado, no tenemos representación conceptual explícita de las clases léxicas. Dado esto, optamos por ubicar los miembros de la clase léxica como colocaciones de *tejer*[+WRITE_00]. Así, aunque no tenemos la representación conceptual de la clase léxica, tenemos la representación léxica en forma de colocaciones. Esta codificación expresa que *tejer* se vinculará con +WRITE_00 cuando se combine con *textos, discurso, conversación, novela, narración, relato, comedia, espectáculo, canción*.

La próxima clase se vincula con +STORY_00. Es una clase estable (tipo I); todos los argumentos pueden lexicalizar a la entidad +STORY_00:

Tabla 270. Comparación entre clase léxica de REDES y +STORY_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|-----------------------|---|---|
| Superodinado: | -- | +SPEECH_00 ^ +WRITING_00 |
| Concepto: | (ARGUMENTO) | +STORY_00 |
| PS: | --- | +(e1: +BE_00 (x1: +STORY_00)Theme (x2: +SPEECH_00 +WRITING_00)Referent) +(e2: +COMPRISE_00 (x1)Theme (x3: +OCCURRENCE_00)Referent) |
| Descripción: | argumento o trama, u otros contenidos articulados que se les asimilan | a piece of fiction that narrates a chain of related events; "he writes stories for the magazines" |
| Manifestación léxica: | <i>intriga, trama, guión, historia, argumento</i> | <i>cuento, historia, relato</i> (español), <i>story, tale</i> (inglés) |

En este caso, para representar conceptualmente la preferencia de selección, formulamos la creación del subconcepto -WRITE_STORY. En la entrada léxica de *tejer*[-WRITE_STORY] no hay que especificar ninguna colocación, pues todos los miembros de la clase léxica están representados conceptualmente como +STORY_00.

(d) *tejer* ‘explicar’

El próximo sentido de *tejer* es ‘explicar’. Nuevamente *tejer* se comporta como verbo liviano: en vez de mantener el sentido de ‘crear explicación’, lo definimos como ‘explicar’, que se vincula con el evento conceptual terminal \$EXPLAIN_00. El MT de este concepto establece preferencias de selección por un Tema ‘humano’, un Referente no especificado y una Meta ‘humana’ o ‘animal’. El PS establece que el Tema le dice el Referente a la Meta, de una Manera ‘clara’.

Tabla 271. Descripción de \$EXPLAIN_00 en FunGramKB

| | |
|------------------------|---|
| Superordinado: | +SAY_00 |
| Concepto: | \$EXPLAIN_00 |
| MT: | (x1: +HUMAN_00)Theme (x2)Referent (x3: +HUMAN_00 ^ +ANIMAL_00)Goal |
| PS: | +(e1: +SAY_00 (x1)Theme (x2)Referent (x3)Goal (f1: +CLEAR_00)Manner) |
| Descripción: | give an explanation for |
| Lexicalización: | <i>explicar</i> (español), <i>account for, clarify, elucidate, exemplify, explain, explicate, expound, interpret</i> (inglés) |

Este sentido ocurre con una clase léxica mixta (tipo III): *explicación, hipótesis* pueden lexicalizar la noción de +EXPLANATION_00, mientras que *filosofía* adquiere ese sentido figuradamente.

Tabla 272. Comparación entre clase léxica de REDES y +EXPLANATION_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|------------------------------|--|---|
| Superordinado: | --- | +MESSAGE_00 |
| Concepto: | RAZONAMIENTO, PENSAMIENTO | +EXPLANATION_00 |
| PS: | --- | +(e1: +BE_00 (x1: +EXPLANATION_00)Theme (x2: +MESSAGE_00)Referent) +(e2: +EXPLAIN_00 (x3)Theme (x1)Referent (x4)Goal) |
| Descripción: | diversas nociones, generalmente dotadas de cierta articulación, susceptibles de ser razonadas o elaboradas | a statement that explains; "he launched into a detailed explanation" |
| Manifestación léxica: | <i>explicación, hipótesis, filosofía</i> | <i>aclaración, clarificación dilucidación, elucidación, explicación, explicación</i> (español), <i>account, clarification, elucidation, explanation, explication, exposition</i> (inglés) |

En este caso, establecer la preferencia de selección +EXPLANATION_00 para el Referente resultaría en un subconcepto redundante: -EXPLAIN_EXPLANATION. Por ello, para poder representar la clase léxica sin establecer una preferencia de selección conceptual, ubicamos los miembros de la clase léxica como colocaciones de *tejer*[+EXPLAIN_00].

(e) tejer ‘desear’

El próximo sentido de *tejer* es ‘desear’, que vinculamos al evento conceptual +DESIRE_01. Nuevamente, en vez de optar por codificar la preferencia de selección ‘crear deseo’, definimos *tejer* como ‘desear’. El MT de +DESIRE_01 establece la preferencia de selección del Tema como ‘humano’ y mantiene el Referente sin especificar:

Tabla 273. Descripción de +DESIRE_01 en FunGramKB

| | |
|------------------------|---|
| Superordinado: | +WANT_00 |
| Concepto: | +DESIRE_01 |
| MT: | (x1: +HUMAN_00)Theme (x2)Referent |
| PS: | +(e1: +WANT_00 (x1)Theme (x2)Referent (f1: +MUCH_00)Quantity) |
| Descripción: | feel or express a desire or hope concerning the future |
| Lexicalización: | <i>desear</i> (español), <i>desire, wish</i> (inglés) |

Tejer adquiere este sentido en combinación con una clase léxica que se vincula con la entidad +DESIRE_00. Es una clase mixta, del tipo III, aunque muy estable. Los argumentos *esperanza, ilusión, deseo, sueño* lexicalizan a +DESIRE_00, mientras que *destino* adquiere ese sentido figuradamente.

Tabla 274. Comparación entre clase léxica de REDES y +DESIRE_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|------------------------------|--|--|
| Superordinado: | --- | +FEELING_00 |
| Concepto: | INCLINACIÓN | +DESIRE_00 |
| PS: | --- | +(e1: +BE_00 (x1: +DESIRE_00)Theme (x2: +FEELING_00)Referent) *(e2: +FEEL_00 (x3: (e3: +DESIRE_01 (x4: +HUMAN_00)Theme (x5)Referent))Agent (x4)Theme (x1)Attribute) |
| Descripción: | sustantivos de carácter prospectivo que designan lo que se anhela o lo que ha de venir | a strong hope or wish |
| Manifestación léxica: | <i>esperanza, ilusión, deseo, sueño, destino</i> | <i>deseo</i> (español), <i>desire, wish</i> (inglés) |

En este caso, al igual que con *tejer*[+WRITE_00] y *tejer*[EXPLAIN], resultaría muy redundante crear el subconcepto -DESIRE_DESIRE. La solución es ubicar los miembros de la clase léxica como colocaciones de *tejer*[+DESIRE_01].

(f) *tejer* ‘sentir’

El próximo sentido de *tejer* es ‘sentir’, que vinculamos al evento +FEEL_00. Este evento conceptual tiene un Agente y un Atributo no especificados y un Tema con preferencia de selección ‘humano’:

Tabla 275. Descripción de +FEEL_00 en FunGramKB

| | |
|------------------------|--|
| Superordinado: | #EMOTION |
| Concepto: | +FEEL_00 |
| MT: | (x1)Agent (x2: +HUMAN_00)Theme (x3)Attribute |
| PS: | sp |
| Descripción: | seem with respect to the sensation given; of physical states, indicating as health, etc.: "My cold is gone--I feel fine today"; "She felt tired after the long hike" |
| Lexicalización: | <i>experimentar, sentir</i> (español), <i>experience, feel, sense</i> (inglés) |

El vínculo con +FEEL_00 ocurre con una clase que se describe por un concepto superordinado (tipo II). La entidad +FEELING_00 abarca a los subordinados *amargura*[+SADNESS_00], *temor*[+FEAR_00], *hastío*[+SADNESS_00], *ira*[+ANGER_00], *paz*[+PLEASURE_00].

Tabla 276. Comparación entre clase léxica de REDES y +FEELING_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|------------------------------|--|--|
| Superordinado: | --- | +PSYCHOLOGICAL_ATT_00 |
| Concepto: | SENTIMIENTO | +FEELING_00 |
| PS: | --- | +(e1: +BE_00 (x1: +FEELING_00)Theme (x2: +PSYCHOLOGICAL_ATT_00)Referent) +(e2: +FEEL_00 (x3)Agent (x4: +HUMAN_00)Theme (x1)Attribute) |
| Descripción: | diversas sensaciones o sentimientos | an emotion that you feel; emotional attribute. |
| Manifestación léxica: | <i>amargura, temor, hastío, ira, paz</i> | <i>ánimo, emoción, sentimiento</i> (español), <i>emotion, feeling, humour, mood,</i> |

| | | |
|--|--|---|
| | | <i>sentiment, spirit, temper</i> (inglés) |
|--|--|---|

Este caso se suma a los tres anteriores. Crear un subconcepto llamado -FEEL_FEELING, para dar cuenta de la preferencia de selección de *tejer*, resultaría muy redundante. Por lo tanto, no registramos conceptualmente la preferencia de selección, pero sí ubicamos los miembros de la clase léxica como colocaciones en *tejer*[+FEEL].

(g) *tejer* ‘crear’

El último sentido que presentamos de *tejer* es ‘crear’. Podemos decir que este es el sentido básico de *tejer* como verbo liviano. Lo vinculamos en FunGramKB con el evento +CREATE_00, que tiene como preferencia de selección ‘humano’ o ‘animal’ en el Tema, mientras que el Referente no está especificado. El PS lee: el Tema realiza el Referente, para un Beneficiario ‘humano’ o ‘animal’, con un Instrumento, y ese Referente en el pasado no existía.

Tabla 277. Descripción de +CREATE_00 en FunGramKB

| | |
|------------------------|---|
| Superordinado: | #CREATION & +DO_00 |
| MT: | (x1: +HUMAN_00 ^ +ANIMAL_00)Theme (x2)Referent |
| PS: | +((e1: +DO_00 (x1)Theme (x3)Referent (f1: +HUMAN_00 ^ +ANIMAL_00)Beneficiary (f2)Instrument (f3: x2)Result)(e2: rpast n +EXIST_00 (x2)Theme)) |
| Descripción: | bring into existence; "The company was created 25 years ago" |
| Lexicalización: | <i>create, make, produce</i> (inglés), <i>crear, hacer, producir</i> (español) |

La primera clase con que se combina *tejer* en el sentido de ‘crear’ es mixta (tipo III). El argumento *crisis* lexicaliza a +PROBLEM_00, mientras que *escándalo, tragedia, tensión* adquieren el sentido de ‘problema’ en combinación con *tejer*.

Tabla 278. Comparación entre clase léxica de REDES y +PROBLEM_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|------------------------------|--|--|
| Superodinado: | --- | +STATE_00 |
| Concepto: | ADVERSIDAD, CONFLICTO | +PROBLEM_00 |
| PS: | --- | +(e1: +BE_00 (x1: +PROBLEM_00)Theme (x2: +STATE_00)Referent) +(e2: pos n +DO_00 (x3: +HUMAN_00)Theme (x4)Referent (f1: (e3: +BE_01 (x4)Theme (x5: +DIFFICULT_00)Attribute))Reason (f2: x1)Result) |
| Descripción: | estados de dificultad, inconveniencia o adversidad | a factor causing trouble in achieving a positive result or tending to produce a negative result; "serious difficulties were encountered in obtaining a pure reagent" |
| Manifestación léxica: | <i>crisis, escándalo, tragedia, tensión</i> | <i>dificultad, problema</i> (español), <i>difficulty, problem, trouble</i> (inglés) |

A partir de esta selección léxica, formulamos la creación del subconcepto -CREATE_PROBLEM. Los argumentos *escándalo, tragedia, tensión* se ubican como colocaciones de *tejer*[+CREATE_PROBLEM].

El sentido de ‘crear’ ocurre también con cuatro clases léxicas que no se pueden vincular a ninguna entidad en FunGramKB (tipo IV). Hay dos clases que REDES asocia a la noción de CONFUSIÓN, y una a INCERTIDUMBRE, que no pudimos vincular en FunGramKB (tipo IV). La primera clase tiene el descriptor ‘red o entramado, en sentido físico y figurado’ y los miembros *red, entramado, telaraña, sistema*. La segunda tiene el descriptor ‘cosas de textura o composición compleja o enrevesada, por extensión metafórica; también nociones caracterizadas generalmente por su complejidad o su confusión’ y los miembros *jeroglífico, enredo, urdimbre, maraña*. La tercera, asociada a INCERTIDUMBRE, se describe como ‘nociones caracterizadas por la inseguridad o la falta de certeza’ y se compone de *sospecha, suposición, especulación, conjetura, incertidumbre, interrogante, rumor, presunción*. La entidad conceptual más cercana a estas tres clases es +MYSTERY_00 (‘something that baffles understanding and cannot be explained; "how it got out is a mystery"’). Sin embargo, aunque la ‘confusión’, la ‘incertidumbre’ y el ‘misterio’ son entidades que entrañan una dificultad para entender algo, la causa de la dificultad es diferente. En la ‘confusión’, la causa es un estado de complejidad o enredo, mientras que en

la ‘incertidumbre’ es una falta de certeza; por otro lado, en el ‘misterio’, la dificultad de entender es una condición intrínseca. Como no podemos vincular estas clases léxicas a ninguna entidad y, por lo tanto, no podemos representarlas conceptualmente, ubicamos los argumentos miembros de las clases léxicas como colocaciones de *tejer*[+CREATE].

La cuarta clase con que se combina *tejer*[+CREATE_00] tampoco se puede vincular a ninguna entidad (tipo IV). Tiene el descriptor ‘diversas relaciones de asociación o vínculo, así como otros que las implican o las caracterizan’. Sus miembros son *relación, vínculo, unidad, consenso, coalición, pacto, asociación, fraternidad, confianza*. REDES asocia esta clase a los conceptos RELACIÓN, ACCIÓN CONCERTADA. No encontramos en FunGramKB ninguna entidad que se acerque a la noción de ‘relación, vínculo, acuerdo’. Los miembros de esta clase léxica, por lo tanto, también se ubican como colocaciones en la plantilla léxica de *tejer*[+CREATE_00].

1.1.1 Resumen TEJER

La tabla siguiente muestra los vínculos de *tejer* en FunGramKB con eventos conceptuales en la ontología. También se identifican las colocaciones que se ubican en las plantillas léxicas de *tejer* en su vínculo con diferentes conceptos, recordando que estos son los argumentos que no están representados por las preferencias de selección del concepto:

Tabla 279. Resumen: TEJER en FunGramKB (vinculación propuesta)

| Evento conceptual | Colocaciones |
|-------------------|--|
| +WEAVE_00 | <i>cesta, estera, canasto, jersey, camisa, sombrero</i> |
| +CONSPIRE_00 | <i>trampa</i> |
| +WRITE_00 | <i>textos, discurso, conversación, novela, narración, relato, comedia, espectáculo, canción</i> |
| -WRITE_STORY | --- |
| +EXPLAIN_00 | <i>explicación, hipótesis, filosofía</i> |
| +DESIRE_00 | <i>esperanza, ilusión, deseo, sueño, destino</i> |
| +FEEL_00 | <i>amargura, temor, hastío, ira, paz</i> |
| +CREATE_00 | <i>red, entramado, telaraña, sistema; jeroglífico, enredo, urdimbre, maraña; relación, vínculo, unidad, consenso, coalición, pacto, asociación, fraternidad, confianza</i> |
| -CREATE_PROBLEM | <i>escándalo, tragedia, tensión</i> |

En la tabla siguiente, se muestra la vinculación de cada clase léxica con que se combina *tejer* con una entidad de FunGramKB, y su clasificación por tipos. Además, se

muestran los argumentos miembros de las clases léxicas que pertenecen al paradigma semántico y los que pertenecen a paradigmas discursivos o pragmáticos:

Tabla 280. Tipología de clases léxicas que se combinan con *tejer* con respecto a FunGramKB

| Tipo de clase | CONCEPTO en FunGramKB | CONCEPTO en REDES | Argumentos CL – paradigma semántico | Argumentos CL – paradigma discursivo o pragmático |
|---------------|-----------------------|-----------------------------|---|--|
| I | +STORY_00 | (ARGUMENTO) | <i>intriga, trama, historia, argumento, guión</i> | --- |
| II | +CLOTH_00 | --- | <i>hilo, punto, mimbre, algodón, lana, seda, tela</i> | --- |
| II | +FEELING_00 | SENTIMIENTO | <i>amargura, temor, hastío, ira, paz</i> | --- |
| III | +WRITING_00 | (TEXTO) | <i>novela, narración, relato, comedia, canción</i> | <i>textos, discurso, conversación, espectáculo</i> |
| III | +EXPLANATION_00 | RAZONAMIENTO, PENSAMIENTO | <i>explicación, hipótesis</i> | <i>filosofía</i> |
| III | +DESIRE_00 | INCLINACIÓN | <i>esperanza, ilusión, deseo, sueño</i> | <i>destino</i> |
| III | +PROBLEM_00 | ADVERSIDAD, CONFLICTO | <i>crisis</i> | <i>escándalo, tragedia, tensión</i> |
| III | +PLAN_00 | PROYECTO | <i>estrategia, plan, proyecto,</i> | <i>trampa</i> |
| IV | --- | --- | --- | <i>cesta, estera, canasto, jersey, camisa, sombrero</i> |
| IV | --- | CONFUSIÓN | --- | <i>red, entramado, telaraña, sistema</i> |
| IV | --- | RELACIÓN, ACCIÓN CONCERTADA | --- | <i>relación, vínculo, unidad, consenso, coalición, pacto, asociación, fraternidad, confianza</i> |
| IV | --- | CONFUSIÓN | --- | <i>jeroglífico, enredo, urdimbre, maraña</i> |
| IV | --- | INCERTIDUMBRE | --- | <i>sospecha, suposición, especulación, conjetura, incertidumbre, interrogante, rumor, presunción</i> |

1.2. TRAMAR

REDES distingue para el verbo *tramar* un solo sentido, ‘urdir’, o sea, ‘conspirar’. Nosotros distinguimos en los datos de REDES un segundo sentido, ‘pensar’. Y para efectos del estudio de polisemia regular, añadimos el sentido etimológico, que no está representado en la combinatoria de REDES, pero sí en los diccionarios tradicionales: ‘atravesar los hilos de la trama por entre los de la urdimbre, para tejer alguna tela’.

La tabla siguiente muestra la combinatoria léxica de *tramar*, distribuida según los dos sentidos ‘urdir’ y ‘pensar’:

Tabla 281. Sentidos y combinatoria léxica de *tramar*

| Sentidos del verbo | CONCEPTO asignado a la CL en REDES | Descriptor de la CL en REDES | Argumentos miembros de la CL en REDES |
|--------------------|------------------------------------|--|---|
| ‘urdir’ | (AGRESIÓN) | diversas formas de agresión | <i>atentado, asesinato, secuestro, venganza, golpe de estado</i> |
| | ACCIÓN HOSTIL | acción oculta o encubierta, con objetivos a menudo ilegítimos | <i>conspiración, complot, conjura, estafa, ardid, argucia</i> |
| | PROYECTO | estrategia, intención o modelo organizado de actuación | <i>plan, estrategia, operación, táctica, proyecto, alternativa</i> |
| | OPOSICIÓN | diversas formas de levantamiento, irrupción o manifestación activa contra algo o alguien | <i>acción, levantamiento, revuelta, insurrección, acción, campaña</i> |
| ‘pensar’ | POSIBILIDAD | idea o posibilidad, generalmente no contrastada | <i>idea, tesis, conjetura</i> |

(a) *tramar* ‘atravesar los hilos de la trama por entre los de la urdimbre, para tejer alguna tela’

Este sentido de *tramar* no está representado en REDES, pero aparece en los diccionarios consultados. Para efectos de examinar la polisemia regular en este grupo de verbos, reconocimos este sentido, y lo vinculamos con +WEAVE_00. El MT y el PS de +WEAVE_00 se describieron arriba (en la sección 1.1a):

Tabla 282. Descripción de +WEAVE_00 en FunGramKB

| | |
|------------------------|---|
| Superordinado: | +CREATE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2: +CLOTH_00 ^ +BASKET_00)Referent |
| PS: | +(e1: +CREATE_00 (x1)Theme (x2)Referent (f1: +HAND_00 +MACHINE_00)Instrument (f2: +THREAD_00 ^ +STRING_00)Means (f3: +HUMAN_00)Beneficiary) |
| Descripción: | create a piece of cloth by interlacing strands of fabric, such as wool or cotton |
| Lexicalización: | <i>crochet, knit, weave</i> (inglés), <i>hilar, tejer, tricotar</i> (español) |

(b) tramar ‘pensar’

El segundo sentido que identificamos para *tramar* es ‘pensar’. Este lo vinculamos con +THINK_00, que especifica el Tema como ‘humano’ y el Referente como ‘pensamiento’. El PS expresa que el Tema ‘humano’ usa un Referente que típicamente es su ‘cerebro’ con la Meta de crear en el Tema (‘humano’) el Referente (‘pensamiento’):

Tabla 283. Descripción de +THINK_00 en FunGramKB

| | |
|------------------------|--|
| Superordinado: | #COGNITION |
| MT: | (x1: +HUMAN_00)Theme (x2: +THOUGHT_00)Referent |
| PS: | +(e1: +USE_00 (x1)Theme (x3: +BRAIN_00)Referent (f1: (e2: +CREATE_00 (x1)Theme (x2)Referent))Goal) |
| Descripción: | use or exercise the mind |
| Lexicalización: | <i>discurrir, pensar, planificar, razonar</i> (español), <i>cogitate, plan, reason, think</i> (inglés) |

Este sentido ocurre con una clase léxica que vinculamos con +THOUGHT_00. Se trata de una clase estable (tipo I): todos los miembros son lexicalizaciones de +THOUGHT_00:

Tabla 284. Comparación entre clase léxica de REDES y +THOUGHT_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|-----------------------|---|---|
| Superordinado: | --- | #PROPOSITION |
| Concepto: | POSIBILIDAD | +THOUGHT_00 |
| PS: | --- | sp |
| Descripción: | idea o posibilidad, generalmente no contrastada | any cognitive content |
| Manifestación léxica: | <i>idea, tesis, conjetura</i> | <i>concepto, idea, mente, pensamiento, reflexión (español), concept, consideración, idea, meditation, notion, relection, thought (inglés)</i> |

Como esta preferencia de selección está representada en el MT de +THINK_00, es suficiente con vincular *tramar* a +THINK_00.

(c) *tramar* ‘urdir’

El tercer sentido de *tramar* es ‘urdir’ o ‘conspirar’, que vinculamos con el evento conceptual +CONSPIRE_00, cuyo MT y PS discutimos arriba (en la sección 1.1b).

Tabla 285. Descripción de \$CONSPIRE_00 en FunGramKB

| | |
|-----------------|---|
| Superordinado: | +THINK_00 |
| MT: | (x1: +HUMAN_00)Theme (x2:+THOUGHT_00)Referent |
| PS: | +(e1: +THINK_00 (x1)Theme (x2)Referent (f1: +SECRET_00)Manner (f2: +HUMAN_00)Company (f3: ((e2: +DO_00 (x1)Theme (x3)Referent (f2)Company) (e3: n +BE_01 (x3)Theme (x4: +LEGAL_00)Attribute))))Purpose) |
| Descripción: | to secretly plan with someone else_to do something illegal |
| Lexicalización: | <i>conspire, plot</i> (inglés), <i>conchabar, confabular, conspirar, maquinari, tramar</i> (español) |

Este sentido ocurre con la mayoría de las clases léxicas que identifica REDES para *tramar*. Veamos la vinculación de estas clases con entidades conceptuales en FunGramKB. La primera clase léxica la vinculamos con +CRIME_00. Se trata de una clase descrita por un concepto superordinado (tipo II), porque todos sus miembros son tipos de crímenes: *atentado, asesinato, secuestro, venganza, golpe de estado*.

Tabla 286. Comparación entre clase léxica de REDES y +CRIME_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|-----------------------|--|--|
| Superodinado: | --- | +OCCURRENCE_00 |
| Concepto: | (AGRESIÓN) | +CRIME_00 |
| PS: | --- | +(e1: +BE_00 (x1: +CRIME_00)Theme (x2: +OCCURRENCE_00)Referent) +(e2: n +BE_01 (x1)Theme (x3: +LEGAL_00)Attribute) |
| Descripción: | diversas formas de agresión | an act punishable by law; usually considered an evil act |
| Manifestación léxica: | <i>atentado, asesinato, secuestro, venganza, golpe de estado</i> | <i>delito, ilegalidad, infracción, quebrantamiento, quebranto, transgresión, vulneración (español), crime, law-breaking (inglés)</i> |

Sería redundante crear el subconcepto –CONSPIRE_CRIME, porque la condición de ‘ilegalidad’ está presente tanto en +CONSPIRE_00 como en +CRIME_00. Por eso, ubicamos los miembros de la clase léxica como colocaciones de *tramar*[+CONSPIRE_00].

Hay dos clases léxicas que se combinan con *tramar*[+CONSPIRE_00], que vinculamos con la entidad conceptual +PLAN_00. La primera es una clase mixta (tipo III). La mayoría de sus miembros puede lexicalizar a +PLAN_00, aunque es cierto que contienen un valor adicional de ‘intención ilegal’: *conspiración, complot, conjura, estafa, ardid*. Por otro lado, *argucia* adquiere el sentido de ‘plan’ de manera figurada. La segunda clase léxica también es mixta (tipo III): *plan, estrategia, táctica, proyecto* lexicalizan la entidad +PLAN_00, mientras que *operación, alternativa* adquieren el sentido de manera figurada.

Tabla 287. Comparación entre clase léxica de REDES y +PLAN_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|------------------------------|--|--|
| Superodinado: | --- | +INFORMATION_00 |
| Concepto: | ACCIÓN HOSTIL// PROYECTO | +PLAN_00 |
| PS: | --- | +(e1: +BE_00 (x1: +PLAN_00)Theme (x2: +INFORMATION_00)Referent) *(e2: +DO_00 (x3)Theme (x4)Referent (f1: x1)Manner) |
| Descripción: | acción oculta o encubierta, con objetivos a menudo ilegítimos // estrategia, intención o modelo organizado de actuación | a series of steps to be carried out or goals to be accomplished; "they drew up a six-step plan"; "they discussed plans for a new bond issue" |
| Manifestación léxica: | <i>conspiración, complot, conjura, estafa, ardid, argucia // plan, estrategia, operación, táctica, proyecto, alternativa</i> | <i>estrategia, plan, programa, proyecto</i> (español), <i>plan, programme, project, strategy</i> (inglés) |

Dado que \$CONSPIRE_00 incluye la preferencia de selección +THOUGHT_00 para el Referente, y +PLAN_00 es subordinado de +THOUGHT_00, no es necesario crear un subconcepto para especificar esa preferencia de selección. Ubicaremos *argucia, operación, alternativa* como colocaciones de *tramar*[+CONSPIRE_00].

La próxima clase léxica que se combina con *tramar*[+CONSPIRE_00] la vinculamos con +VIOLENCE_00. Se trata de una clase mixta, del tipo III: *levantamiento, revuelta, insurrección* son lexicalizaciones de +VIOLENCE_00, y *acción* y *campana* adquieren el sentido figuradamente.

Tabla 288. Comparación entre clase léxica de REDES y +VIOLENCE_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|-----------------------|--|---|
| Superodinado: | --- | +OCCURRENCE_00 |
| Concepto: | OPOSICIÓN | +VIOLENCE_00 |
| PS: | --- | +(e1: +BE_00 (x1: +VIOLENCE_00)Theme (x2: +OCCURRENCE_00)Referent) +(e2: +ATTACK_00 (x3: +HUMAN_00)Theme (x4)Referent (f1: x1)Scene) |
| Descripción: | diversas formas de levantamiento, irrupción o manifestación activa contra algo o alguien | an act of aggression (as one against a person who resists); "he may accomplish by craft in the long run what he cannot do by force and violence in the short one" |
| Manifestación léxica: | <i>acción, levantamiento, revuelta, insurrección, campaña</i> | <i>ataque, brutalidad, fuerza, violencia</i> (español), <i>attack, brutality, force, violence</i> (inglés) |

En este caso, sí es necesario especificar la preferencia de selección +VIOLENCE_00, así que formulamos la creación del subconcepto -CONSPIRE_VIOLENCE. En la entrada léxica del *tramar*[-CONSPIRE_VIOLENCE] se añaden como colocaciones *acción* y *campaña*.

1.2.1 Resumen TRAMAR

La tabla siguiente muestra los vínculos de *tramar* en FunGramKB con eventos conceptuales en la ontología. También se identifican las colocaciones que se ubican en las plantillas léxicas de *tramar* en su vínculo con diferentes conceptos, recordando que se trata de los argumentos que no se pudieron vincular a las preferencias de selección conceptuales:

Tabla 289. Resumen: TRAMAR en FunGramKB (vinculación propuesta)

| Evento conceptual | Colocaciones |
|--------------------|---|
| +WEAVE_00 | --- |
| +THINK_00 | --- |
| +CONSPIRE | <i>atentado, asesinato, secuestro, venganza, golpe de estado, argucia, operación, alternativa</i> |
| -CONSPIRE_VIOLENCE | <i>acción, campaña</i> |

La tabla siguiente muestra la vinculación de cada clase léxica seleccionada por *tramar* con una entidad de FunGramKB, y su clasificación por tipos. Además, se muestran los argumentos miembros de las clases léxicas que pertenecen al paradigma semántico y los que pertenecen a paradigmas discursivos o pragmáticos:

Tabla 290. Tipología de clases léxicas que se combinan con *tramar* con respecto a FunGramKB

| Tipo de clase | CONCEPTO en REDES | CONCEPTO en FunGramKB | Argumentos CL – paradigma semántico | Argumentos CL – paradigma discursivo o pragmático |
|---------------|-------------------|---------------------------|--|---|
| I | +THOUGHT_00 | POSIBILIDAD | <i>idea, tesis, conjetura</i> | --- |
| II | +CRIME_00 | (AGRESIÓN) | <i>atentado, asesinato, secuestro, venganza, golpe de estado</i> | --- |
| III / III | +PLAN_00 | ACCIÓN HOSTIL // PROYECTO | <i>conspiración, complot, conjura, estafa, ardid,</i> | <i>argucia</i> |
| III | +VIOLENCE_00 | OPOSICIÓN | <i>levantamiento, revuelta, insurrección,</i> | <i>acción, campaña</i> |

1.3. URDIR

REDES no distingue ningún sentido para *urdir*. A la luz de la combinatoria, nosotros identificamos los sentidos ‘escribir’, ‘explicar’ y ‘conspirar’. Al igual que *tejer*, *urdir* tiene rasgos de verbo liviano, que le permiten asumir el sentido léxico verbal de sus complementos nominales. Aparte de esos sentidos que se desprenden de la combinatoria, para efectos del estudio sobre polisemia regular, añadimos el sentido etimológico de *urdir*: ‘preparar los hilos en la urdidera para pasarlos al telar’. La combinatoria que se presenta en REDES se reproduce a continuación:

Tabla 291. Combinatoria léxica de *urdir*, en REDES

| Sentido del verbo | CONCEPTO asignado a la CL en REDES | Descriptor de la CL en REDES | Argumentos miembros de la CL en REDES |
|-------------------|------------------------------------|---|--|
| ‘escribir’ | TEXTO | diversos géneros narrativos, más frecuentemente el relato; también múltiples objetos de información que se asocian con ellos | <i>historia, relato, intriga, narración, argumento, libro, crónica, carta, anécdota</i> |
| ‘explicar’ | RAZONAMIENTO | razonamientos diversos, más frecuentemente si se proponen para dar sentido a algo | <i>teoría, hipótesis, interpretación, razonamiento, explicación, comentario</i> |
| ‘acordar’ | ACUERDO | acuerdo o afinidad; también otras relaciones de proximidad y diversos resultados de las acciones concertadas | <i>acuerdo, pacto, arreglo, eje, amistad, intercomunidad</i> |
| ‘conspirar’ | CONFUSIÓN | acciones concebidas generalmente para ocasionar algún perjuicio a personas o cosas, más o menudo si se realizan con ardides o diversos artificios encubiertos | <i>plan, trama, conspiración, complot, intriga, operación, estrategia, trampa, maniobra, artimaña, estratagema, conjura, jugada, celada, chantaje, enredo, truco</i> |
| | AGRESIÓN, ACCIÓN HOSTIL | ataque o agresión, u otras acciones ofensivas | <i>venganza, crimen, agresión, ataque, acusación, atentado, golpe de estado, secuestro, amenaza</i> |
| | ENGAÑO | falsedad o engaño, así como ciertas acciones que conllevan tales prácticas | <i>mentira, embuste, infamia, estafa, chanchullo, falsificación, ficción, apaño, patraña, equívoco, artificio</i> |

(a) *urdir* ‘preparar los hilos en la urdidera para pasarlos al telar’

El primer sentido que asignamos a *urdir*, que no está representado en REDES pero sí en los diccionarios, es ‘preparar los hilos en la urdidera para pasarlos al telar’. Este sentido lo vinculamos con el evento conceptual +WEAVE_00, cuyo MT y PS se describieron arriba (en la sección 1.1a):

Tabla 292. Descripción de +WEAVE_00 en FunGramKB

| | |
|------------------------|---|
| Superordinado: | +CREATE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2: +CLOTH_00 ^ +BASKET_00)Referent |
| PS: | +(e1: +CREATE_00 (x1)Theme (x2)Referent (f1: +HAND_00 +MACHINE_00)Instrument (f2: +THREAD_00 ^ +STRING_00)Means (f3: +HUMAN_00)Beneficiary) |
| Descripción: | create a piece of cloth by interlacing strands of fabric, such as wool or cotton |
| Lexicalización: | <i>crochet, knit, weave</i> (inglés), <i>hilar, tejer, tricotar</i> (español) |

(b) *urdir* ‘escribir’

El verbo *urdir* adquiere el sentido de ‘escribir’ en combinación con ciertas clases léxicas. El sentido de ‘escribir’ lo vinculamos con el evento +WRITE_00, cuyo MT y PS se discutieron arriba, en 1.1c:

Tabla 293. Descripción de +WRITE_00 en FunGramKB

| | |
|------------------------|---|
| Superordinado: | +CREATE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2)Referent |
| PS: | +(e1: +CREATE_00 (x1)Theme (x2)Referent (f1: +PENCIL_00 ^ +PEN_00 ^ +MACHINE_00)Instrument (f2: +SURFACE_00)Location (f3: +ON_00)Position) |
| Descripción: | to produce writing or images on paper |
| Lexicalización: | <i>anotar, apuntar, autografiar, caligrafiar, componer, copiar, dactilografiar, dibujar, escribir, estampar</i> (español) <i>annotate, caricature, chalk, compose, copy, crayon, describe, imprint, jot down, note down, paint, pencil, print, record, run off, sign, stamp, take down, trace, type, write, write down</i> (inglés) |

Este sentido ocurre con una clase léxica que vinculamos a la entidad conceptual +STORY_00. Es una clase mixta, del tipo III, porque contiene miembros que lexicalizan a +STORY_00: *historia, relato, narración, crónica, anécdota, intriga, argumento*, y otras que se están cerca semánticamente, pero que adquieren el sentido de +STORY_00 en combinación con *urdir*: *libro, carta*.

Tabla 294. Comparación entre clase léxica de REDES y +STORY_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|-----------------------|--|---|
| Superordinado: | --- | +SPEECH_00 ^ +WRITING_00 |
| Concepto: | TEXTO | +STORY_00 |
| PS: | --- | +(e1: +BE_00 (x1: +STORY_00)Theme (x2: +SPEECH_00 +WRITING_00)Referent) +(e2: +COMPRISE_00 (x1)Theme (x3: +OCCURRENCE_00)Referent) |
| Descripción: | diversos géneros narrativos, más frecuentemente el relato; también múltiples objetos de información que se asocian con ellos | a piece of fiction that narrates a chain of related events; "he writes stories for the magazines" |
| Manifestación léxica: | <i>historia, relato, intriga, narración, argumento, libro, crónica, carta, anécdota</i> | <i>cuento, historia, relato</i> (español), <i>story, tale</i> (inglés) |

Para representar conceptualmente la preferencia de selección, creamos el subconcepto *urdir*[-WRITE_STORY]. Colocamos como colocaciones en la plantilla léxica de *urdir*[-WRITE_STORY] a los argumentos: *libro, carta*.

(c) *urdir* ‘explicar’

El próximo sentido de *urdir* es ‘explicar’, que vinculamos a +EXPLAIN_00, cuyos MT y PS se explicaron en 1.1d:

Tabla 295. Descripción de \$EXPLAIN_00 en FunGramKB

| | |
|-----------------|---|
| Superordinado: | +SAY_00 |
| Concepto: | \$EXPLAIN_00 |
| MT: | (x1: +HUMAN_00)Theme (x2)Referent (x3: +HUMAN_00 ^ +ANIMAL_00)Goal |
| PS: | +(e1: +SAY_00 (x1)Theme (x2)Referent (x3)Goal (f1: +CLEAR_00)Manner) |
| Descripción: | give an explanation for |
| Lexicalización: | <i>explicar</i> (español), <i>account for, clarify, elucidate, exemplify, explain, explicate, expound, interpret</i> (inglés) |

Este sentido ocurre en combinación con una clase léxica que se vincula con +EXPLANATION_00. Se trata de una clase estable (tipo I): todos los argumentos miembros pueden lexicalizar a la entidad +EXPLANATION_00.

Tabla 296. Comparación entre clase léxica de REDES y +EXPLANATION_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|------------------------------|---|--|
| Superordinado: | --- | +MESSAGE_00 |
| Concepto: | RAZONAMIENTO | +EXPLANATION_00 |
| PS: | --- | +(e1: +BE_00 (x1: +EXPLANATION_00)Theme (x2: +MESSAGE_00)Referent) +(e2: +EXPLAIN_00 (x3)Theme (x1)Referent (x4)Goal) |
| Descripción: | razonamientos diversos, más frecuentemente si se proponen para dar sentido a algo | a statement that explains; "he launched into a detailed explanation" |
| Manifestación léxica: | <i>teoría, hipótesis, interpretación, razonamiento, explicación, comentario</i> | <i>aclaramción, clarificación, dilucidación, elucidación, explicación, explicación (español), account, clarification, elucidation, explanation, explication, exposition (inglés)</i> |

Como el subconcepto -EXPLAIN_EXPLANATION resulta redundante, vinculamos a *urdir* a +EXPLAIN_00 y ubicamos todos los argumentos de la clase léxica como colocaciones.

(d) *urdir* ‘acordar’

El próximo sentido de *urdir* es ‘acordar’, que vinculamos con el evento conceptual +AGREE_00. El MT establece como preferencias de selección ‘humano’ para el Tema y para la Meta; el Referente no se especifica:

Tabla 297. Descripción de +AGREE_00 en FunGramKB

| | |
|------------------------|--|
| Superordinado: | +SAY_00 |
| MT: | (x1: +HUMAN_00)Theme (x2)Referent (x3: +HUMAN_00)Goal |
| PS: | +(e1: +SAY_00 (x1)Theme (x2)Referent (x3)Goal) +(e2: +THINK_00 (x1)Theme (x4: (e3: +BE_01 (x2)Theme (x5: +TRUE_00)Attribute))Referent) +(e4: +THINK_00 (x3)Theme (x4)Referent) |
| Descripción: | to have the same opinion, or to accept a suggestion or idea |
| Lexicalización: | <i>estar de acuerdo</i> (español), <i>accede, accord, acquiesce, admit, agree, allow, assent, concur, consent, grant</i> (inglés) |

Este sentido ocurre con una clase léxica del tipo IV, que no pudimos vincular con ninguna entidad en FunGramKB. REDES le asigna el concepto ACUERDO, y el descriptor: ‘acuerdo o afinidad; también otras relaciones de proximidad y diversos resultados de las acciones concertadas’. Sus miembros son *acuerdo, pacto, arreglo, eje, amistad, intercomuni3n*. En este caso, en la entrada léxica de *urdir*[+AGREE_00] deberán ubicarse todos los miembros de la clase léxica como colocaciones.

(e) *urdir* ‘conspirar’

El próximo sentido de *urdir* es ‘conspirar’, que vinculamos con el evento conceptual terminal \$CONSPIRE_00:

Tabla 298. Descripción de \$CONSPIRE_00 en FunGramKB

| | |
|------------------------|--|
| Superordinado: | +THINK_00 |
| MT: | (x1: +HUMAN_00)Theme (x2:+THOUGHT_00)Referent |
| PS: | +(e1: +THINK_00 (x1)Theme (x2)Referent (f1: +SECRET_00)Manner (f2: +HUMAN_00)Company (f3: ((e2: +DO_00 (x1)Theme (x3)Referent (f2)Company) (e3: n +BE_01 (x3)Theme (x4: +LEGAL_00)Attribute)))Purpose) |
| Descripción: | to secretly plan with someone else_to do something illegal |
| Lexicalización: | <i>conspire, plot</i> (inglés), <i>conchabar, confabular, conspirar, maquinari, tramar</i> (español) |

Este sentido se da con una clase léxica que desdoblamos en dos entidades conceptuales. Vinculamos la primera con la entidad +PLAN_00, que es lexicalizada por los argumentos *conspiraci3n, complot, estrategia, stratagema, conjura*. Los demás argumentos

los adjudicamos a una clase del tipo IV, que no se puede vincular a ninguna entidad: *intriga, operación, trampa, maniobra, truco, artimaña, jugada, celada, chantaje, enredo*.

Tabla 299. Comparación entre clase léxica de REDES y +PLAN_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|------------------------------|---|--|
| Superordinado: | --- | +INFORMATION_00 |
| Concepto: | CONFUSIÓN | +PLAN_00 |
| PS: | --- | +(e1: +BE_00 (x1: +PLAN_00)Theme (x2: +INFORMATION_00)Referent) *(e2: +DO_00 (x3)Theme (x4)Referent (f1: x1)Manner) |
| Descripción: | acciones concebidas generalmente para ocasionar algún perjuicio a personas o cosas, más o menudo si se realizan con ardides o diversos artificios encubiertos | a series of steps to be carried out or goals to be accomplished; "they drew up a six-step plan"; "they discussed plans for a new bond issue" |
| Manifestación léxica: | <i>conspiración, complot, intriga, operación, estrategia, trampa, maniobra, truco, artimaña, estratagema, conjura, jugada, celada, chantaje, enredo</i> | <i>estrategia, plan, programa, proyecto (español), plan, programme, project, strategy (inglés)</i> |

Como hemos visto antes, no es necesario especificar la preferencia de selección de \$CONSPIRE_00 con +PLAN_00, porque +THOUGHT_00 es un concepto superordinado. Solo tenemos que añadir los argumentos no representados conceptualmente *–intriga, operación, trampa, maniobra, truco, artimaña, jugada, celada, chantaje, enredo–* a las colocaciones de *urdir*[\$CONSPIRE_00].

La próxima clase léxica con que se combina *urdir*[+CONSPIRE_00] no se puede vincular con ninguna entidad de FunGramKB (tipo IV). REDES la asocia al concepto ENGAÑO y le da el descriptor ‘falsedad o engaño, así como ciertas acciones que conllevan tales prácticas’. Sus miembros incluyen *mentira, embuste, infamia, estafa, chanchullo, falsificación, ficción, apaño, patraña, equívoco, artificio*. No hay ninguna entidad conceptual en la ontología de FunGramKB que exprese la noción de ‘mentira, falsedad, engaño’. Por lo tanto, en la entrada de *urdir*[+CONSPIRE], añadimos como colocaciones a los argumentos: *mentira, embuste, infamia, estafa, chanchullo, falsificación, ficción, apaño, patraña, equívoco, artificio*.

La próxima clase léxica se vincula con la entidad +VIOLENCE_00. Se trata de una clase léxica descrita por un superordinado (tipo II): *ataque, atentado* lexicalizan a +VIOLENCE_00, mientras que *venganza, crimen, agresión, acusación, golpe de estado, secuestro, amenaza* son subordinados de ese concepto.

Tabla 300. Comparación entre clase léxica de REDES y +VIOLENCE_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|-----------------------|---|---|
| Superordinado: | --- | +OCCURRENCE_00 |
| Concepto: | AGRESIÓN, ACCIÓN HOSTIL | +VIOLENCE_00 |
| PS: | --- | +(e1: +BE_00 (x1: +VIOLENCE_00)Theme (x2: +OCCURRENCE_00)Referent) +(e2: +ATTACK_00 (x3: +HUMAN_00)Theme (x4)Referent (f1: x1)Scene) |
| Descripción: | ataque o agresión, u otras acciones ofensivas | an act of aggression (as one against a person who resists); "he may accomplish by craft in the long run what he cannot do by force and violence in the short one" |
| Manifestación léxica: | <i>venganza, crimen, agresión, ataque, acusación, atentado, golpe de estado, secuestro, amenaza</i> | <i>ataque, brutalidad, fuerza, violencia</i> (español), <i>attack, brutality, force, violence</i> (inglés) |

A raíz de esta combinación, formulamos la creación del subconcepto - CONSPIRE_VIOLENCE. En la entrada léxica de *urdir*[-CONSPIRE_VIOLENCE], no hay que añadir colocaciones, porque la clase léxica está representada conceptualmente.

1.3.1 Resumen URDIR

La tabla siguiente muestra los vínculos de *urdir* con eventos conceptuales en la ontología de FunGramKB. También se identifican las colocaciones que se ubican en las plantillas léxicas de *urdir* en su vínculo con diferentes conceptos:

Tabla 301. Resumen: URDIR en FunGramKB (vinculación propuesta)

| Evento conceptual | Colocaciones |
|--------------------------|--|
| +WEAVE_00 | --- |
| -WRITE_STORY | <i>libro, carta</i> |
| +EXPLAIN_00 | <i>teoría, hipótesis, interpretación, razonamiento, explicación, comentario</i> |
| +AGREE_00 | <i>acuerdo, pacto, arreglo, eje, amistad, intercomuni3n</i> |
| \$CONSPIRE_00 | <i>intriga, operaci3n, trampa, maniobra, truco, artimaña, jugada, celada, chantaje, enredo; mentira, embuste, infamia, estafa, chanchullo, falsificaci3n, ficci3n, apaño, patraña, equívoco, artificio</i> |
| -CONSPIRE_VIOLENCE | --- |

En la tabla siguiente, se muestra la vinculaci3n de cada clase léxica con que se combina *urdir* con una entidad de FunGramKB, y su clasificaci3n por tipos. Adem3s, se muestran los argumentos miembros de las clases léxicas que pertenecen al paradigma semántico y los que pertenecen a paradigmas discursivos o pragmáticos.

Tabla 302. Tipología de clases léxicas que se combinan con *urdir* con respecto a FunGramKB

| Tipo de clase | CONCEPTO en FunGramKB | CONCEPTO en REDES | Argumentos CL – paradigma semántico | Argumentos CL – paradigma discursivo o pragmático |
|---------------|-----------------------|-------------------------|---|---|
| I | +PLAN_00 | CONFUSIÓN | <i>conspiración, complot, estrategia, estratagema, conjura</i> | --- |
| I | +EXPLANATIO N_00 | RAZONAMIENTO | <i>teoría, hipótesis, interpretación, razonamiento, explicación, comentario</i> | --- |
| II | +VIOLENCE_00 | AGRESIÓN, ACCIÓN HOSTIL | <i>ataque, atentado</i> | <i>venganza, crimen, agresión, golpe de estado, secuestro, amenaza</i> |
| III | +STORY_00 | TEXTO | <i>historia, relato, intriga, narración, argumento, crónica, anécdota</i> | <i>libro, carta</i> |
| IV | --- | CONFUSIÓN | --- | <i>intriga, operación, trampa, maniobra, truco, artimaña, jugada, celada, chantaje, enredo</i> |
| IV | --- | ENGAÑO | --- | <i>mentira, embuste, infamia, estafa, chanchullo, falsificación, ficción, apaño, patraña, equívoco, artificio</i> |
| IV | --- | ACUERDO | --- | <i>acuerdo, pacto, arreglo, eje, amistad, intercomuni3n</i> |

1.4 COCINAR(SE)

REDES distingue un sentido ‘literal’ y uno ‘figurado’ para *cocinar(se)*. A la luz de su combinatoria léxica, equiparamos el sentido literal de *cocinar(se)* a ‘guisar los alimentos’ y desdoblamos el sentido figurado en: ‘acordar’, ‘escribir’, ‘conspirar’ y ‘crear’. El verbo *cocinar(se)*, al igual que otros que hemos visto antes, tiene rasgos de verbo liviano, en tanto que asume los sentidos léxicos de sus complementos. La combinatoria que presenta REDES para *cocinar(se)* se reproduce a continuación:

Tabla 303. Combinatoria léxica de *cocinar(se)*, en REDES

| Sentido del verbo | CONCEPTO asignado a la CL en REDES | Descriptor de la CL en REDES | Argumentos miembros de la CL en REDES |
|-------------------|------------------------------------|--|---|
| ‘guisar’ | (ALIMENTO) | alimentos o comidas | <i>arroz, pescado, carne, guiso, paella, gazpacho</i> |
| ‘acordar’ | ACUERDO | resolución o acción concertada | <i>acuerdo, decisión, consenso, pacto, preacuerdo</i> |
| ‘escribir’ | TEXTO | diversos tipos de textos; también varias formas de creación no verbal | <i>documento, texto, novela, música, disco</i> |
| ‘conspirar’ | CONFUSIÓN | estados de cosas conflictivos, confusos o enredados | <i>entuerto, drama, marrón, conflicto, entresijo</i> |
| ‘crear’ | NORMA | leyes o normas, a menudo de cierta importancia | <i>ley, constitución, decreto</i> |
| | CAMBIO | cambio | <i>reforma, cambio, renovación</i> |
| | PROYECTO | proyecto o propuesta, también otras formas en que se presentan las intenciones de actuar | <i>proyecto, propósito, objetivo, medida, idea, propuesta, sugerencia</i> |

(a) *cocinar(se)* ‘guisar los alimentos’

El primer sentido de *cocinar(se)* es ‘guisar los alimentos’. Vinculamos este sentido al evento conceptual +COOK_00:

Tabla 304. Descripción de +COOK_00 en FunGramKB

| | |
|------------------------|--|
| Superordinado: | +CREATE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2: +FOOD_00)Referent |
| PS: | +(e1: +CREATE_00 (x1)Theme (x2)Referent (f1: +HUMAN_00)Beneficiary (f2)Instrument (f3: +FOOD_00)Means (f4: (e2: +HEAT_00 (x1)Theme (x3: f3)Referent))Manner) |
| Descripción: | prepare for eating by applying heat; "Cook me dinner, please"; "can you make me an omelette?" |
| Lexicalización: | cocinat, estofar, guisar, preparar (español), braise, cook, fix, prepare, stew (inglés) |

El sentido recto de *cocinar* ocurre cuando este se combina con una clase léxica que vinculamos con +FOOD_00. Esta clase léxica pertenece al tipo II, porque está descrita por un concepto superordinado. Todos los argumentos miembros son subordinados de +FOOD_00: *arroz, pescado, carne, guiso, paella, gazpacho*.

Tabla 305. Comparación entre clase léxica de REDES y +FOOD_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|-----------------------|---|---|
| Superordinado: | --- | +LIQUID_00 ^ +SOLID_00 |
| Concepto: | (ALIMENTO) | +FOOD_00 |
| PS: | --- | +(e1: +BE_00 (x1: +FOOD_00)Theme (x2: +LIQUID_00 ^ +SOLID_00)Referent) +(e2: +INGEST_00 (x3: +HUMAN_00 ^ +ANIMAL_00)Agent (x1)Theme (f1: (e3: +BE_01 (x3)Theme (x4: +HUNGRY_00)Attribute))Scene) |
| Descripción: | alimentos o comidas | any substance that can be metabolized by an organism to give energy and build tissue |
| Manifestación léxica: | <i>arroz, pescado, carne, guiso, paella, gazpacho</i> | <i>comida</i> (español), <i>food, foodstuff, nutrient</i> (inglés) |

Como la clase léxica está representada por la preferencia de selección de +COOK_00, los vínculos conceptuales de *cocinar* y su clase léxica está completa.

(b) *cocinar(se)* ‘acordar’

El segundo sentido de *cocinar(se)* es ‘acordar’, que vinculamos con el evento conceptual +AGREE_00:

Tabla 306. Descripción de +AGREE_00 en FunGramKB

| | |
|-----------------|--|
| Superordinado: | +SAY_00 |
| MT: | (x1: +HUMAN_00)Theme (x2)Referent (x3: +HUMAN_00)Goal |
| PS: | +(e1: +SAY_00 (x1)Theme (x2)Referent (x3)Goal) +(e2: +THINK_00 (x1)Theme (x4: (e3: +BE_01 (x2)Theme (x5: +TRUE_00)Attribute))Referent) +(e4: +THINK_00 (x3)Theme (x4)Referent) |
| Descripción: | to have the same opinion, or to accept a suggestion or idea |
| Lexicalización: | <i>estar de acuerdo</i> (español), <i>accede, accord, acquiesce, admit, agree, allow, assent, concur, consent, grant</i> (inglés) |

Cocinar(se) adquiere este sentido en combinación con una clase léxica que no se puede vincular satisfactoriamente con ninguna entidad conceptual. En REDES, la clase léxica está asociada al concepto ACUERDO y el descriptor ‘resolución o acción concertada’; sus miembros son *acuerdo, decisión, consenso, pacto, preacuerdo*. En este caso, en la plantilla léxica de *cocinar(se)*[+AGREE_00], ubicamos las colocaciones *acuerdo, decisión, consenso, pacto, preacuerdo*.

(c) cocinar(se) ‘escribir’

El tercer sentido que adquiere *cocinar(se)* es ‘escribir’, que vinculamos a +WRITE_00:

Tabla 307. Descripción de +WRITE_00 en FunGramKB

| | |
|------------------------|---|
| Superordinado: | +CREATE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2)Referent |
| PS: | +(e1: +CREATE_00 (x1)Theme (x2)Referent (f1: +PENCIL_00 ^ +PEN_00 ^ +MACHINE_00)Instrument (f2: +SURFACE_00)Location (f3: +ON_00)Position) |
| Descripción: | to produce writing or images on paper |
| Lexicalización: | <i>anotar, apuntar, autografiar, caligrafiar, componer, copiar, dactilografiar, dibujar, escribir, estampar</i> (español) <i>annotate, caricature, chalk, compose, copy, crayon, describe, imprint, jot down, note down, paint, pencil, print, record, run off, sign, stamp, take down, trace, type, write, write down</i> (inglés) |

Este sentido se da con una clase léxica que vinculamos a la entidad +WRITING_00. Se trata de una clase léxica mixta, del tipo III. Los argumentos *documento*, *texto* lexicalizan a +WRITING_00 y *novela* sería un subordinado. Por su parte, *música* y *disco* adquieren el sentido de ‘texto’ en combinación con *cocinar(se)*.

Tabla 308. Comparación entre clase léxica de REDES y +WRITING_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|------------------------------|---|---|
| Superordinado: | --- | +INFORMATION_OBJECT_00 |
| Concepto: | TEXTO | +WRITING_00 |
| PS: | --- | +(e1: +BE_00 (x1: +WRITING_00)Theme (x2: +INFORMATION_OBJECT_00)Referent) +(e2: +COMPRISE_00 (x1)Theme (x3: +LETTER_01 +NUMBER_00)Referent) *(e3: pos +READ_00 (x4: +HUMAN_00)Theme (x1)Referent) |
| Descripción: | diversos tipos de textos; también varias formas de creación no verbal | reading matter; anything expressed in letters of the alphabet |
| Manifestación léxica: | <i>documento, texto, novela, música, disco</i> | <i>escrito, texto</i> (español), <i>text, writing</i> (inglés) |

Como con los otros verbos que se vincularon a +WRITE_00, para evitar la creación de un subconcepto redundante como -WRITE_WRITING, la entrada de *cocinar(se)*[+WRITE_00] tendría como colocaciones a todos los miembros de la clase léxica.

(d) *cocinar(se)* ‘conspirar’

El verbo *cocinar(se)* también adquiere el sentido de ‘conspirar’, que se vincula a +CONSPIRE_00 en FunGramKB:

Tabla 309. Descripción de \$CONSPIRE_00 en FunGramKB

| | |
|------------------------|--|
| Superordinado: | +THINK_00 |
| MT: | (x1: +HUMAN_00)Theme (x2:+THOUGHT_00)Referent |
| PS: | +(e1: +THINK_00 (x1)Theme (x2)Referent (f1: +SECRET_00)Manner (f2: +HUMAN_00)Company (f3: ((e2: +DO_00 (x1)Theme (x3)Referent (f2)Company) (e3: n +BE_01 (x3)Theme (x4: +LEGAL_00)Attribute)))Purpose) |
| Descripción: | to secretly plan with someone else_to do something illegal |
| Lexicalización: | <i>conspire, plot</i> (inglés), <i>conchabar, confabular, conspirar, maquinari, tramar</i> (español) |

Este sentido ocurre con una clase léxica que no se puede vincular a la subontología de entidades conceptuales. En REDES, la clase se asocia al concepto CONFUSIÓN, tiene el descriptor ‘estados de cosas conflictivos, confusos o enredados’ y se compone de los miembros: *entuerto, drama, marrón, conflicto, entresijo*.

(e) *cocinar(se)* ‘crear’

El próximo sentido que asume *cocinar(se)* es +CREATE_00:

Tabla 310. Descripción de +CREATE_00 en FunGramKB

| | |
|------------------------|--|
| Superordinado: | #CREATION & +DO_00 |
| MT: | (x1: +HUMAN_00 ^ +ANIMAL_00)Theme (x2)Referent |
| PS: | +(e1: +DO_00 (x1)Theme (x3)Referent (f1: +HUMAN_00 ^ +ANIMAL_00)Beneficiary (f2)Instrument (f3: x2)Result)(e2: rpast n +EXIST_00 (x2)Theme)) |
| Descripción: | bring into existence; "The company was created 25 years ago" |
| Lexicalización: | <i>create, make, produce</i> (inglés), <i>crear, hacer, producir</i> (español) |

El verbo adquiere este sentido en combinación con una clase que vinculamos a +PLAN_00. Es una clase léxica del tipo III, en la que *proyecto, medida, propuesta*

lexicalizan a +PLAN_00, mientras que *propósito, objetivo, idea, sugerencia* adquieren el sentido figuradamente.

Tabla 311. Comparación entre clase léxica de REDES y +PLAN_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|------------------------------|--|--|
| Superodinado: | --- | +INFORMATION_00 |
| Concepto: | PROYECTO | +PLAN_00 |
| PS: | --- | +(e1: +BE_00 (x1: +PLAN_00)Theme (x2: +INFORMATION_00)Referent) *(e2: +DO_00 (x3)Theme (x4)Referent (f1: x1)Manner) |
| Descripción: | proyecto o propuesta, también otras formas en que se presentan las intenciones de actuar | a series of steps to be carried out or goals to be accomplished; "they drew up a six-step plan"; "they discussed plans for a new bond issue" |
| Manifestación léxica: | <i>proyecto, propósito, objetivo, medida, idea, propuesta, sugerencia</i> | <i>estrategia, plan, programa, proyecto</i> (español), <i>plan, programme, project, strategy</i> (inglés) |

A partir de estas selección léxica, formulamos la creación del subconcepto -CREATE_PLAN. En la entrada léxica de *cocinar(se)*[-CREATE_PLAN], se ubicarían como colocaciones los miembros de la clase léxica que no se pueden representar conceptualmente por +PLAN_00: que *propósito, objetivo, idea, sugerencia*.

El verbo *cocinar(se)* también mantiene el sentido de ‘crear’ en combinación con otra clase léxica, que vinculamos a +LAW_00. Esta clase es una clase estable (tipo I):

Tabla 312. Comparación entre clase léxica de REDES y +LAW_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|----------------------|--|---|
| Superodinado: | --- | +GROUP_00 |
| Concepto: | NORMA | +LAW_00 |
| PS: | --- | +(e1: +BE_00 (x1: +LAW_00)Theme (x2: +GROUP_00)Referent)(e2: +COMPRISE_00 (x2)Theme (x3: +RULE_00)Referent)) *(e3: +CREATE_00 (x4: +GOVERNMENT_00)Theme (x1)Referent) *(e4: obl +OBEY_00 (x5: +HUMAN_00)Theme (x1)Referent) |
| Descripción: | leyes o normas, a menudo de cierta importancia | the collection of rules imposed by authority; "civilization presupposes respect for the law" |
| Manifestación | <i>ley, constitución,</i> | <i>derecho, legislación, ley</i> (español), <i>law, legislation</i> |

| | | |
|----------------|----------------|----------|
| léxica: | <i>decreto</i> | (inglés) |
|----------------|----------------|----------|

A partir de esta selección léxica, proponemos el subconcepto -CREATE_LAW, al que se vincularía *cocinar(se)*. Como los miembros de la clase léxica están bien representados por la preferencia de selección, no hay que añadirlos como colocaciones en la plantilla léxica de *cocinar(se)[-CREATE_LAW]*.

La tercera clase léxica con que se combina *cocinar(se)* bajo el sentido de +CREATE_00 no se puede vincular a ninguna entidad (tipo IV). Tiene el concepto CAMBIO, el descriptor ‘cambio’ y los argumentos: *reforma, cambio, renovación*. Estos argumentos aparecerían como colocaciones en la entrada léxica de *cocinar(se)[+CREATE_00]*.

1.4.1 Resumen COCINAR(SE)

La tabla siguiente muestra los vínculos de *cocinar(se)* en FunGramKB con eventos conceptuales en la ontología. También se identifican las colocaciones que se ubican en las plantillas léxicas de *cocinar(se)* en su vínculo con diferentes conceptos. Las colocaciones, como sabemos, son los argumentos que no se pueden representar por las preferencias de selección conceptuales:

Tabla 313. Resumen: COCINAR(SE) en FunGramKB (vinculación propuesta)

| Evento conceptual | Colocaciones |
|--------------------------|---|
| +COOK_00 | --- |
| +AGREE_00 | <i>acuerdo, decisión, consenso, pacto, preacuerdo</i> |
| +WRITE_00 | <i>documento, texto, novela, música, disco</i> |
| +CONSPIRE_00 | <i>entuerto, drama, marrón, conflicto, entresijo</i> |
| +CREATE_00 | <i>reforma, cambio, renovación</i> |
| -CREATE_PLAN | <i>propósito, objetivo, idea, sugerencia</i> |
| -CREATE_LAW | --- |

En la tabla siguiente, se muestra la vinculación de cada clase léxica con que se combina *cocinar(se)* con una entidad de FunGramKB, y su clasificación por tipos. Además, se muestran los argumentos miembros de las clases léxicas que pertenecen al paradigma semántico y los que pertenecen a paradigmas discursivos o pragmáticos:

Tabla 314. Tipología de clases léxicas que se combinan con *cocinar(se)* con respecto a FunGramKB

| Tipo de clase | CONCEPTO en FunGramKB | CONCEPTO en REDES | Argumentos CL – paradigma semántico | Argumentos CL – paradigma discursivo pragmático |
|---------------|-----------------------|-------------------|---|---|
| I | +LAW_00 | NORMA | <i>ley, constitución, decreto</i> | --- |
| II | +FOOD_00 | (ALIMENTO) | <i>arroz, pescado, carne, guiso, paella, gazpacho</i> | --- |
| III | +PLAN_00 | PROYECTO | <i>proyecto, medida, propuesta</i> | <i>propósito, objetivo, idea, sugerencia</i> |
| III | +WRITING_00 | TEXTO | <i>documento, texto, novela,</i> | <i>música, disco</i> |
| IV | --- | ACUERDO | --- | <i>acuerdo, decisión, consenso, pacto, preacuerdo</i> |
| IV | --- | CONFUSIÓN | --- | <i>entuerto, drama, marrón, conflicto, entresijo</i> |
| IV | --- | CAMBIO | --- | <i>reforma, cambio, renovación</i> |

1.5 MAQUINAR

REDES no distingue sentidos para *maquinar*. A partir de su combinatoria, le adjudicamos el sentido de ‘conspirar’. Para efectos del estudio de polisemia regular, añadimos el sentido etimológico de ‘trabajar una pieza metalúrgica por medio de una máquina’, que no está presente en REDES pero sí en el DRAE. La combinatoria de *maquinar*, según la presenta REDES, es:

Tabla 315. Combinatoria léxica de *maquinar*, en REDES

| Sentido del verbo | CONCEPTO asignado a la CL en REDES | Descriptor de la CL en REDES | Argumentos miembros de la CL en REDES |
|-------------------|------------------------------------|---|--|
| ‘conspirar’ | (PROYECTO) | planes y otras formas de organizar las actuaciones previstas | <i>plan, estrategia, tesis</i> |
| | OPERACIÓN | operación, muy frecuentemente de carácter delictivo, ilícito o encubierto | <i>operación, crimen, intriga, conspiración, trama, enredo, estafa, fuga</i> |

(a) *maquinar* ‘trabajar una pieza metálica por medio de una máquina’

El sentido de ‘trabajar una pieza metálica por medio de una máquina’ lo vinculamos con +OPERATE_00. El MT establece las preferencias de selección de +OPERATE_00 como un Tema ‘humano’ y un Referente ‘artefacto’:

Tabla 316. Descripción de +OPERATE_00 en FunGramKB

| | |
|------------------------|--|
| Superordinado: | +USE_00 |
| MT: | (x1: +HUMAN_00)Theme (x2: +ARTEFACT_00)Referent |
| PS: | +(e1: +USE_00 (x1)Theme (x2)Referent (f1)Instrument (f2: (e2: +DO_00 (x2)Theme (x3)Referent)))Purpose) |
| Descripción: | to use and control a machine or equipment |
| Lexicalización: | <i>arrancar, controlar, encender, funcionar, operar, poner en marcha</i> (español), <i>control, function, go, handle, operate, run, start, start up, switch on, turn on, work</i> (inglés) |

(b) *maquinar* ‘conspirar’

El sentido principal de *maquinar* es ‘conspirar’, que vinculamos en FunGramKB a \$CONSPIRE_00:

Tabla 317. Descripción de \$CONSPIRE_00 en FunGramKB

| | |
|------------------------|--|
| Superordinado: | +THINK_00 |
| MT: | (x1: +HUMAN_00)Theme (x2:+THOUGHT_00)Referent |
| PS: | +(e1: +THINK_00 (x1)Theme (x2)Referent (f1: +SECRET_00)Manner (f2: +HUMAN_00)Company (f3: ((e2: +DO_00 (x1)Theme (x3)Referent (f2)Company) (e3: n +BE_01 (x3)Theme (x4: +LEGAL_00)Attribute)))Purpose) |
| Descripción: | to secretly plan with someone else_to do something illegal |
| Lexicalización: | <i>conspire, plot</i> (inglés), <i>conchabar, confabular, conspirar, maquinar, tramar</i> (español) |

El verbo *maquinar* se combina con dos clases léxicas. Vinculamos la primera a la entidad conceptual +PLAN_00. Se trata de una clase mixta (tipo III): *plan, estrategia* lexicalizan a +PLAN_00, mientras que *tesis* asume ese sentido figuradamente.

Tabla 318. Comparación entre clase léxica de REDES y +PLAN_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|-----------------------|--|--|
| Superodinado: | --- | +INFORMATION_00 |
| Concepto: | (PROYECTO) | +PLAN_00 |
| PS: | --- | +(e1: +BE_00 (x1: +PLAN_00)Theme (x2: +INFORMATION_00)Referent) *(e2: +DO_00 (x3)Theme (x4)Referent (f1: x1)Manner) |
| Descripción: | planes y otras formas de organizar las actuaciones previstas | a series of steps to be carried out or goals to be accomplished; "they drew up a six-step plan"; "they discussed plans for a new bond issue" |
| Manifestación léxica: | <i>plan, estrategia, tesis</i> | <i>estrategia, plan, programa, proyecto</i> (español), <i>plan, programme, project, strategy</i> (inglés) |

Como hemos dicho antes, no es necesario especificar la preferencia de selección del Referente como +PLAN_00, porque el Referente actual, +THOUGHT_00, es superordinado de +PLAN_00.

La próxima clase léxica se vincula con +CRIME_00. Es una clase mixta (tipo III) también: *crimen* lexicaliza a +CRIME_00, mientras que *conspiración, estafa, fuga* son subordinados. Los demás argumentos: *operación, intriga, trama, enredo* adquieren el sentido figuradamente.

Tabla 319. Comparación entre clase léxica de REDES y +CRIME_00 en FunGramKB

| Tipo de información | REDES | FUNGRAMKB |
|-----------------------|--|---|
| Superodinado: | --- | +OCCURRENCE_00 |
| Concepto: | OPERACIÓN | +CRIME_00 |
| PS: | --- | +(e1: +BE_00 (x1: +CRIME_00)Theme (x2: +OCCURRENCE_00)Referent) +(e2: n +BE_01 (x1)Theme (x3: +LEGAL_00)Attribute) |
| Descripción: | operación, muy frecuentemente de carácter delictivo, ilícito o encubierto | an act punishable by law; usually considered an evil act |
| Manifestación léxica: | <i>operación, crimen, intriga, conspiración, trama, enredo, estafa, fuga</i> | <i>delito, ilegalidad, infracción, quebrantamiento, quebranto, transgresión, vulneración</i> (español), <i>crime, law-breaking</i> (inglés) |

Crear el subconcepto -CONSPIRE_CRIME sería redundante, así que añadimos los argumentos miembros de la clase léxica a la entrada léxica de *maquinar*[+CONSPIRE_00].

1.5.1 Resumen MAQUINAR

La tabla siguiente muestra los vínculos de *maquinar* en FunGramKB con eventos conceptuales en la ontología. También se identifican las colocaciones que se ubican en las plantillas léxicas de *maquinar* en su vínculo con diferentes conceptos:

Tabla 320. Resumen: MAQUINAR en FunGramKB (vinculación propuesta)

| Evento conceptual | Colocaciones |
|-------------------|---|
| +OPERATE_00 | --- |
| +CONSPIRE_00 | <i>operación, intriga, trama, enredo; operación, crimen, intriga, conspiración, trama, enredo, estafa, fuga</i> |

En la tabla siguiente, se muestra la vinculación de cada clase léxica con que se combina *maquinar* con una entidad de FunGramKB, y su clasificación por tipos. Presentamos también los argumentos miembros de las clases léxicas que pertenecen al paradigma semántico y los que pertenecen a paradigmas discursivos o pragmáticos:

Tabla 321. Tipología de clases léxicas que se combinan con *maquinar* con respecto a FunGramKB

| Tipo de clase | CONCEPTO en FunGramKB | CONCEPTO en REDES | Argumentos CL – paradigma semántico | Argumentos CL – paradigma discursivo o pragmático |
|---------------|-----------------------|-------------------|--|---|
| III | +PLAN_00 | (PROYECTO) | <i>plan, estrategia</i> | <i>tesis</i> |
| III | +CRIME_00 | OPERACIÓN | <i>crimen, conspiración, trama, estafa, fuga</i> | <i>operación, intriga, trama, enredo</i> |

2. Análisis

2.1 Sentidos verbales

La próxima tabla muestra los vínculos de los sentidos de los cinco verbos con eventos conceptuales en FunGramKB.

Tabla 322. Vínculos de los sentidos predicativos con eventos conceptuales en FunGramKB

| <i>tejer</i> | <i>tramar</i> | <i>urdir</i> | <i>cocinar(se)</i> | <i>maquinar</i> |
|---------------------|------------------------|------------------------|--------------------|-----------------|
| +WEAVE | +WEAVE | +WEAVE | +COOK | +OPERATE |
| \$CONSPIRE | \$CONSPIRE | \$CONSPIRE | \$CONSPIRE | \$CONSPIRE |
| | -CONSPIRE _VIOLENCE | -CONSPIRE _VIOLENCE | | |
| +WRITE | | | +WRITE | |
| -WRITE _STORY | | -WRITE _STORY | | |
| +CREATE | | | +CREATE | |
| -CREATE _PROBLEM | | | -CREATE _PLAN | |
| | | | -CREATE _LAW | |
| +EXPLAIN | | +EXPLAIN | | |
| | | | | |
| | | +AGREE | +AGREE | |
| | +THINK_00 | | | |
| +DESIRE_00 | | | | |
| +FEEL_00 | | | | |

Como refleja la Tabla, hay un vínculo compartido entre los cinco predicados con el evento \$CONSPIRE_00. De los cinco verbos, hay tres *–tejer, tramar, urdir–* que tienen también vínculo con +WEAVE_00. Existe, por lo tanto, un caso de polisemia regular directa entre *tejer, tramar, urdir*, y los sentidos +WEAVE_00 y \$CONSPIRE_00. Ese patrón se puede validar para las palabras inglesas *weave* y *conspire*. Ahora bien, si consideramos que +CREATE_00 es el superordinado tanto de +WEAVE_00 como de +COOK_00, podemos añadir a *cocinar(se)* en un patrón de polisemia regular indirecta: *tejer, tramar, urdir* y *cocinar(se)* alternan entre +CREATE_00 y \$CONSPIRE_00. Existe otro patrón de polisemia regular directa entre los sentidos +WEAVE_00, \$CONSPIRE_00 y +WRITE_00, que es manifestado por *tejer, urdir, cocinar(se)*. La próxima tabla muestra los tres patrones:

Tabla 323. Polisemia regular (directa e indirecta) de cuatro verbos

| PREDICADO | CONCEPTO | | | |
|--------------------|------------|--------|-------------------------|---------------------|
| | \$CONSPIRE | +WEAVE | +CREATE | +WRITE |
| <i>tejer</i> | + | + | + | + |
| <i>tramar</i> | + | + | + | |
| <i>urdir</i> | + | + | +(+WEAVE) ⁹⁴ | +(-WRITE _STORY) |
| <i>cocinar(se)</i> | + | | +(+COOK) | + |

Se podría postular que la alternancia semántica de este grupo es de naturaleza metafórica: va de un primer sentido literal general de ‘confeccionar algo’ a un segundo sentido de ‘conspirar’. En términos de FunGramKB, sería +CREATE_00 → \$CONSPIRE_00. Más específicamente, en los verbos *urdir*, *tramar*, la extensión va de ‘tejer un textil o preparar los hilos para ello’ o ‘armar algo con cuidado y paciencia’ a ‘elaborar una artimaña’ o ‘conspirar’. En términos de FunGramKB, sería +WEAVE_00 (subordinado de +CREATE_00) → \$CONSPIRE_00. En el grupo de *tejer*, *urdir*, *cocinar(se)*, hay una alternancia entre ‘escribir’ y ‘conspirar’, que en términos conceptuales se ve como +WRITE_00 y \$CONSPIRE_00. Aquí también la relación que media es la idea de ‘armar algo con cuidado y paciencia’, que se puede manifestar en la escritura o en una conspiración. Con *cocinar* y *maquinar* pasa algo similar: hay un sentido recto que tiene que ver con una confección compleja que toma tiempo, que pasa a ser una conspiración. Aquí el movimiento semántico sería +COOK_00 /+OPERATE_00 --> \$CONSPIRE_00.

2.2 Preferencias de selección

En esta sección, miramos las preferencias de selección de los verbos, especialmente de los que participan en los patrones de polisemia regular. En la tabla siguiente, presentamos los subconceptos que resultaron necesarios para reflejar algunas preferencias de selección de los verbos. No hay, sin embargo, casos de polisemia regular en el nivel del subconcepto:

⁹⁴ Marcamos entre paréntesis +WEAVE_00 y +COOK_00, los conceptos subordinados a +CREATE_00, con lo que se vinculan los verbos *urdir* y *cocinar(se)*, y el subconcepto -WRITE_STORY, con que se vincula *urdir*.

Tabla 324. Subconceptos que reflejan preferencias de selección

| PREDICADO | SUBCONCEPTO (propuesto) | | | | |
|--------------------|-------------------------|------------------|---------------------|------------------|-----------------|
| | -CONSPIRE _VIOLENCE | -WRITE _STORY | _CREATE _PROBLEM | -CREATE _PLAN | -CREATE _LAW |
| <i>tejer</i> | | + | + | | |
| <i>tramar</i> | + | | | | |
| <i>urdir</i> | + | + | | | |
| <i>cocinar(se)</i> | | | | + | + |

Los datos de selección léxica de estos verbos validaron preferencias de selección ya presentes en los conceptos de FunGramKB. Por ejemplo, los verbos *tejer*, *tramar*, *urdir*, *maquinar* se vinculan a +CONSPIRE_00 y tienen como preferencia de selección a +PLAN_00, que está contenida ya en la preferencia de selección +THOUGHT_00. También hubo algunos casos de preferencias de selección que resultaban redundantes, porque se repetía el mismo sentido básico en el evento y en la entidad. Nos parece que esta evidencia apunta al hecho de que estos verbos tienen rasgos de verbos livianos: al combinarse con ciertos complementos, asumen el sentido léxico verbal del complemento nominal. Este fue el caso de: +WRITE_00 y +WRITING_00 (*tejer*, *cocinarse*), +EXPLAIN_00 y +EXPLANATION_00 (*tejer*, *urdir*), +DESIRE_01 y +DESIRE_00 (*tejer*), +FEEL_00 y +FEELING_00 (*tejer*), +THINK_00 y +THOUGHT_00 (*tramar*). También hubo dos verbos –*urdir*, *cocinarse*– que se vincularon con +AGREE_00 en combinación con clases del tipo IV que REDES identifica como ACUERDO.

2.3 Clases léxicas y entidades conceptuales

Los cinco verbos se combinan con un total de 32 clases léxicas diferentes. Al distribuir las entre los cuatro tipos, obtuvimos los siguientes resultados:

Tabla 325. Distribución de clases léxicas por tipo

| Tipo de clase | Cantidad | Porcentaje |
|---------------|----------|------------|
| Tipo I | 5 | 16% |
| Tipo II | 5 | 16% |
| Tipo III | 13 | 41% |
| Tipo IV | 9 | 28% |

La distribución de estas clases léxicas por tipos coincide parcialmente con la distribución de las clases léxicas de los nueve verbos que vimos antes. En ambos grupos, el tipo más productivo fueron las clases mixtas (tipo III) y el segundo tipo más productivo fueron las clases que no se podían vincular (tipo IV). Los tipos I y II, que sentarían las bases para una reducción de inventario de clases léxicas, sumaron un 32%. Las clases léxicas del tipo I se vincularon con: +PLAN_00, +EXPLANATION_00, +THOUGHT_00, +STORY_00, +LAW_00, mientras que las del tipo II se vincularon con: +VIOLENCE_00, +FOOD_00, +CRIME_00, +CLOTH_00, +FEELING_00.

Las 32 clases léxicas con que se combinaron estos cinco verbos se pudieron vincular a 13 entidades conceptuales, lo que constituye una reducción en el inventario. Cuando miramos las entidades conceptuales que eligen los predicados, hay varias selecciones compartidas: +PLAN_00, +CRIME_00, +EXPLANATION_00, +VIOLENCE, +WRITING_00 y +STORY_00.

Tabla 326. Preferencias de selección de predicados

| <i>tejer</i> | <i>tramar</i> | <i>urdir</i> | <i>cocinar(se)</i> | <i>maquinar</i> |
|--------------|---------------|--------------|--------------------|-----------------|
| +PLAN | +PLAN | +PLAN | +PLAN | +PLAN |
| | +CRIME | | | +CRIME |
| +EXPLANATION | | +EXPLANATION | | |
| | +VIOLENCE | +VIOLENCE | | |
| +WRITING | | | +WRITING | |
| | +STORY | +STORY | | |
| +FEELING | | | | |
| +DESIRE | | | | |
| +PROBLEM | | | | |
| +CLOTH | | | | |
| | +THOUGHT | | | |
| | | | +LAW | |
| | | | +FOOD | |

Cabe señalar también que hubo algunas clases léxicas compartidas por los cinco verbos que no se pudieron vincular a entidades conceptuales. Por ejemplo, *tejer*, *urdir* y *cocinar(se)* eligieron clases léxicas que REDES asocia a CONFUSIÓN, mientras que *urdir* y *maquinar* escogieron clases descritas como ACUERDO.

CONCLUSIONES

Las conclusiones de esta tesis se organizan contestando las preguntas de trabajo que se presentaron en la Introducción. En la primera parte, “Cómo puede servir FunGramKB a REDES”, presentamos los beneficios que aporta el marco ontológico de FunGramKB a los datos de restricciones de selección léxica de REDES. En la segunda parte, “Cómo puede servir REDES a FunGramKB”, consideramos la contribución de los datos lingüísticos de REDES a la base de conocimiento léxico conceptual FunGramKB. En el apartado tercero y final, contemplamos futuras líneas de investigación.

1. Cómo puede servir FunGramKB a REDES

1.1 ¿Cómo sirve el entorno ontológico de FunGramKB para sacar el máximo provecho de los datos de REDES, es decir, para reducir, organizar y relacionar conceptualmente la casuística lingüística de REDES?

El entorno ontológico de FunGramKB permite hacer una traducción conceptual de los datos de REDES, tanto de la semántica de los predicados como de la semántica de las clases léxicas, y de la relación de selección que media entre ambas. Los sentidos de los predicados se vinculan con eventos conceptuales, mientras que las clases léxicas se vinculan con entidades conceptuales. Las restricciones semánticas en la selección léxica de los predicados pasa a ser parte integral de la definición de los eventos en FunGramKB, tomando la forma de preferencias de selección conceptuales. Las selecciones léxicas que no se pueden codificar conceptualmente –porque adquieren de manera figurada el sentido que define la clase léxica–, se ubican como colocaciones en la plantilla léxica del predicado. De estas maneras, la base de conocimiento léxico conceptual FunGramKB permite representar la información de REDES –reorganizando hasta cierto punto el material–, sin que se pierda ningún dato.

La codificación de la semántica de los predicados resulta más sencilla que la codificación de la semántica de las clases léxicas. Esto se debe a que la semántica de una unidad léxica (en este caso, cada verbo) es más discreta –por compleja que sea– que la semántica de una clase compuesta por múltiples unidades léxicas polisémicas agrupadas por múltiples criterios. Estos criterios no son solo semánticos, sino también pragmáticos e incluyen extensiones semánticas variables.

En el marco de FunGramKB, logramos hacer una representación reducida de los sentidos verbales: la mayoría de los 15 verbos con que trabajamos presentaron vinculaciones

con tres eventos conceptuales. El inventario de clases léxicas también fue reducido significativamente: la selección léxica de nueve verbos en su vínculo con +WASTE_00 se redujo de 46 clases léxicas a 12 entidades conceptuales, y la selección léxica de los verbos que presentaban polisemia regular se redujo de 32 clases léxicas a 13 entidades conceptuales. Aunque hubo reducción en la representación de los datos en la ontología conceptual, hubo ampliación en otros niveles. Por ejemplo, para dar cuenta de sus preferencias de selección, los verbos generaron 3.5 subconceptos nuevos cada uno en promedio. Por otro lado, aunque las clases léxicas se vincularon a un inventario reducido de entidades conceptuales, los argumentos que no se pudieron representar conceptualmente fueron incluidos como colocaciones en la plantilla léxica del predicado. Esto quiere decir que la información de REDES se redistribuye y reacomoda en diferentes espacios de la base de conocimiento, pero mantiene una representación reducida en el eje central conceptual.

Organizar la información a partir de un eje conceptual compartido nos permite visualizar claramente la semántica compartida y diferente de las unidades predicativas y de sus selecciones léxicas. Esto representa una posibilidad valiosa para la lexicografía electrónica moderna. Por ejemplo, en torno a \$WASTE_00, se organizan nueve verbos: algunos monosémicos (*malgastar*, *desaprovechar*, *dilapidar*, *despilfarrar*, *desperdiciar*) y otros polisémicos, que se vinculan a su vez con otros conceptos: +FAIL_00 (*perder* y *malograr*), \$ABOUND_00 (*derrochar*), +SPLIT_00 (*disipar*), +FINISH_00 (*disipar*), \$MISS_00 (*perder*) y +LOSE_00 (*perder*). El enfoque conceptualista nos permite considerar no solo las relaciones entre predicados, sino también las relaciones entre conceptos. Podemos formular preguntas como, por ejemplo, ¿cuál es la relación entre +WASTE_00 y +FAIL_00, o entre +WASTE_00 y \$ABOUND_00? Asimismo, nos permite explorar posibles patrones de polisemia regular: partiendo de dos conceptos, como +WEAVE_00 y \$CONSPIRE_00, o +WRITE_00 y \$CONSPIRE_00, podemos conocer la nómina de verbos que alternan entre estos dos sentidos.

Los vínculos de las clases léxicas con entidades conceptuales nos permiten explorar la información de selección de maneras nuevas. Por ejemplo, una preferencia de selección +WASTE_00 (presente en seis verbos: *malgastar*, *desaprovechar*, *dilapidar*, *malograr*, *despilfarrar*, *desperdiciar*) es +ABILITY_00. Si orientamos la búsqueda desde +ABILITY_00, descubrimos que también es seleccionada por \$ABOUND_00 (visto en un verbo: *derrochar*). La orientación conceptualista nos permite explorar qué otros predicados se predicán de cada entidad.

El módulo léxico de FunGramKB completa la información lingüística de REDES que no se puede representar conceptualmente. El hecho es que las entidades conceptuales reflejan satisfactoriamente la semántica de las clases léxicas, pero no necesariamente pueden proyectar todas las palabras que las constituyen. Para eso está la plantilla léxica del predicado en el módulo léxico. Por ejemplo, en su sentido de +WASTE_00, *desaprovechar* se combina con argumentos que lexicalizan a +ABILITY_00, como *talento, capacidad, facultad, facilidad, potencial, cualidad*. Pero *desaprovechar*, en ese mismo sentido, también se combina con *conocimiento, disposición, carrera*, argumentos que adquieren el sentido de ‘habilidad’ de manera figurada y no pueden lexicalizar a +ABILITY_00. Por lo tanto, en la plantilla léxica de *desaprovechar*[-WASTE_ABILITY] ubicaremos estos argumentos (*conocimiento, disposición, carrera*) como colocaciones.

El marco ontológico de FunGramKB dota la información de restricciones de selección léxica de REDES de una arquitectura conceptual común –que permite reducir, organizar y relacionar la información, de maneras que no son posibles en el formato actual del diccionario–, a la vez que provee un módulo léxico que guarda aspectos particulares de cada unidad léxica. Aunque requiere una reorganización de los datos, FunGramKB es capaz de albergar toda la información de REDES y hacerla disponible para la explotación lexicográfica.

1.2 ¿Pueden vincularse los predicados de REDES –o, mejor, los sentidos que asumen los predicados– a la ontología conceptual de FunGramKB?

Se logró representar conceptualmente la monosemia o polisemia de los 15 predicados de manera satisfactoria, con una cantidad de conceptos reducida. Como dijimos antes, cada verbo tuvo vínculos con un promedio de tres eventos conceptuales. Ello satisface el desiderátum de Nirenburg y Raskin (2004) de mantener el número de sentidos en el mínimo posible para efectos del PLN. FunGramKB nos permite, por lo tanto, alejarnos del lexicón enumerativo, sin llevarnos a un reduccionismo extremo:

Tabla 327. Cantidad de eventos conceptuales vinculados con cada verbo

| Verbo | Cantidad de eventos conceptuales | Verbo | Cantidad de eventos conceptuales |
|----------------------|----------------------------------|------------------|----------------------------------|
| <i>dictar</i> | 3 | <i>tejer</i> | 7 |
| <i>perder</i> | 5 | <i>tramar</i> | 3 |
| <i>disipar</i> | 3 | <i>urdir</i> | 5 |
| <i>derrochar</i> | 2 | <i>cocinarse</i> | 5 |
| <i>malograr</i> | 2 | <i>maquinar</i> | 2 |
| <i>malgastar</i> | 2 | | |
| <i>dilapidar</i> | 1 | | |
| <i>desaprovechar</i> | 1 | | |
| <i>despilfarrar</i> | 1 | | |
| <i>desperdiciar</i> | 1 | | |

En muchos casos, los conceptos básicos o terminales existentes daban cuenta del sentido del verbo y de sus restricciones de selección. En otros casos, fue necesario formular nuevos conceptos terminales o subconceptos para dar cabida a los componentes particulares de sentido o a la información de selección léxica que aporta REDES. Trabajar con grupos de verbos relacionados semánticamente fue importante para efectos de la creación de estos nuevos conceptos terminales o subconceptos, porque provee un marco de referencia que evidencia los rasgos de sentido comunes y destaca los diferentes.

De nuestro trabajo, emergieron cuatro nuevos conceptos terminales: \$DICTATE_00, \$MISS_00, \$DISSIPATE_00, \$SQUANDER_00. El primero, \$DICTATE_00, es subordinado de +SAY_00. Este concepto terminal es necesario para codificar uno de los sentidos del predicado *dictar*, un verbo que se analizó por su cuenta. Para adquirir ese sentido, *dictar* se combina con una clase léxica que vinculamos a la entidad +WRITING_00 (descrita en REDES como ‘textos’ y compuesta de *carta*, *texto*, *informe*, *párrafo*). Para codificar \$DICTATE, establecemos que el Referente del concepto básico +SAY_00 tiene una preferencia de selección por +WRITING_00, y añadimos un parámetro que explica que un participante humano le dice a otro participante humano algo que está escrito para que el segundo lo escriba. El segundo concepto terminal que proponemos sale de la combinación del verbo *perder* con +VEHICLE_00 o +TRAVEL_00 (*perder el tren*, *perder el vuelo*). Ese nuevo concepto terminal se llama \$MISS_00, y es subordinado de +FAIL_00. Además de establecer la preferencia de selección del Referente de +FAIL_00 como +VEHICLE_00 o +TRAVEL_00, este concepto terminal articula dos restricciones de sentido en el PS de

+FAIL_00 que informan de que el participante humano no logra llegar a ese transporte a tiempo. El tercer concepto terminal es \$DISSIPATE_00, y es subordinado de +SPLIT_00. Lo usamos para codificar uno de los sentidos del verbo *disipar*, que ocurre cuando este se combina con una clase léxica compuesta por sustantivos que designan masas de gas o energía (*viento, nube, aire, humo, gas, niebla, luz, sonido, fuerza, imagen, música*). En ese caso, además de especificar la preferencia de selección del Referente de +SPLIT_00 a +GAS_00^+ENERGY_00, se añadió un parámetro que establecía que las partes en que se dividía este Referente eran extremadamente pequeñas. El cuarto concepto terminal es \$SQUANDER_00, subordinado de +WASTE_00. Este concepto terminal es motivado por los verbos *despilfarrar* y *derrochar* en combinación con +MONEY_00 (*despilfarrar, derrochar dinero*). En este caso se añade la preferencia de selección por +MONEY_00 al MT de +WASTE_00 y un parámetro de exceso al PS de +WASTE_00.

El recurso del subconcepto –que restringe los participantes del MT de un concepto existente, pero mantiene el mismo PS– fue particularmente útil para dar cuenta de las preferencias de selección de los predicados. Los subconceptos nos permiten ampliar la extensión de un concepto, manteniendo igual su intensidad. Por otro lado, como los subconceptos no constituyen nuevos eventos, no sobrecargan la ontología. La próxima tabla da cuenta de la cantidad de subconceptos que se formularon para cada verbo:

Tabla 328. Cantidad de subconceptos para cada verbo

| Verbo | Cantidad de subconceptos | Verbo | Cantidad de subconceptos |
|----------------------|--------------------------|------------------|--------------------------|
| <i>dictar</i> | 3 | <i>tejer</i> | 2 |
| <i>perder</i> | 4 | <i>tramar</i> | 1 |
| <i>disipar</i> | 5 | <i>urdir</i> | 2 |
| <i>derrochar</i> | 3 | <i>cocinarse</i> | 2 |
| <i>malograr</i> | 5 | <i>maquinar</i> | 0 |
| <i>malgastar</i> | 6 | | |
| <i>dilapidar</i> | 6 | | |
| <i>desaprovechar</i> | 4 | | |
| <i>despilfarrar</i> | 4 | | |
| <i>desperdiciar</i> | 6 | | |

Los subconceptos añaden información más precisa a cada evento en la ontología, dando cuenta de sus diferentes preferencias de selección y de los predicados que lexicalizan

esas preferencias de selección. Por ejemplo, con +WASTE_00 se vinculan nueve predicados diferentes, pero el subconcepto -WASTE_MONEY se lexicaliza con más naturalidad como *dilapidar* o *disipar dinero*, mientras que -WASTE_FOOD se expresa más naturalmente como *despilfarrar* o *desperdiciar comida*.

La selección léxica de un solo predicado es suficiente para validar la creación de un subconcepto, pero en ocasiones varios predicados comparten la misma selección. Se formularon 12 subconceptos de +WASTE_00, que reflejan las diferentes preferencias de selección de los verbos. Los siguientes tres subconceptos están lexicalizados por cinco verbos o más: -WASTE_ABILITY, -WASTE_TIME, -WASTE_MONEY. Esto sugiere que existe una “concordancia semántica” fuerte entre el evento +WASTE_00 y esas tres entidades.

Es pertinente hacer un comentario sobre la representación de los sentidos rectos y los sentidos figurados de los predicados. Los sentidos rectos y figurados de los verbos usualmente están ubicados en ramas diferentes de la ontología de FunGramKB. Aunque no hay conexión explícita en el nivel de los eventos, se puede establecer la relación a través del predicado. Por ejemplo, en su sentido recto, *disipar* se vincula con un concepto terminal de +SPLIT_00 (\$DISSIPATE_00) cuando se combina con +GAS_00 y +ENERGY_00 (*disipar la niebla*), pero con un subconcepto de +WASTE_00 (-WASTE_MONEY) cuando se combina con +MONEY_00 (*disipar una fortuna*) y con varios subconceptos de +FINISH_00 cuando se combina con otras nueve clases léxicas (*disipar las dudas*, *disipar la tensión*, etc.). Sería imposible relacionar explícitamente estos tres eventos en la ontología, porque tienen rutas conceptuales diferentes. Sería igualmente imposible ubicar estos tres sentidos como subconceptos de +SPLIT_00, usando diferentes extensiones, porque +SPLIT_00 tiene una ruta conceptual material, que no da cabida al sentido de *disipar dudas*. El vínculo común entre estos tres conceptos radica pues en el lexicon, en el predicado mismo: *disipar*.

El marco ontológico de FunGramKB posibilitó la representación conceptual de la monosemia o polisemia de los 15 predicados, con una granularidad semántica manejable por una máquina. El recurso del subconcepto de FunGramKB fue clave para representar la extensión de los verbos sin variar su intensidad, lo que coincide con planteamientos de Bosque (2004). A través de los subconceptos, pudimos codificar las restricciones de selección de los verbos en REDES como preferencias de selección conceptuales en FunGramKB, sin necesariamente sobrecargar con un nuevo concepto a la ontología. El subconcepto, en efecto,

nos permitió unir la intensión y la extensión de un predicado en una misma unidad conceptual.

1.3 ¿Puede reducirse el inventario de clases léxicas de REDES usando la ontología conceptual de FunGramKB? ¿Se mantiene la clase léxica como unidad ante la ontología de FunGramKB?

El inventario de clases léxicas de REDES se puede reducir por criterios semánticos usando la ontología de FunGramKB, pero para dar cuenta de todos los miembros léxicos de la clase a veces es necesario contar con el módulo léxico de FunGramKB. Esto quiere decir que, en muchos casos, la clase léxica de REDES no se mantiene como unidad de trabajo íntegra en FunGramKB; se divide en un paradigma semántico y un paradigma pragmático o discursivo, los cuales mantienen una conexión por medio de los módulos de FunGramKB.

Las clases léxicas de REDES se constituyen porque se combinan con un mismo predicado; están compuestas por múltiples unidades léxicas polisémicas agrupadas no solo por criterios semánticos sino también pragmáticos y discursivos. FunGramKB está fundamentada en una ontología de conceptos con significados unívocos y lexicalizaciones inequívocas. Esto quiere decir que, si bien los componentes semánticos y conceptuales de las clases léxicas apuntan a conceptos de FunGramKB, los conceptos de FunGramKB no siempre pueden apuntar a todos los miembros de la clase léxica de REDES.

Nuestra clasificación tipológica de las clases léxicas presenta un panorama de la relación entre la semántica de la clase y sus miembros léxicos. Las clases léxicas seleccionadas por los 15 verbos estudiados se distribuyeron como sigue: las clases estables, en que los miembros lexicalizan el concepto (tipo I) sumaron un 23%; las clases cuyos miembros léxicos son descritos por un concepto superordinado (tipo II) alcanzaron un 19%; las clases mixtas, en donde algunos miembros pertenecen al paradigma semántico y otros al paradigma discursivo o pragmático (tipo III) obtuvieron el 31%; y las clases que no se podían describir por ninguna entidad conceptual (tipo IV) constituyeron el 27%. Esto quiere decir que el 73% de las clases léxicas se pudo vincular en la ontología (grupos I-III), mientras que el 27 % (tipo IV) no. El 58% (grupos III y IV) requieren dar cuenta de algunos miembros léxicos (III) o de todos (IV) en la plantilla léxica del predicado. El grupo II (19%) requiere un manejo cuidadoso, porque para algunos fines la representación por un concepto superordinado es apropiada, pero para otros no.

Como se trata de clases léxicas seleccionadas por predicados, la codificación en FunGramKB se hace en función del evento conceptual al que se vincula el predicado. La clase léxica se vincula a una entidad y esta se codifica como preferencia de selección del evento. Los argumentos del paradigma semántico quedan representados por la entidad, pero los que acceden a la clase léxica por otras vías son ubicados como colocaciones del predicado. Por ejemplo, el verbo *dictar* selecciona una clase léxica que vinculamos con la entidad +PLAN_00. Los argumentos *modelo*, *patrón*, *pauta*, *estrategia* y *esquema* lexicalizan a +PLAN_00, por lo que no tienen que explicitarse, pero los argumentos *mecanismo*, *método* y *criterio* advienen a la clase por mecanismos discursivos. Así, en la entrada léxica de *dictar*[-COMMAND_PLAN], figuran como colocaciones solo los argumentos que no son predecibles semánticamente: *mecanismo*, *método* y *criterio*. La conexión entre las dos partes de la clase léxica radica ahí: *dictar* selecciona léxicamente a *mecanismo*, *método* y *criterio* cuando selecciona conceptualmente a +PLAN_00.

El ejercicio deliberado de reducción del inventario de clases léxicas requiere que se comparen las clases cuya descripción semántica y conceptual es similar. Hemos observado que los verbos relacionados semánticamente usualmente seleccionan clases léxicas con sentidos similares⁹⁵. El grupo de nueve verbos (*derrochar*, *desaprovechar*, *desperdiciar*, *despilfarrar*, *dilapidar*, *disipar*, *malgastar*, *malograr* y *perder*), en su sentido de ‘malgastar’ o +WASTE_00, se combinó con 46 clases léxicas diferentes en REDES, que se redujeron a 12 entidades conceptuales en FunGramKB. Por su parte, los cinco verbos que evidenciaban polisemia regular (*tejer*, *tramar*, *urdir*, *cocinarse*, *maquinar*) se combinaron con 32 clases léxicas en REDES, que se redujeron a 13 entidades en FunGramKB. El hecho de que las clases pertenezcan a los tipos I, II o III no afecta el proceso de reducir el inventario.

Cabe señalar que para vincular las clases léxicas con entidades conceptuales nos limitamos a usar los conceptos existentes en la subontología de entidades, sin proponer nuevos conceptos terminales o subconceptos. Los conceptos terminales y subconceptos, que tan bien sirvieron para codificar los sentidos predicativos verbales en la subontología de eventos, sin duda serán valiosos para una más precisa codificación de las clases léxicas en la subontología de entidades.

Para manejar datos como las clases léxicas de REDES es necesario contar con una plataforma que dé cuenta de la regularidad semántica de las clases, pero también de la

⁹⁵ Otra estrategia para trabajar con clases léxicas similares es partir del índice conceptual de clases léxicas en REDES y agrupar todas las clases léxicas que estén vinculadas con un mismo concepto.

idiosincrasia de la combinatoria con cada predicado. La plataforma ontológica de FunGramKB nos permitió realizar eso justamente, porque viabilizó el desdoblamiento de algunas clases entre el módulo conceptual y el léxico.

1.4 ¿Pueden estos vínculos revelar patrones o generalizaciones sobre las relaciones sistemáticas entre los sentidos de un predicado y su combinatoria?

Los vínculos entre los datos lingüísticos de REDES y los conceptos de la ontología de FunGramKB permiten visualizar claramente la relación entre cada predicado y su combinatoria léxica. En la gran mayoría de los casos, los sentidos de un predicado se corresponden nítidamente con las clases léxicas. Por esta razón, nos parece que la combinatoria léxica, debidamente traducida al nivel conceptual, puede servir como indicador para desambiguar un predicado en el PLN.

- ¿Puede tener un predicado más de un sentido con la misma combinatoria?

Un predicado puede tener más de un sentido con la misma combinatoria, aunque no es muy frecuente. Ello requiere que el verbo pueda asumir más de un sentido en una misma combinación; de este fenómeno dependen las bromas de “doble sentido”. En nuestros materiales, solo encontramos un caso posible de dos sentidos en una misma combinación. El verbo *perder* en *perder el tren/el avión/el autobús* puede significar ‘no llegar a tiempo’, cuando se trata de un medio de transporte, o ‘dejar de tener’, cuando se trata de un juguete. Sin embargo, al codificar la combinación en la ontología de FunGramKB, no hay posibilidad de ambigüedad. Al vincular la información a conceptos en FunGramKB, tendríamos que elegir entre [\$MISS_00] y ([+TRANSPORT_00]), o [+LOSE_00] y ([+TOY_00]).

- ¿Puede tener un predicado el mismo sentido con diferentes combinaciones?

Un predicado puede tener el mismo sentido con diferentes combinaciones. Esto queda extremadamente claro en los datos de REDES. Es especialmente evidente en las palabras monosémicas, como *malgastar*, *desaprovechar*, *dilapidar*, *despilfarrar* y *desperdiciar*, que se combinan con múltiples clases léxicas pero mantienen el vínculo a un solo concepto: +WASTE_00. Por otro lado, el recurso de los subconceptos permite

establecer diferencias dentro de un mismo concepto a partir de las distintas preferencias de selección del predicado.

En el caso de las palabras polisémicas, cada sentido se asocia a un grupo definido de clases léxicas. Por lo tanto, los diferentes sentidos sí se corresponden a grupos determinados de clases léxicas, pero las diferentes clases léxicas por sí solas no necesariamente implican diferentes sentidos.

1.5 ¿Pueden diversos predicados compartir la misma alternancia de sentidos, es decir, participar en patrones de polisemia regular? ¿Qué relación tiene la polisemia regular de los predicados con su selección léxica sistemática?

Diversos predicados verbales pueden compartir una alternancia de sentidos, y FunGramKB es capaz de dar cuenta de ello. Se encontró un caso de polisemia regular, con patrón directo (vínculo directo a los mismos conceptos), en *tejer*, *tramar*, *urdir*, entre +WEAVE_00 y +CONSPIRE_00. Hubo un segundo patrón de polisemia directa entre +WEAVE_00 y +WRITE_00 entre *tejer*, *urdir* y *cocinar*. Además se encontró un caso de polisemia regular con patrón indirecto (vínculo a conceptos en la misma ruta conceptual), entre +CREATE_00 y +CONSPIRE_00: *tejer*, *tramar*, *urdir*, *cocinarse*.

En el grupo de nueve verbos, encontramos un caso de polisemia regular entre dos verbos: *malograr* y *perder*, que se vinculan con +WASTE_00 y +FAIL_00. Este patrón, sin embargo, no se mantiene a nivel de subconcepto, porque cada predicado tiene preferencias distintas.

1.6 ¿Cómo nos ayudan las nociones de ‘preferencia de selección’ –de naturaleza conceptual– y ‘colocación’ –de naturaleza léxica–, establecidas por FunGramKB, a entender mejor los datos de REDES?

Las nociones de preferencia de selección y de colocación establecidas en FunGramKB nos permiten codificar toda la información provista por las clases léxicas de REDES, sin perder datos ni ser redundantes. Las restricciones de selección léxica que se pueden codificar conceptualmente están presentes como preferencias de selección en el MT de los conceptos, mientras que las que no, pasan a ser colocaciones en el lexicón.

La noción de preferencia de selección nos permite además crear conceptos terminales o subconceptos nuevos para alojar formalmente las preferencias de selección reveladas por

las restricciones léxicas. Mientras, el espacio de colocación de los predicados, en el nivel léxico, nos permite ubicar todos los argumentos que no se puedan vincular a un concepto.

Los espacios de preferencia de selección y de colocación compensan por el hecho de que no es posible capturar en FunGramKB los sentidos figurados que asumen los argumentos para entrar en una combinación dada. Por ejemplo, con el verbo *malgastar* se combinan los argumentos *saliva, línea, palabra, página, óleo, tinta, cinta*, que pertenecen a una clase léxica que vinculamos con +INFORMATION_OBJECT_00. Los argumentos *línea, palabra* lexicalizan a +INFORMATION_OBJECT_00, pero *saliva, página, óleo, tinta, cinta* no se pueden codificar en la ontología como +INFORMATION_OBJECT_00, porque adquieren el sentido de ‘objeto de información’ de forma figurada. Por eso, los ubicamos como colocaciones en la plantilla léxica del verbo *malgastar*[-WASTE_INFORMATION_OBJECT].

1.7 ¿Constituye el vínculo de REDES a FunGramKB un primer paso para convertir a REDES es un diccionario informatizado inteligente?

El vínculo de los datos de REDES puede suponer un primer paso hacia la conversión de REDES en un diccionario informatizado inteligente. Una plataforma adecuada podría combinar los vínculos con la ontología de FunGramKB con información lexicográfica más tradicional.

El diccionario informatizado REDES podría tener la siguiente forma: el verbo se define en un primer nivel con eventos conceptuales (incluidos subconceptos). La descripción conceptual del evento en FunGramKB está disponible apretando un botón. Al apretar cada evento, aparecen una o varias glosas de sentido, que se acercan a definiciones tradicionales. A la vez, cada evento o subconcepto hace claras sus preferencias de selección en forma de entidades conceptuales (tal vez destacadas en otro color). Esas entidades pueden desplegar su información conceptual en FunGramKB, o bien desplegar su manifestaciones léxicas, que se podrían desdoblar en los miembros del paradigma semántico y los miembros del paradigma discursivo o pragmático. Las búsquedas se podrían orientar desde los sustantivos, que se definirían a través de entidades conceptuales. En la entrada también se deben proveer los eventos que seleccionan la entidad. Las búsquedas se deben poder hacer desde las palabras o desde los conceptos, al igual que en FunGramKB. Así, se podría aprovechar al máximo la aportación de cada recurso: la organización ontológica conceptual de FunGramKB y la riqueza lingüística sobre combinatoria sistemática léxica de REDES.

2. Cómo puede servir REDES a FunGramKB

2.1 ¿Cómo pueden los datos de REDES enriquecer a FunGramKB, tanto en el nivel conceptual, por medio de nuevos conceptos y preferencias de selección, como en el nivel léxico, con lexicalizaciones de conceptos y colocaciones?

La información que ofrece REDES sobre restricciones sistemáticas de selección léxica de predicados a argumentos en el español constituye un importante inventario de información lingüística proveniente de corpus y organizado por criterios semánticos, que no tiene par en la lingüística o lexicografía actuales. Luego de completar el trabajo de transferir nuestra muestra de REDES a FunGramKB, concluimos que los datos de REDES pueden ayudar a poblar y enriquecer a FunGramKB, tanto en el módulo conceptual como en el módulo léxico.

Los sentidos de los predicados se vinculan a eventos en la ontología conceptual de FunGramKB. Las restricciones de selección léxica de esos predicados se pueden traducir a preferencias de selección de los eventos. En ocasiones, los conceptos existentes en FunGramKB ya poseen en su definición las preferencias de selección que se reflejan en REDES. Por ejemplo, *tejer* se combina en REDES con una clase léxica descrita como ‘telas y otros materiales, y objetos formados con ellos’ y compuesta por *hilo, punto, mimbre, algodón, lana, seda, tela, cesta, estera, canasto, jersey, camisa, sombrero*; tanto el sentido de *tejer* como su selección léxica están perfectamente representados por el concepto +WEAVE_00, cuyo MT es: (x1: +HUMAN_00)Theme (x2: +CLOTH_00 ^ +BASKET_00)Referent y cuyo PS es: +(e1: +CREATE_00 (x1)Theme (x2)Referent (f1: +HAND_00 | +MACHINE_00)Instrument (f2: +THREAD_00 ^ +STRING_00)Means (f3: +HUMAN_00)Beneficiary). En esos casos, los datos de REDES sirven para validar las definiciones conceptuales de FunGramKB.

En otras ocasiones, las restricciones de selección léxica de REDES pueden apuntar a preferencias de selección conceptuales que no están presentes todavía en la ontología de FunGramKB. Para codificar estas nuevas preferencias de selección conceptuales en FunGramKB, se requiere la creación de nuevos conceptos terminales o subconceptos. Aunque la ontología de FunGramKB se considera terminada en los niveles metaconceptual y básico, se continúa poblando el nivel de conceptos terminales y de subconceptos. Por lo tanto, este procedimiento enriquece la ontología conceptual de FunGramKB. Cabe señalar que, aparte de las particularidades de selección léxica, la información de REDES puede revelar también alguna matización semántica específica del predicado. Esto puede justificar

la creación de nuevos conceptos terminales y subconceptos, usualmente por medio de añadir o ajustar un parámetro a un concepto básico. Del trabajo salieron cuatro nuevos conceptos terminales (\$DICTATE_00, \$MISS_00, \$DISSIPATE_00 y \$SQUANDER_00) y 37 nuevos subconceptos, por ejemplo, bajo *perder*[+FAIL_00]: -FAIL_TRIAL, -FAIL_BATTLE, -FAIL_COMPETITION.

Los datos de REDES también aportan información léxica que ayuda a poblar el módulo léxico de FunGramKB. Recordemos que el trabajo fundacional de FunGramKB se ha hecho en el módulo conceptual, y el módulo léxico está en proceso de ser poblado. Los predicados verbales se convierten en lexicalizaciones de eventos de la ontología, según cada sentido que adquieran. Si el verbo no tenía entrada léxica en el lexicón, su adición constituye una nueva entrada. Por ejemplo, el evento conceptual +WASTE_00 solo contaba con una lexicalización en español: *malgastar*. A partir de nuestro trabajo, se pueden añadir ocho verbos más (*peder, disipar, derrochar, malograr, desperdiciar, dilapidar, desaprovechar, despilfarrar*), cada uno de los cuales constituirá una nueva entrada léxica en el lexicón. Por su parte, las clases léxicas nominales se equiparan a entidades conceptuales: los argumentos que pertenecen al paradigma semántico de la clase léxica se convierten en lexicalizaciones de la entidad conceptual. Por ejemplo, una clase léxica que se combina con *tejer*[+CONSPIRE_00] se vincula a +PLAN_00. La mayoría de los argumentos miembros de la clase léxica pueden lexicalizar a +PLAN_00: *estrategia, plan, proyecto*. Los argumentos que no puedan adherirse al concepto asignado a la clase léxica –porque adquieren ese sentido de manera figurada– se ubican en el lexicón, como colocaciones del predicado. En el ejemplo anterior, solo el argumento *trampa* se ubica como colocación en la plantilla léxica de *tejer*[+CONSPIRE_00].

Los datos de REDES ponen a prueba el modelo de FunGramKB, que busca representar la totalidad del conocimiento lingüístico humano. La prueba demuestra que FunGramKB es capaz de acomodar los datos de restricción de selección léxica de REDES.

2.2 ¿Tiene la relación entre un predicado y su combinatoria léxica en REDES un correlato conceptual en la ontología de FunGramKB?

La relación entre un predicado y su combinatoria léxica sistemática tiene un correlato conceptual en la relación entre un evento y las entidades que constituyen sus preferencias de selección. La preferencia de selección, de hecho, fija la relación entre un evento y una entidad en el MT y en el PS de un concepto. Por ejemplo, *perder* tiene un sentido que se

vincula a +LOSE_00, cuyo MT es: (x1: +HUMAN_00^+ANIMAL_00)Theme (x2: +CORPUSCULAR_00)Referent. En el concepto +LOSE_00, existe una relación conceptual, pues, entre un Tema humano o animal y un Referente corpuscular. La clase léxica con que se combina *perder* en ese sentido se vinculó con +SOLID_00, que es subordinado de +CORPUSCULAR_00. Esto quiere decir que pudimos vincular ese sentido de *perder* a +LOSE_00 sin hacer ningún ajuste. A veces, sin embargo, la combinatoria léxica presentada en REDES no cabe dentro de las preferencias de selección existentes en FunGramKB. En esos casos, echamos mano de los conceptos terminales o subconceptos para codificarla. Por ejemplo, la combinación sistemática de *perder* con nombres como *partido*, *campeonato*, *encuentro* da pie al subconcepto -FAIL_COMPETITION, mientras que las selecciones léxicas sistemáticas compartidas por *malgastar*, *desperdiciar*, *dilapidar*, etc. generan los subconceptos -WASTE_ABILITY, -WASTE_TIME y -WASTE_MONEY.

Además de los patrones de selección léxica de los verbos individuales, el eje conceptual revela patrones de selección léxica compartidos por múltiples verbos. De los nueve predicados vinculados a +WASTE_00, seis escogieron +ABILITY_00, seis a TIME_00, cinco a +MONEY_00, etc. Estos vínculos hacen claro que la combinatoria sistemática entre predicados y grupos de argumentos responden, en gran medida, a relaciones de selección, o de concordancia semántica, entre eventos y entidades conceptuales. Dada la relación que se revela entre +WASTE_00 y +ABILITY_00, TIME_00, +MONEY_00, podríamos decir que, a nivel cognitivo, la noción de ‘malgastar’ elige prototípicamente nociones como ‘talento’, ‘tiempo’ y ‘dinero’.

2.3 ¿Pueden los datos de combinatoria léxica de REDES aportar al tratamiento de la polisemia en la ontología de FunGramKB?

Los datos de REDES pueden aportar al tratamiento de la polisemia en el PLN. En las palabras polisémicas, hemos encontrado una relación clara entre los sentidos y la combinatoria léxica establecida en REDES (no necesariamente clases léxicas solas, sino conjuntos de clases léxicas). Por ejemplo, *disipar* tiene tres sentidos –‘esparcir’, ‘malgastar’, ‘hacer desaparecer’– y se combina con 12 clases léxicas; la distribución de las clases léxicas está definida entre los tres sentidos: ‘esparcir’–2 clases léxicas, ‘malgastar’–1 clase léxica, ‘hacer desaparecer’–9 clases léxicas. En la medida que la combinatoria léxica sistemática se codifique en términos conceptuales, la desambiguación automática se hace viable, además de que adquiere poder predictivo. Sin embargo, dada la naturaleza de las clases léxicas de

REDES, parte de la información de combinatoria estará codificada en términos conceptuales, como preferencias de selección, y parte en términos léxicos, como colocaciones, por lo que un programa de WSD para el PLN tendrá que acceder a ambos módulos.

Hemos observado que los verbos polisémicos con los que trabajamos tienen un sentido –que llamamos “sentido dominante” en la discusión– que ocurre con la mayoría de las combinaciones léxicas. Esto confirma el hecho de que las palabras tienen un sentido principal, que ha sido señalado por otros investigadores (Hanks, 2013; Velardi, 1991). Los segundos, terceros y cuartos sentidos se dan con combinatorias más restringidas. La tabla siguiente hace claro cuál es el sentido dominante para cuatro verbos:

Tabla 329. Cantidad de clases léxicas por sentido por predicado

| Sentido | Predicado | | | |
|-------------|------------------|-----------------|----------------|---------------|
| | <i>derrochar</i> | <i>malograr</i> | <i>disipar</i> | <i>perder</i> |
| +WASTE_00 | 2 | 2 | 2 | 2 |
| +FAIL_00 | | 6 | | 2 |
| \$ABOUND_00 | 15 | | | |
| +SPLIT_00 | | | 2 | |
| +FINISH_00 | | | 10 | |
| \$MISS_00 | | | | 2 |
| +LOSE_00 | | | | 11 |

Ante este hecho, se podrían establecer sentidos “por defecto” o “no marcados”, y programar las aplicaciones de WSD para la detección de los sentidos “marcados” solamente, que se darían con la combinatoria más restringida.

2.4. ¿Puede la estructura de FunGramKB identificar patrones de polisemia regular?

La ontología de FunGramKB es capaz de dar cuenta de patrones de polisemia regular directa e indirecta en verbos. El trabajo en FunGramKB reveló dos casos de polisemia regular directa en los predicados que estudiamos: entre +WEAVE_00 y +CONSPIRE_00 alternan los verbos *tejer*, *urdir*, *tramar*, y entre +WEAVE_00 y +WRITE_00 están los verbos *tejer*, *urdir* y *cocinar*. Hubo un caso de polisemia regular indirecta, entre +CREATE y +CONSPIRE_00, en los verbos *tejer*, *urdir*, *tramar* y *cocinar(se)*.

No hay mucho trabajo sobre la polisemia regular en verbos, por lo que es valioso constatar que FunGramKB es una herramienta que puede dar cuenta de la polisemia regular

en esta categoría gramatical. La jerarquía ontológica de FunGramKB demostró que puede dar cuenta tanto de alternancias directas, donde los predicados están vinculados directamente a los conceptos, como de alternancias indirectas, es decir, de alternancias que se dan entre rutas conceptuales, como es el caso del patrón +CREATE_00 y \$CONSPIRE_00. Las alternancias indirectas requieren un trato cuidadoso; habría que comprobar, por ejemplo, si todas las palabras que se vinculan con +CREATE_00 (el concepto superordinado común a varios verbos) se podrían también vincular con \$CONSPIRE_00, lo cual parece improbable.

El modelo de FunGramKB también nos permite explorar la polisemia regular en el nivel del subconcepto. Aunque no encontramos ningún caso con el mínimo de tres verbos, hay dos verbos (*tejer*, *urdir*) que alternan entre el concepto terminal \$CONSPIRE_00 y el subconcepto -WRITE_STORY. La validación de un patrón de polisemia regular a nivel de subconcepto establece un nivel mayor de detalle y precisión en la alternancia. Por otra parte, un patrón de polisemia regular que se da en el nivel de conceptos básicos o terminales, se puede invalidar en el nivel de subconcepto. Por ejemplo, los verbos *perder* y *malograr* alternan entre +WASTE_00 y +FAIL_00. Sin embargo, bajo +WASTE_00, *perder* se vincula al subconcepto -WASTE_PERIOD y *malograr* a -WASTE_ABILITY y -WASTE_TIME. Bajo +FAIL_00, *perder* se vincula a -FAIL_TRIAL, -FAIL_BATTLE y -FAIL_COMPETITION, mientras que *malograr* se vincula con -FAIL_PLAN, -FAIL_EFFORT y -FAIL_VICTORY. Los subconceptos, por lo tanto, nos ofrecen un nivel de mayor precisión para la constatación de la polisemia regular: no solo consideran la alternancia entre significados de conceptos sino también entre sus preferencias de selección.

El marco ontológico de FunGramKB es capaz de dar cuenta de la polisemia regular directa e indirecta en verbos, en múltiples niveles: concepto básico y terminal, e incluso subconceptos. El subconcepto presenta un nivel de mayor precisión en el estudio de la polisemia regular, pues no solo considera el significado del concepto sino también sus preferencias de selección específicas.

2.5 ¿Nos permiten los datos de REDES identificar principios y patrones regulares que den consistencia a las nociones clave para el PLN de “preferencia de selección” y “colocación”, establecidas por FunGramKB?

Los datos de REDES validan la noción de “preferencia de selección”, y ajustan la noción de “colocación” según establecida por FunGramKB, aprovechando el apartado de colocaciones en FunGramKB de una manera particular.

Los datos semánticos de selección léxica se acomodan bien en las preferencias de selección de nivel conceptual y, viceversa, las preferencias de selección dan el espacio adecuado para codificar las restricciones semánticas de los predicados, a partir de su selección léxica. Por ejemplo, el que *perder* se combine con ‘capacidades’ como *vista*, *oído* y *olfato* (*perder la vista*, *el oído*, *el olfato*) se puede traducir a [+LOSE_00] y a [+SENSE_00], lo que da pie al subconcepto -LOSE_SENSE_00. El marco ontológico nos permite codificar las preferencias de selección conceptualmente y crear un nuevo evento semánticamente restringido.

Existe una diferencia entre las preferencias de selección de nuestro trabajo y las codificadas por Jiménez y Pérez (2011) en casos como *calzar*[-TAKE_SHOES_00] o *empuñar*[-WIELD] o *traspapelar*[-MISPLACE_00]. En el caso de *calzar*, *empuñar* y *traspapelar*, las restricciones conceptuales están presentes en el significado léxico de los verbos mismos: *calzar* implica ‘zapatos’, *empuñar* implica ‘un arma con puño’ y *traspapelar* implica ‘documentos’. Esas preferencias de selección también se manifiestan en selecciones léxicas: *calzar zapatos*, *empuñar el cuchillo* y *traspapelar el contrato*, respectivamente. En la mayoría de nuestros verbos, la preferencia de selección se manifiesta en la combinatoria léxica, pero no está presente tan obviamente en el significado léxico del verbo. Por ejemplo, tanto *perder* como *malgastar* tienen una preferencia de selección por +TIME_00, pero esto no es deducible del significado de *perder* o *malgastar*, a diferencia de la preferencia de *calzar* por +SHOES_00. No obstante lo anterior, los datos de REDES validan esas restricciones de selección léxica como fenómenos semánticos sistemáticos, por lo que consideramos que se trata del mismo tipo de preferencia de selección conceptuales.

La noción de colocación en FunGramKB se refiere a las ‘combinaciones de lexemas que común y frecuentemente co-ocurren en la lengua’. Esa definición amplia puede incluir todos los argumentos de todas las clases léxicas de REDES. Como hemos visto a lo largo del trabajo, nosotros hemos dividido la clase léxica de REDES entre los argumentos que se pueden describir conceptualmente en la ontología y los que no. Ese segundo grupo, en nuestro tratamiento, se vacía en el espacio de colocaciones de FunGramKB. De este modo, no perdemos ninguna información sobre la clase léxica. En nuestro trabajo, por lo tanto, esos “lexemas que común y frecuentemente co-ocurren en la lengua” los hemos dividido entre la preferencia de selección conceptual y las colocaciones léxicas. Este procedimiento refuerza la noción de “colocación” en FunGramKB como un fenómeno particular a cada lengua, por

algunos “colocados” se pueden describir en términos conceptuales universales, y lo que no, se consideran privativos de la lengua en cuestión.

3. Futuras líneas de investigación

El trabajo realizado en esta tesis sugiere múltiples líneas de investigación futura. En primer lugar, sería útil someter a pruebas de WSD a los verbos polisémicos con los que hemos trabajado –por ejemplo, *perder* o *disipar*–, usando la codificación hecha en FunGramKB. Así podríamos determinar si la codificación de las preferencias de selección de los eventos, basadas en las restricciones de selección léxica de los predicados en REDES, viabilizan la desambiguación automática de los predicados en corpus.

Con respecto a la polisemia regular, sería útil poner a prueba las alternancias de sentidos descubiertas, con el objetivo de ver hasta qué punto pueden auxiliar los procesos de WSD. Habiendo identificado varios patrones, sería provechoso seguir investigándolos partiendo desde los conceptos, para contestar preguntas como, por ejemplo: ¿qué otros predicados, en español y en otras lenguas, pueden lexicalizar la alternancia entre +WEAVE_00 y +CONSPIRE_00, o las otras alternancias identificadas? Asimismo, nos gustaría seguir explorando en FunGramKB otros grupos verbales que sugieren alternancias regulares de sentidos. Finalmente, se debe auscultar hasta qué punto los patrones de alternancia indirecta son válidos.

Conviene también explorar en otras lenguas todas las preferencias de selección identificadas. Si bien entendemos que los datos de REDES –por reflejar sistematicidad en la lengua– justifican la creación de preferencias de selección de nivel conceptual, sería provechoso estudiarlas en las otras lenguas incluidas en FunGramKB, y añadir las lexicalizaciones correspondientes.

Se debería extender este trabajo a otros grupos verbales relacionados semánticamente, que amplíen la red de los eventos que demostraron puntos de contacto con +WASTE_00, por ejemplo, +LOSE_00, +FAIL_00, \$ABOUND, etc. Al igual que exploramos exhaustivamente los verbos que se vinculan con +WASTE_00 y su combinatoria, nos gustaría hacerlo con otros verbos. Un objetivo sería constatar si los verbos relacionados semánticamente mantienen patrones similares en su selección léxica, algo que confirmamos en los verbos vinculados a +WASTE_00 y en los verbos vinculados a \$CONSPIRE_00.

Nos gustaría poner el foco en las entidades conceptuales, y a partir de ellas estudiar las clases léxicas en REDES. Por ejemplo, nos acercamos a las clases léxicas que están vinculadas a +ABILITY_00 en virtud de su combinación con +WASTE_00 y \$ABOUND_00. Podemos preguntarnos: ¿qué otros verbos se predicán de +ABILITY_00? ¿Y qué adjetivos? ¿Cómo se manifiesta +ABILITY_00 entonces?

Los vínculos establecidos entre clases léxicas y entidades pueden someterse a examen más riguroso. Por ejemplo, ¿son las clases del tipo II válidas para el PLN? Es posible que al estar descritas por conceptos superordinados, el concepto resulte demasiado amplio para la realidad léxica de la clase. Por otro lado, en las clases del tipo III y IV, valdría la pena acercarse a los argumentos que advienen a las clases léxicas por mecanismos discursivos y pragmáticos, y caracterizar con más detalle esos mecanismos. Esos datos podrían dar pie a un estudio sistemático de metonimia, metáfora, extensión semántico y otros procesos pragmáticos. Asimismo, podríamos tratar de precisar la codificación conceptual de las clases léxicas, usando el recurso de conceptos terminales y subconceptos; tal vez se podrían reducir las clases léxicas que no se pudieron vincular (tipo IV), formulando nuevos conceptos terminales o subconceptos en la subontología de entidades.

Finalmente, se podrían usar los resultados de este estudio para comenzar a traducir el diccionario REDES a un formato electrónico. Una plataforma adecuada combinaría los vínculos de predicados y clases léxicas con conceptos ontológicos con información lexicográfica más tradicional. Las búsquedas se harían tanto desde la palabra como desde el concepto, y se podrían explorar mediante enlaces electrónicos las relaciones predicado-argumento, y entre evento-entidad. Así se aprovecharía al máximo la aportación de cada recurso: la organización ontológica conceptual de FunGramKB y la riqueza lingüística sobre combinatoria sistemática léxica de REDES.

BIBLIOGRAFÍA

- Abel, A. y Weber, V. (2000). ELDIT. A Prototype of an Innovative Dictionary. En U. Heid, S. Evert, E. Lehmann y C. Rohrer (Eds.), *Proceedings of the Ninth Euralex International Congress, EURALEX 2000* (pp. 807–18). Stuttgart, Alemania: Universitat Stuttgart.
- ADESSE: *Alternancias de Diátesis y Esquemas Sintáctico-Semánticos del Español*. Universidad de Vigo. En <http://adesse.uvigo.es/>.
- Agirre, E. y Edmonds, P. (2007). Introduction. En E. Agirre y P. Edmonds (Eds.), *Word Sense Disambiguation: Algorithms and Applications* (pp. 1–28). Springer.
- Agirre, E. y Martínez, D. (2001). Learning class-to-class selectional preferences. *Proceedings of the International Conference on Computational Linguistics (COLING)* (pp. 16-22). Copenhagen, Dinamarca.
- Agirre, E. y Stevenson, M. (2007). Knowledge sources for WSD. En E. Agirre y P. Edmonds (Eds.), *Word Sense Disambiguation: Algorithms and Applications* (pp. 217-251). Springer.
- Allen, J. F. (1983). Maintaining knowledge about temporal intervals. *Communications of the ACM*, 26(11), 832-843.
- Alonso Ramos, M. (Directora). *Diccionario de colocaciones del español*. Universidad de A Coruña. En <http://stel.ub.edu/dice/paginas>
- Altman, J. y Polguère, A. (2003). La BDéf: base de définitions dérivée du Dictionnaire explicatif et combinatoire. En *Proceedings of the First International Conference on Meaning-Text Theory* (pp. 43-54). París, Francia.
- Andersen, B. y Nielsen, S. (2009). Ten Key Issues in Lexicography for the Future. En H. Bergenholtz, S. Nielsen y S. Tarp (Eds.), *Lexicography at a Crossroads. Dictionaries and Encyclopedias today, Lexicographical Tools tomorrow* (pp. 355-365). Berna, Suiza: Peter Lang.
- Apresjan, J. (1973). Regular Polysemy. *Linguistics*, 142(5), 5-32.
- Atkins, B. T. S. (1996). Bilingual Dictionaries: Past, Present and Future. En M. Gellerstam, J. Jarborg, S. G. Malmgren, K. Noren, L. Rogstrom y C. R. Pappmehl (Eds.), *Proceedings I-II, Papers submitted to the Seventh EURALEX International Congress on Lexicography* (pp. 515–46). Gotemburgo, Suecia: Goteborg Universitet.
- Baker, C. y Rupperhofer, J. (2002). FrameNet's Frames Vs. Levin's Classes. Obtenido de: <http://www.icsi.berkeley.edu/cgi-bin/pubs/publication.pl?ID=116>
- Bar-Hillel, Y. (1960). The present status of automatic translation of languages. En F. Alt (Ed.), *Advances in Computers* (pp. 91–163). NY: Academic Press.
- Baroni, M. y Zamparelli, R. (2010). Nouns are vectors, adjectives are matrices: representing

- adjective-noun constructions in semantic space. En *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, EMNLP '10 (pp.1183–1193), Stroudsburg, PA: Association for Computational Linguistics.
- Barque, L. (2007). The fickle polysemy of emotions. En *Third International Conference on Meaning-Text Theory*. Klagenfurt, Austria.
- Barque, L. y Chaumartin, F. R. (2006). Regular Polysemy in WordNet. *Journal for Language Technology and Computational Linguistics*, 24(2), 1-14.
- Barque, L. y Polguère, A. (2005). Application du métalangage définitionnel de la BDéf au traitement formel de la polysémie. En *Actes de TALN 2005* (pp. 391-397). Dourdan, Francia.
- Base de Datos Sintácticos del Español Actual*. Universidad de Santiago de Compostela. En <http://www.bds.usc.es/>
- Bateman, J. A., Henschel, R. y Rinaldi, F. (1995). The Generalized Upper Model 2.0. Technical report. IPSI/GMD, Darmstadt. Obtenido de: <http://www.fb10.uni-bremen.de/anglistik/.../gum-2.pdf>
- Batiukova, O. (2009a). Aplicaciones lexicográficas de la teoría del lexicón generativo. En E. de Miguel & et al. (Eds.), *Fronteras de un diccionario: las palabras en movimiento* (pp. 233-270). Fundación San Millán de la Cogolla.
- Batiukova, O. (2009b). La teoría del léxico en los nuevos diccionarios. En de Miguel, E. (Ed.), *Panorama de la lexicología* (pp. 487-519). Barcelona: Ariel.
- Benson, M., Benson, E. y Ilson, R. (1986). *The BBI Combinatory Dictionary of English. A Guide to Word Combinations*. Amsterdam: John Benjamins.
- Bergenholtz, H., Nielsen, S. y Tarp, S. (Eds.). (2009). *Lexicography at a Crossroads: Dictionaries and Encyclopedias Today, Lexicographical Tools Tomorrow*. Berna, Suiza: Peter Lang.
- Bierwisch, M. (1982). Linguistics and language error. En A. Cutler (Ed.), *Slips of the tongue and language production* (pp. 29-72). Amsterdam: Mouton Publishers.
- Boas, H. C. (2003). *A constructional approach to resultatives*. Stanford: CSLI Publications.
- Boas, H. C. (2005). Semantic Frames as Interlingual Representations for Multilingual Lexical databases. *International Journal of Lexicography*, 1-34.
- Boas, H. C. (2008a). Resolving Form-meaning Discrepancies in Construction Grammar. En J. Leino (Ed.), *Constructional reorganization* (pp. 11-36). Amsterdam & Philadelphia: John Benjamins.
- Boas, H. C. (2008b). Determining the structure of lexical entries and grammatical constructions in Construction Grammar. *Annual Review of Cognitive Linguistics* 6, 113-144.

- Boas, H. C. (Ed.) (2009a). *Multilingual FrameNets in Computational Lexicography: Methods and Applications*. Berlín: Mouton de Gruyter.
- Boas, H. (2009b). Semantic Frames as Interlingual Representations for Multilingual Lexical Databases. En H. C. Boas (Ed.), *Multilingual FrameNets in Computational Lexicography: Methods and Applications* (pp. 59-99). Berlín: Mouton de Gruyter.
- Boas, H. C. (2009c). Introduction: Recent trends in multilingual computational lexicography. En H. C. Boas (Ed.), *Multilingual FrameNets in Computational Lexicography: Methods and Applications* (pp. 1-26). Berlín: Mouton de Gruyter.
- Boas, H. C., Langer, S. y Schnorbusch, D. (2005). From Theory to Practice: Frame Semantics and the Design of FrameNet. En S. Langer y D. Schnorbusch (Eds.), *Semantisches Wissen im Lexikon* (pp. 129-160). Tübinga: Narr.
- Boleda, G., Padó, S. y Utt, J. (2012). Regular polysemy: A distributional model. En *First Joint Conference on Lexical and Computational Semantics (SEM 2012)* (pp. 151-160). Montreal, Canadá: Association for Computational Linguistics.
- Bosque, I. (2001a). On the weight of light predicates. En J. Herschenson, E. Mallén y K. Zagona (Eds.), *Features and interfaces in romance. Essays in honor of Heles Contreras* (pp. 23-38). Amsterdam: John Benjamins.
- Bosque, I. (2001b). Bases para un diccionario de restricciones léxicas. *Moenia*, (7), 11-52.
- Bosque, I. (2001c). Sobre el concepto de colocación y sus límites. *Lingüística Española Actual*, 23(1), 9-40.
- Bosque, I. (2001d). Sobre el concepto de 'lugar común' desde el punto de vista gramatical. En M. Franco & M. Olmos (Eds.), *Lieu(x) commun(s)*, número temático de *Pandora Revue d'études hispaniques du Département d'Espagnol*. París: Universidad de París-8.
- Bosque, I. (2004a). La direccionalidad en los diccionarios combinatorios y el problema de la selección léxica. En T. Cabré (Ed.), *Lingüística teòrica: anàlisi i perspectives (Catalan Journal of Linguistics, Monografies)* (pp. 13-58). Bellaterra: Universitat Autònoma de Barcelona.
- Bosque, I. (2004b). *REDES Diccionario combinatorio del español contemporáneo*. Madrid, SM.
- Bosque, I. (2006). *Diccionario combinatorio práctico del español contemporáneo*. Madrid, SM.
- Bosque, I. y Mairal, R. (2012a). Definiciones Mínimas. En F. Rodríguez González (Ed.), *Estudios de lingüística española. Homenaje a Manuel Seco* (pp. 123-136). Universidad de Alicante.

- Bosque, I. y Mairal, R. (2012b). Hacia una organización conceptual del *definiens*. Capas nocionales del adverbio *arriba*. En D. Corbella *et al.* (Coords.), *Lexicografía hispánica del siglo XXI: nuevos proyectos y perspectivas. Homenaje al profesor Cristóbal Corrales Zumbado*. Madrid: ArcoLibros.
- Bouaud, J., Bachimont, B., Charlet, J. y Zweigenbaum, P. (1995). Methodological principles for structuring an ontology. CHU-Pitié-Salpetriere París 6: Departement Intelligence Artificielle et Medicine.
- Briscoe, T. y Copestake, A. (1991). Sense extensions as lexical rules. En D. Fass, E. Hinkelman y J. Martin (Eds.), *Proceedings of IJCAI Workshop on Computational Approaches to Non-literal Language* (pp. 12-20). University of Colorado at Boulder: Department of Computer Science.
- Briscoe, T. y Copestake, A. (1999). Lexical rules in constraint-based grammar. *Computational Linguistics*, 25(4), 487-526.
- British National Corpus* (BNC). Brigham Young University. En <http://corpus.byu.edu/bnc/>
- Briz, A. (Director). *Diccionario de partículas discursivas del español*. En <http://textodigital.com/P/DDPD/>
- Brockman, C. y Lapata, M. (2003). Evaluating and combining approaches to selectional preference acquisition. En *Proceedings of the European Association for Computational Linguistics (EACL)* (pp. 27-34). Budapest, Hungría.
- Brown Corpus*, creado por Francis, N. y Kucera, H. (1964). Brown University.
- Brugman, C. (1988). *The story of over: Polysemy, semantics and the structure of the lexicon*. NY: Garland Press.
- Buitelaar, P. (1998a). *CoreLex: Systematic Polysemy and Underspecification* (Tesis doctoral, Brandeis University, MA).
- Buitelaar, P. (1998b). CoreLex: an ontology of systematic polysemous classes. En *Proceedings of FOIS98, International Conference on Formal Ontology in Information Systems*. Trento, Italia.
- Butler, C. S. (2009). The Lexical Constructional Model: Genesis, strengths and challenges. En C. S. Butler y J. Martín Arista (Eds.), *Deconstructing Constructions* (pp. 117-152). Amsterdam/Philadelphia: John Benjamins.
- Butler, C. S. y Martín Arista, J. (Eds.). (2009). *Deconstructing Constructions*. Amsterdam/Philadelphia: John Benjamins.
- Cambridge Dictionaries Online*. En <http://dictionary.cambridge.org>
- Chomsky, N. (1986). *Knowledge of language*. NY: Praeger.

- CLAVE. *Diccionario de Uso del Español Actual*. (2000). Madrid: SM. En página web: <http://clave.librosvivos.net>
- Collins Cobuild English Language Dictionary*. (1987). Londres: Collins.
- Copetake, A. y Briscoe, T. (1991). Lexical Operations in a Unification Based Framework. En *Proceedings of ACL SIGLEX Workshop on Lexical Semantics and Knowledge Representation* (pp. 88-101). Berkeley, CA.
- Copetake, A. y Briscoe, T. (1995). Semi-productive polysemy and sense extension. *Journal of Semantics*, 12, 15-67.
- Corominas, J. (1954). *Diccionario crítico etimológico de la lengua castellana*. Madrid: Gredos.
- Corpus of Contemporary American English (COCA)*. Brigham Young University. En página web: <http://corpus.byu.edu/coca/>.
- Corpus Diacrónico del Español (CORDE)*. Real Academia Española. En <http://corpus.rae.es/cordenet.html>.
- Corpus de Referencia del Español Actual (CREA)*. Real Academia Española. En <http://corpus.rae.es/creanet.html>.
- Croft, W. (2012). *Verbs: Aspect and Causal Structure*. Oxford: Oxford University Press.
- Davies, M. *Corpus del Español*. Brigham Young University. En <http://www.corpusdelespanol.org/>.
- de Miguel, E. (Ed). (2009). *Panorama de la lexicología*. Barcelona: Ariel.
- de Miguel, E. (2009b). La teoría del lexicón generativo. En E. de Miguel (Ed.), *Panorama de la lexicología* (pp. 339-369). España: Ariel.
- De Schryver, M. (2003). Lexicographers' dreams in the electronic-dictionary age. *International Journal of Lexicography*, 16(2). Oxford: Oxford University Press.
- De Schryver, G. y Prinsloo, D. (2001). Fuzzy SF: Towards the ultimate customised dictionary. *Studies in Lexicography*, 11(1), 97-111.
- Diccionario de Americanismos*. (2010). Asociación de Academias de la Lengua Española. Madrid: Santillana.
- Diccionario de la Real Academia Española* (XXII edición, actualizada en Internet). Real Academia Española y Asociación de Academias de la Lengua Española. Madrid: Espasa. En <http://rae.es/rae.html>.
- Diccionario Salamanca de la lengua española*. (1996). Madrid: Santillana. En <http://fenix.cnice.mec.es/diccionario>.

- Dolan, W. (1994). Word Sense Ambiguation: Clustering Related Senses. *Association for Computational Linguistics*. Obtenido de: acl.ldc.upenn.edu/C/C94/C94-2113.pdf.
- Dodd, W. S. (1989). Lexicomputing and the Dictionary of the Future. En R. R. K. Hartman (Ed.). (2003) *Lexicography: Critical Concepts*. NY: Routledge.
- Dowty, D. (1979). *Word Meaning and Montage Grammar*. Dordrech, Holanda: Reidel.
- Dowty, D. (2000). The Garden Swarms with Bees and the Fallacy of Argument Alternation. En Y. Ravin y C. Leacock (Eds.), *Polysemy: Theoretical and Computational Approaches* (pp. 111-128). Oxford: Oxford University Press.
- DSO Corpus of Sense-Tagged English*. En <http://www.ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC97T12>
- EuroWordNet*. En <http://www.illc.uva.nl/EuroWordNet/>.
- Faber, P. y Mairal, R. (2003). Representación léxica y esquemas léxicos. En M. A. Martí Antonín, A. Fernández Montraveta y G. Vázquez García (Eds.), *Lexicografía computacional y semántica* (pp. 35-59). Barcelona: Universidad de Barcelona.
- Fauconnier, G. (1985). *Mental Spaces*. Cambridge, MA: MIT Press.
- Fellbaum, C. (Ed.). (1998). *WordNet: An Electronic Lexical Database*. Cambridge, MA: MIT Press. Obtenido de: <http://wordnet.princeton.edu/>.
- Fellbaum, C. (2000). Autotropony. En Y. Ravin y C. Leacock (Eds.), *Polysemy: Theoretical and Computational Approaches* (pp. 91-110). Oxford: Oxford University Press.
- Fillmore, C. J. (1976). Frame semantics and the nature of language. *Annals of the New York Academy of Sciences: Conference on the Origin and Development of Language and Speech*, 280, 20-32.
- Fillmore, C. J. (1977). Scenes-and-frames semantics. En A. Zampolli (Ed.), *Linguistic Structures Processing*. Holanda: North Holland Publishing.
- Fillmore, C. J. (1982). Frame semantics. *Linguistics in the Morning Calm* (pp. 111-137). Seúl, Sur Korea: Hanshin Publishing Co.
- Fillmore, C. J. (1985). Frames and the semantics of understanding. *Quaderni di Semantica*, 6, 222-254.
- Fillmore, C. J. (1992). 'Corpus linguistics' vs. 'computer-aided armchair linguistics'. *Directions in Corpus Linguistics: Proceedings from a 1991 Nobel Symposium on Corpus Linguistics* (pp. 35-66). Estocolmo: Mouton de Gruyter.
- Fillmore, C. J. y Atkins, B. T. S. (1992). Towards a frame-based lexicon: The semantics of RISK and its neighbors. En A. Lehrer y E. Kittay (Eds.), *Frames, Fields and Contrasts: New Essays in Semantics and Lexical Organization* (pp. 75-102). Hillsdale: Lawrence Erlbaum Associates.

- Fillmore, C. J. y Atkins, B. T. S. (2000). Describing Polysemy: The Case of Crawl. En Y. Ravin y C. Leacock (Eds.), *Polysemy: Theoretical and Computational Approaches* (pp. 91-110). Oxford: Oxford University Press.
- Fillmore, C. J. y Baker, C. F. (2001). Frame Semantics for Text Understanding. En *Proceedings of WordNet and Other Lexical Resources Workshop*. (North American Association for Computational Linguistics.) Pittsburgh. En [5http://framenet.icsi.berkeley.edu/*framenet/Papers.html4](http://framenet.icsi.berkeley.edu/*framenet/Papers.html4).
- Fillmore, C. J., Johnson, C. R. y Petruck, M. R. L. (2003). Background to Framenet. *International Journal of Lexicography*, 16(3), 235-250.
- Fillmore, C. J., Petruck, M. R. L., Ruppenhofer, J. y Wright, A. (2003b). FrameNet in Action: The Case of Attaching. *International Journal of Lexicography* 16(3): 297-332.
- Fontenelle, T. (2000). Introduction: Dictionaries, Thesauri and Lexical-Semantic Relations. *International Journal of Lexicography* 13(4), 229-31.
- FrameNet*. En <https://framenet.icsi.berkeley.edu/fndrupal/>.
- Functional Grammar Knowledge Base (FunGramKB)*. En <http://www.fungramkb.com/>.
- Fuertes-Olivera, P. A. y Tarp, S. (2011). Lexicography for the third millennium: Cognitive-oriented specialised dictionaries for learners. *Iberica*, 21, 141-161.
- Fuertes-Olivera, P. A. y Nielsen, S. (2011). Online dictionaries for assisting translators of LSP texts: the accounting dictionaries. *International Journal of Lexicography*, 25(2), 191-215.
- Gale, W., Church, K. y Yarowsky, D. (1992). Estimating upper and lower bounds on the performance of word-sense disambiguation programs. En *Proceedings of the 30th Annual Meeting of the Association for Computational Linguistics (ACL)* (pp. 249-256). Newark, DE.
- Gangemi, A., Guarino, N., Masolo, C., Oltramari, A. y Schneider, L. (2002). Sweetening ontologies with DOLCE. En A. Gómez y V. Richards (Eds.) *Knowledge engineering and knowledge management. Ontologies and the Semantic Web: 13th International Conference* (pp. 166-181). Sigüenza, España.
- Garrido, N. y Ruiz de Mendoza, F. (2011). La modelación del conocimiento procedimental en el Cognición de FunGramKB: una propuesta desde los supuestos del Modelo Léxico Construccional. *AngloGermanica online*, 8, 106-120.
- Global WordNet*. En <http://www.globalwordnet.org>
- Global Glossary*. En <http://www.globalglossary.org>
- Goldberg, A. (1995). *Constructions. A Construction Grammar approach to argument structure*. Chicago: University of Chicago Press.

- Goldberg, A. (2006). *Constructions at work: The nature of generalization in language*. New York: Oxford University Press.
- Grefenstette, G. (1998). The Future of Linguistics and Lexicographers: Will there be Lexicographers in the year 3000? *Euralex '98 Proceedings, 1*, 25-41. Lieja, Bélgica.
- Grimshaw, J. (1979). Complement Selection and the Lexicon. *Linguistic Inquiry, 10*, 279-326.
- Guillaume P. 2009. Cross-lingual labeling of semantic predicates and roles: A low-resource method based on bilingual L(atent) S(ematic) A(nalysis), En Boas, H. (ed). *Multilingual FrameNets in computational lexicography: methods and applications* (pp. 245-287). Berlín: Mouton de Gruyter.
- Lexicom* (Modelo Léxico Construccional). En www.lexicom.es.
- Hanks, P. (2004). The syntagmatics of metaphor and idiom. *International Journal of Lexicography, 17(3)*: 245-274.
- Hanks, P. (2013). *Lexical Analysis: Norms and Exploitations*. Cambridge, MA: MIT Press.
- Hermann, K. M., Dyer, C., Blunsom, P. y Pulman, S. (2012). Learning Semantics and Selectional Preference of Adjective-Noun Pairs. En *SemEval 12 - The First Joint Conference on Lexical and Computational Semantics (SEM 2012)* (pp. 70-74). Obtenido de: <http://ixa2.si.ehu.es/starsem/proc/index.html>.
- Herpio, M. (2002). Benedict: an EU Project for an Intelligent Dictionary. *Kernerman Dictionary News, 10(9)*. Obtenido de: <http://kictionaries.com/newsletter.html>.
- Hill, J. y Lewis, M. (Eds.). (1997). *LTP Dictionary of Selected Collocations*. Londres: English Teaching Publications.
- Hunter, A. y Marten, L. (1999). Context sensitive reasoning with lexical and world knowledge. En G.-J. Scott, E. K.-M. Mui y H.-J. Lee (Eds.), *SOAS Working Papers in Linguistics and Phonetics, 9*, 373-386. Londres.
- Ide, N. y Véronis, J. (1998). Word sense disambiguation: The state of the art. *Computational Linguistics, 24(1)*, 1-40.
- Ide, N. y Wilks, Y. (2006). Making Sense about Sense. En E. Agirre y P. Edmonds (Eds.), *Word Sense Disambiguation: Algorithms and Applications* (pp. 47-74). Springer.
- Instituto Cervantes*. En <http://www.cervantes.es/default.htm>.
- Jiménez, R. y Luzondo, A. (2011). Building ontological meaning in a lexico-conceptual knowledge base. *Onomázein, 23*, 11-40.
- Jiménez, R. y Pérez, B. (2011). An account of selection restrictions in Role and Reference Grammar. *Revista Canaria de Estudios Ingleses, 62*, 99-122.

- Jorgensen, J. (1990). The Psychological Reality of Word Senses. *Journal of Psycholinguistic Research*, 19, 167-190.
- Kilgarriff, A. (1997). I don't believe in word senses. *Computers in the Humanities*, 31(2), 91-113.
- Kilgarriff, A. (2006). Word Senses. En E. Agirre y P. Edmonds (Eds.), *Word Sense Disambiguation: Algorithms and Applications* (pp. 47-74). Springer.
- Kilgarriff, A., Rychly, P., Smrz, P. y Tugwell, D. (2004). The Sketch Engine. En <http://www.sketchengine.co.uk/sketch-engine-eLx04.pdf>.
- Koike, K. (2000). *Colocaciones léxicas en el español actual. Análisis formal y léxico semántico*. Tesis doctoral, Universidad de Alcalá de Henares.
- Koike, K. (2009). Trabajos sobre colocaciones publicados desde el año 2000 (actualizado el 01/05/2009). Obtenido de: <http://www.ner.takushoku-u.ac.jp/~kkoike/Colocaciones2000.pdf>.
- Lapata, M. A. (2001a). Corpus-based Account of Regular Polysemy: The Case of Context-sensitive Adjectives. *Annual Conference of the North American Chapter of the Association for Computational Linguistics, 2001*.
- Lapata, M. A. (2001b). *The Acquisition and Modeling of Lexical Knowledge: A Corpus-based Investigation of Systematic Polysemy* (Tesis doctoral). Universidad de Edinburgo, Escocia.
- Leacock, C., Miller, G. A. y Chodorow, M. (1998). Using corpus statistics and WordNet relations for sense identification. *Computational Linguistics*, 24(1), 1-40.
- Leech, G. N. ([1974] 1981). *Semantics*. Middlessex, Inglaterra: Penguin.
- Lenci, A. (2000). Building an ontology for the lexicon: semantic types and word meaning. En P. A. Jensen, P. R. Skadhauge (Eds.), *Ontology-Based Interpretation of Noun Phrases* (103-120). University of Southern Denmark, Kolding: Department of Business Communication and Information Science.
- Lenci, A., Busa, Ruimy, N., Gola, E., Monachini, M., Calzolari, Zampolli, A. (2000). *SIMPLE Work Package 2-Linguistic Specifications*. Obtenido de: http://www.ilc.cnr.it/AZ_bibliography/Z176.PDF
- Levin, B. (1993). *English Verb Classes and Alternations: A Preliminary Investigation*, Chicago, IL: University of Chicago Press.
- Lingüística Española Actual* (Número especial dedicado a las colocaciones), 23(1). (2001).
- Longman Dictionary of Contemporary English*. (1987). Londres: Longman.

- Luzondo, A. (2011). *English Resultative Constructions in the Lexical Constructional Model: Implications for Constructional Modeling within a Lexical Conceptual Knowledge Base*. Tesis doctoral, Universidad de La Rioja.
- Mahesh, K. (1996). *Ontology development for machine translation: ideology and methodology* (Technical report MCCS-96-292). New Mexico State University: Computing Research Laboratory.
- Mahesh, K. y Nirenburg, S. (1995). Semantic classification for practical natural language processing. En *6th ASIS SIG/CR Classification Research Workshop: An interdisciplinary meeting* (pp. 79-94). Chicago, IL.
- Mairal, R. (2011 [inédito]). Implementación de un lexicón enumerativo robusto: un enfoque ontológico a la polisemia en el marco del PLN.
- Mairal, R. (2013). La arquitectura de una base de conocimiento léxico conceptual: implicaciones lingüísticas. En M. Giammatteo, L. Ferrari y H. Albano (Eds.), *Léxico y Sintaxis. Volumen temático de la serie editada por la Sociedad Argentina de Lingüística* (pp. 165-192). Mendoza, Argentina: Editorial FFyL, UNCuyo.
- Mairal, R. y Faber, P. (2002). Functional Grammar and Lexical Templates. En R. Mairal y M. J. Pérez Quintero (Eds.), *New perspectives on predicate argument structure in Functional Grammar* (pp. 41-98). Berlín y NY: Mouton de Gruyter.
- Mairal, R., L. Guerrero y C. González (Eds.). (2012). *El funcionalismo en la teoría lingüística. La Gramática del Papel y la Referencia. Introducción, avances y aplicaciones*. Madrid: Akal.
- Mairal, R. y Periñán, C. (2009a). The anatomy of the lexicon component within the framework of a conceptual knowledge base. *Revista Española de Lingüística Aplicada*, 22, 217-244.
- Mairal, R. y Periñán, C. (2009b). Role and Reference Grammar and Ontological Engineering. *Volumen Homenaje a Enrique Alcaraz*. Alicante: Universidad de Alicante.
- Mairal, R. y Periñán, C. (2010). Teoría lingüística y representación del conocimiento: una discusión preliminar. En D. García Padrón y M. C. Fumero Pérez (Eds.), *Tendencias en lingüística general y aplicada* (pp. 155-168). Berlin: Peter Lang.
- Mairal, R., C. Periñán y M. B. Pérez. (2012). La representación léxica. Hacia un enfoque ontológico. En R. Mairal, L. Guerrero y C. González (Eds.), *El funcionalismo en la teoría lingüística. La Gramática del Papel y la Referencia. Introducción, avances y aplicaciones* (pp. 85-102). Madrid: Akal.
- Mairal, R. y F. Ruiz de Mendoza. (2008a). Internal and external constraints in meaning construction: the lexicon-grammar continuum. En *Estudios de Filología Inglesa: Homenaje a la Dra. Asunción Alba Pelayo*. Madrid: UNED.
- Mairal, R. y F. Ruiz de Mendoza. (2008b). New challenges for lexical representation within the Lexical-Constructional Model. *Revista Canaria de Estudios Ingleses* 57, 137-158.

- Mairal, R. y F. Ruiz de Mendoza. (2009). Levels of description and explanation in meaning construction. En Butler, C. y Martín Arista, J. (Eds.), *Deconstructing Constructions* (pp. 153–198). Amsterdam y Filadelfia: John Benjamins.
- Mairal, R., Ruiz de Mendoza, F. J. y Periñán, C. (2011). Constructions within a Natural Language Processing Knowledge Base. En H. Boas y F. González-García (Eds.), *Construction Grammar goes Romance*. Amsterdam y Filadelfia: John Benjamins.
- María Moliner. *Diccionario de Uso del Español*. (2002). Madrid: Gredos.
- Martínez, H., Pedersen, B. S. y Bel, N. (2011). Identification of Sense Selection in Regular Polysemy Using Shallow Features. En B. Sandford Pedersen, G. Nespore e I. Skadina (Eds.), *NODALIDA 2011 Conference Proceedings* (pp. 18–25).
- McMillan Dictionary*. En <http://www.macmillandictionary.com>.
- Mel'čuk, I. (1995). The Future of the Lexicon in Linguistic Description: The Explanatory Combinatorial Dictionary. En I. H. LEE (Ed.), *Linguistics in the Morning Calm*, 3. Seúl, Sur Korea.
- Mel'čuk, I. *et al.* *Dictionnaire explicatif et combinatoire du français contemporain, Recherches lexico-sémantiques*, Vol 1 (1984), Vol 2 (1988), Vol 3 (1992), Vol 4 (1999). Montréal: Les Presses de L'Université de Montréal.
- Merriam-Webster Dictionary*. En <http://www.merriam-webster.com>.
- Mihalcea, R. (2007). Knowledge-Based Methods for WSD. En E. Agirre y P. Edmonds (Eds.), *Word Sense Disambiguation: Algorithms and Applications* (pp. 107-132). Springer.
- Mihalcea, R. y Edmonds P. (Eds.). (2004). *Proceedings of Senseval-3: Third International Workshop on the Evaluation of Systems for the Semantic Analysis of Text*. Barcelona, España.
- Miller, G. A. (Ed). (1990). Special Issue, WordNet: An online lexical database. *International Journal of Lexicography*, 3(4).
- Miller, G., Beckwith, R., Fellbaum, C., Gross, D., y Miller, K. (1990). WordNet: An on-line lexical database. *International journal of lexicography*, 3(4), 235-244.
- Morante, R. 2011. Extended Meanings of Verbs: a Proposal of Formalization. Obtenido de: <http://cswww.essex.ac.uk/cluk/morante.PDF>
- MultiSemCor Corpus*. En <http://multisemcor.fbk.eu/index.php>
- Navigli, R. (2009). Word Sense Disambiguation: a Survey. *ACM Computing Surveys*, 41(2), 1-69. ACM Press.

- Nesi, H. (2000). Electronic Dictionaries in Second Language Vocabulary Comprehension and Acquisition: the State of the Art. En U. Heid, Evert, S., Lehmann, E. y Rohrer, C. (Eds.), *Proceedings of the Ninth Euralex International Congress, EURALEX 2000* (pp. 839– 47). Stuttgart, Alemania: Universitat Stuttgart.
- Nesi, H. (2009). Dictionaries in electronic form. En A. P. Cowie (Ed). *The Oxford History of English Lexicography* (pp. 458-478). NY: Oxford University Press.
- Niles, I. y Pease, A. (2001). Origins of the Standard Upper Merged Ontology: a proposal for the IEEE Standard Upper Ontology. En *Working Notes of the IJCAI-2001 Workshop on the IEEE Standard Upper Ontology*. Seattle, WA.
- Nirenburg, S. y Raskin, V. (1996). Ten choices for lexical semantics. *Memoranda in Computer and Cognitive Science* (MCCS-96-304), New Mexico State University: Computing Research Laboratory.
- Nirenburg, S. y Raskin, V. (2001). Ontological semantics, formal ontology, and ambiguity. En *Proceedings of the international conference on Formal Ontology in Information Systems*. NY.
- Nirenburg, S. y Raskin, V. (2004). *Ontological Semantics*. Cambridge, MA: MIT Press.
- Norvig, P. y Lakoff, G. (1987). Taking: a study in lexical network theory. *Proceedings of the 13th meeting of the Berkeley Linguistics Society*, 195-206.
- Noy, N. F. y McGuinness, D. L. (2001). *Ontology development 101: a guide to creating your first ontology* (Technical report KSL-01-05). Stanford University: Stanford Knowledge Systems Laboratory.
- Nunberg, G. (1979). The non-uniqueness of semantic solutions: polysemy. *Linguistics and Philosophy*, 3(2), 143-184.
- Nunberg, G. y Zaenen, A. (1992). Systematic polysemy in lexicology and lexicography. En H. Tommola, K. Varantola, T. Salmi-Tolonen y J. Schopp (Eds.), *Proceedings of Euralex II*, University of Tampere, Tampere, Finland. Obtenido de: <http://people.ischool.berkeley.edu/~nunberg/Euralex.html>
- Ó Séaghdha, D. (2010). Latent variable models of selectional preference. En *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics* (pp. 435– 444). Stroudsburg, PA.
- O'Hara, T., Mahesh, K., Niremburg, S. (1998). Lexical Acquisition with WordNet and the Mikrokosmos Ontology. En *Proceedings of the COLING/ACL Workshop on Usage of WordNet in Natural Language Processing Systems*.
- Ó Séaghdha, D. y Korhonen, A. (2012). Modelling selectional preferences in a lexical hierarchy. En *Proceedings of the First Joint Conference on Lexical and Computational Semantics (SEM 2012)*. Montreal, Canadá.

- Ostler, N. y Atkins, B. (1991). Predictable meaning shift: some linguistic properties of lexical implication rules. En J. Pustejovsky y S. Bergler (Eds.), *Lexical Semantics and Knowledge Representation*. Association for Computational Linguistics.
- Oxford Advanced Learner's Dictionary, 4th edition*. (1989). Oxford: Oxford University Press.
- Pajsz, J. 2009. On the Possibility of Creating Multifunctional Lexicographical Databases. En H. Bergenholtz, S. Nielsen, S. Tarp (Eds.), *Lexicography at a Crossroads. Dictionaries and Encyclopedias today, Lexicographical Tools tomorrow*. Bern [etc.]: Peter Lang.
- Palmer, M. (1998). Are WordNet sense distinctions appropriate for computational lexicons? En *SIGLEX-98, SENSEVAL*. Sussex, Inglaterra.
- Periñán, C. (2012). The situated common-sense knowledge in FunGramKB. *Review of Cognitive Linguistics 10 (1)*, 184-214.
- Periñán, C. y Arcas, F. (2004). Meaning postulates in a lexico-conceptual knowledge base. *Fifteenth International Workshop on Databases and Expert Systems Applications, IEEE* (pp. 38-42). Los Alamitos, CA.
- Periñán, C. y Arcas, F. (2005). Microconceptual-Knowledge Spreading in FunGramKB. *Proceedings on the 9th IASTED International Conference on Artificial Intelligence and Soft Computing* (pp. 239-244). Anaheim-Calgary-Zurich: ACTA Press.
- Periñán, C. y Arcas, F. (2006). Reusing computer-oriented lexica as foreign-language electronic dictionaries. *Anglogermánica Online, 4*, 69-93.
- Periñán, C. y Arcas, F. (2007a). Cognitive modules of an NLP knowledge base for language understanding. *Procesamiento del Lenguaje Natural, 39*, 197-204.
- Periñán, C. y Arcas, F. (2007b). Deep semantics in an NLP knowledge base. *Twelfth Conference of the Spanish Association for Artificial Intelligence* (pp. 279-288). Salamanca, España: Universidad de Salamanca.
- Periñán, C. y Arcas, F. (2008). A cognitive approach to qualities for NLP. *Procesamiento del Lenguaje Natural, 41*, 137-144.
- Periñán, C. y Arcas, F. (2010a). Ontological commitments in FunGramKB. *Procesamiento del Lenguaje Natural, 44*, 27-34.
- Periñán, C. y Arcas, F. (2010b). The architecture of FunGramKB. *Proceedings of the Seventh International Conference on Language Resources and Evaluation, European Language Resources Association (ELRA)* (pp. 2667-2674). Valeta, Malta.
- Periñán, C. y Carrión, Ll. (2011). FunGramKB y el conocimiento cultural. *Anglogermanica Online 2011*. Obtenido de: <http://www.fungramkb.com/resources/papers/fgkb06.pdf>

- Periñán, C. y Mairal, R. (2009) Bringing Role and Reference Grammar to natural language understanding. *Procesamiento del Lenguaje Natural*, 43, 265-273.
- Periñán, C. y Mairal, R. (2010). La gramática de COREL: un lenguaje de representación conceptual. *Onomázein*, 21(1), 11-45.
- Periñán, C. y Mairal, R. (2011). The COHERENT Methodology in FunGramKB. *Onomazein*, 24(2), 13-33.
- Periñán, C. y R. Mairal. (2012). La dimensión computacional de la Gramática del Papel y la Referencia: la estructura lógica conceptual y su aplicación en el procesamiento del lenguaje natural. En R. Mairal, L. Guerrero y C. González (Eds.), *El funcionalismo en la teoría lingüística. La Gramática del Papel y la Referencia. Introducción, avances y aplicaciones* (pp. 333-348). Madrid: Akal.
- Periñán, C. (2012). The situated common-sense knowledge in FunGramKB. *Review of Cognitive Linguistics*, 10(1), 184-214.
- Pesetsky, D. (1982). *Paths and Categories*. Tesis doctoral, Massachusetts Institute of Technology. Cambridge, MA.
- Peters, W. (2004). *Detection and Characterization of Figurative Language Use in Wordnet*. Tesis doctoral, Universidad de Sheffield, Inglaterra.
- Peters, W. (2006). In Search for More Knowledge: Regular Polysemy and Knowledge Acquisition. En P. Sojka, K-S. Choi, C. Fellbaum, P. Vossen (Eds.), *GWC 2006 Proceedings* (pp. 245–250). Masaryk University, República Checa.
- Peters, W. y Peters, I. (2000). Lexicalised Systematic Polysemy in WordNet. *LREC 2000 2nd International Conference on Language Resources & Evaluation*. Atenas, Grecia.
- Peters, W. y Wilks, Y. (2001). Distribution-oriented extension of WordNet's ontological framework. En *Proceedings of Recent Advances in Natural Language Processing*. Tzigov Chark, Bulgaria.
- Peters, W., I. Peters y P. Vossen. (1998) The Reduction of Semantic Ambiguity in Linguistic Resources. *Proceedings of the First International Conference on Language Resources and Evaluation* (pp. 409-416). Granada.
- Pinker, S. (1999). *How the mind works*. London: Penguin Books.
- Procter, P. (1978). *Longman Dictionary of Contemporary English (LDOCE)*. Londres: Longman. En <http://www.ldoceonline.com/dictionary>
- Pruvost, J. 2000. Colloquium report: Des dictionnaires papier aux dictionnaires électroniques. VII^{me} Journée des dictionnaires. *International Journal of Lexicography*, 13(3), 187–93.
- Pustejovsky, J. (1991). The generative lexicon. *Computational Linguistics*, 17.

- Pustejovsky, J. (1993). Type Coercion and Lexical Selection. En J. Pustejovsky (Ed.), *Semantics and the Lexicon*. Holanda: Kluwer Academia Publishers.
- Pustejovsky, J. (1995). *The Generative Lexicon*. Cambridge, MA: The MIT Press.
- Pustejovsky, J. (2008). From concepts to meaning. The role of lexical knowledge. En P. Sterkenburg (Ed.), *Unity and diversity of languages*. Amsterdam: John Benjamins.
- Pustejovsky, J. y Boguraev, B. (1993). Lexical Knowledge Representation and Natural Language Processing. *Artificial Intelligence*, 63, 193-223.
- Pustejovsky, J. y Busa, F. (1995). *Semantic Parameters and Lexical Universals*. Cambridge, MA: MIT Press.
- Ravin, Y. y Leacock, C. (Eds.). (2000). *Polysemy: Theoretical and Computational Approaches* (pp. 91-110). Oxford: Oxford University Press.
- Resnik, P. (1992). A class-based approach to lexical discovery. *Proceedings of the 30th Annual Meeting of the Association for Computational Linguists* (pp. 327-329).
- Resnik, P. (1997). Selectional preference and sense disambiguation. *Proceedings of ACL Workshop on Tagging Text with Lexical Semantics, Why, What, and How?* (pp. 52-57). Washington.
- Reuters Corpus*. En <http://about.reuters.com/researchandstandards/corpus/>
- Roget's International Thesaurus of English Words and Phrases*. (2000). N.Y.: Bartleby. En <http://thesaurus.com/Roget-Alpha-Index.html>
- Ruhl, C. (1989). *On monosemy: A study in linguistic semantics*. Albany, NY: State University of New York Press.
- Ruiz de Mendoza, F. y Mairal, R. (2007). High-level metaphor and metonymy in meaning construction. En G. Radden, K. M. Köpcke, T. Berg y P. Siemund (Eds.), *Aspects of Meaning Construction in Lexicon and Grammar* (pp. 33-49). Amsterdam y Filadelfia: John Benjamins.
- Ruiz de Mendoza, F. y Mairal, R. (2008). Levels of description and constraining factors in meaning construction: an introduction to the Lexical Constructional Model. *Folia Linguistica*, 42(2), 355-400.
- Ruiz de Mendoza, F. y Mairal, R. (2009). Constructing meaning: a brief overview of the Lexical Constructional Model. En Brdar, M. (Ed.) *Converging and diverging tendencies in Cognitive Linguistics*. Amsterdam y Filadelfia: John Benjamins.
- Ruiz de Mendoza, F. y Mairal, R. (2011). Constraints on syntactic alternation: lexical-constructional subsumption in the Lexical Constructional Model. En P. Guerrero (Ed.), *Morphosyntactic Alternations in English: Functional and Cognitive Perspectives* (pp. 62-82). Inglaterra: Equinox Publishing Books.

- Ruppenhofer, J., M. Ellsworth, M. L. Petruck, C. R. Johnson y J. Scheffczyk. (2010). *FrameNet II: Extended Theory and Practice*. Berkeley, California: International Computer Science Institute.
- Sag, I. (1981). Formal semantics and extralinguistic context. En P. Cole (Ed). *Radical Pragmatics*. NY: Academic Press.
- Sánchez. A. (2001). *Gran Diccionario de Uso del Español Actual*. Madrid: Sociedad General Española de Librería, S. A.
- Schütze, H. (2000). En Y. Ravin y C. Leacock (Eds.), *Polysemy: Theoretical and Computational Approaches* (pp. 91-110). Oxford: Oxford University Press.
- Seco, M., Andrés, O. y Ramos, G. (1999). *Diccionario del Español Actual*. Madrid: Aguilar.
- Selva, T. y Chanier, T. (1998). Apport de l'informatique pour l'accès lexical dans les dictionnaires pour apprenants: projet Alexia. En T. Fontenelle, P. Hiligsmann, A. Michiels, A. Moulin y S. Theissen (Eds.), *Papers submitted to the Eighth EURALEX International Congress on Lexicography* (pp. 631-42). Lieja, Bélgica.
- Selva, T., Verlinde, S. y Binon, J. (2002). Le DAFLES, un nouveau dictionnaire électronique pour apprenants du français. En A. Braasch y C. Povlsen (Eds.), *Proceedings of the Tenth EURALEX International Congress* (pp. 773-783). Copenhagen, Dinamarca.
- SemCor*. Princeton University. En <http://www.cse.unt.edu/~rada/downloads.html#semcor>
- Seto, K. (1996). On the cognitive triangle: the relation of metaphor, metonymy and synecdoque. En A. Burkhardt y N. Norrich (Eds), *Tropic Truth*. Berlín/NY: Mouton De Gruyter.
- Slator, B. y Y. Wilks. (1990). Towards semantic structures from dictionary entries. En A. Kunz y U. Schmitz (eds.), *Linguistic Approaches to Artificial Intelligence*. Frankfurt: Peter Lang.
- Sowa, J. (2000). Ontology, Metadata, and Semiotics. *ICCS '00 Proceedings of the Linguistic on Conceptual Structures: Logical Linguistic, and Computational Issues*, 55-81.
- Stern, G. ([1931], 1964). *Meaning and Change of Meaning*. Blomington, IA: University Press.
- Stevenson, M. y Wilks, Y. (2001). The interaction of knowledge sources in word sense disambiguation. *Computational Linguistics*. 27(3): 321-349
- Subirats, C. (2009). Spanish FrameNet: A frame-semantic analysis of the Spanish lexicon. En Boas, H. (Ed). *Multilingual FrameNets in computational lexicography; methods and applications* (pp. 135-164). Berlín: Mouton de Gruyter.
- Subirats, C. and M. R. L. Petruck. (2003). Surprise: Spanish FrameNet! *Workshop on Frame Semantics. International Congress of Linguists*. Praga, República Checa.

- Tarp, S. (2008). The Third Leg of Two-Legged Lexicography. *Hermes Journal of Linguistics and Communication Studies*, 40, 117-131.
- Tarp, S. (2009). Beyond lexicography: New Visions and Challenges in the Information Age. En H. Bergenholtz, S. Nielsen, S. Tarp (Eds.), *Lexicography at a Crossroads. Dictionaries and Encyclopedias today, Lexicographical Tools tomorrow* (pp. 17-32). Bern [etc.]: Peter Lang.
- Taylor, J. R. (1992). How many meanings does a word have? *Stellenbosch Papers in Linguistics*, 25, 133-168.
- The American Heritage of the English Language (4th Edition)*. Boston: Houghton Mifflin. En <http://dictionary1.classic.reference.com/help/ahd4.html>.
- Tomuro, N. (1998). Semi-Automatic Induction of Systematic Polysemy from WordNet. *Proceedings ACL-98 Workshop on the Use of WordNet in NLP*.
- Tulving, E. (1985). How many memory systems are there? *American Psychologist*, 40, 385-398.
- Van Valin, R. (2005). *The syntax-semantics-pragmatics interface: an introduction to Role and Reference Grammar*. Cambridge: Cambridge University Press.
- Van Valin, R. y LaPolla, R. (1997). *Syntax: structure, meaning and function*. Cambridge: Cambridge University Press.
- Van Valin, R. y R. Mairal. (En prensa). Interfacing the Lexicon and an Ontology in a Linking Algorithm. En M. Ángeles Gómez, F. Ruiz de Mendoza y F. González-García (Eds.), *Form and Function in Language: Functional, Cognitive and Applied Perspectives. Essays in Honour of Christopher S. Butler*. Amsterdam: John Benjamins.
- Varantola, K. (2002). Use and Usability of Dictionaries: Common Sense and Context Sensibility? En M. H. Corraed (Ed.), *Natural Language Processing: A Festschrift in Honour of B. T. S. Atkins. EURALEX* (pp. 30– 44). Obtenido de: <http://www.euralex.org/conferences/>.
- Veale, T. (2004). A non-distributional approach to polysemy detection in WordNet. *Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC 2004)*. Lisboa, Portugal.
- Velardi, P. (1991). Acquiring a Semantic Lexicon for Natural Language Processing. En U. Zernik (Ed.), *Lexical Acquisition: Exploiting On-line Resources To Build A Lexicon* (pp. 343-349). Psychology Press.
- Velardi, P. y Paziienza, M. T. (1991). How to Encode Semantic Knowledge: A Method for Meaning Representation and Computer-Aided Acquisition. *Computational Linguistics*, 17(2), 153-170.

- Vendler, Z. (1967). Verbs and times. En Z. Vendler (Ed.), *Linguistics in Philosophy* (pp. 97-121). Ithaca: Cornell University Press.
- VisuWords*. En <http://www.visuwords.com>.
- Visual Thesaurus*. En <http://www.visualthesaurus.com>.
- Vossen, P. (2004). EuroWordNet: a multilingual database of autonomous and language-specific wordnets connected via an Inter-Lingual-Index. Semi-special issue on multilingual databases. *International Journal of Linguistics*, 17/2.
- Vossen, P. y C. Fellbaum. (2009). Universals and Idiosyncracies in Multilingual WordNets. En F. Boas (ed). *Multilingual FrameNets in Computational Lexicography: Methods and Applications*. Berlín: Walter de Gruyter.
- Weaver, W. (1949). Translation. Reproducido en W. N. Locke y D. A. Booth (1955), *Machine Translation of Languages* (pp. 15–23). NY: John Wiley & Sons.
- Webster's New World Dictionary*. (1989). Estados Unidos: Simon and Schuster.
- Weinreich, U. (1964). Webster's third: A critique of its semantics. *International Journal of American Linguistics*, 30, 405-409.
- Wierzbicka, A. (1972). *Semantic Primitives*. Frankfurt: Athenäum.
- Wilensky, R. (1991). Extending the lexicon by exploiting subregularities (U. C. Berkeley Technical Report). Obtenido de: <http://www.eecs.berkeley.edu/Pubs/TechRpts/1991/6376.html>
- Wilks, Y., Slator, B. y Guthrie, L. (1996). *Electric words: dictionaries, computers and meanings*. Cambridge, MA: MIT Press.
- Wing, B. (2002). Cross-Linguistic Discovery of Semantic Regularity. *Proceedings of the First Global WordNet Conference*. Mysore, India. En http://www.globalwordnet.org/gwa/gwa_conf_01.htm
- WordNet*. En <http://wordnet.princeton.edu/>.
- WordNet Domains*. En <http://wndomains.fbk.eu/>.
- WordNik*. En <http://www.wordnik.com/>.
- Yarowsky, D. (1993). One sense per collocation. En *Proceeding of ARPA Human Language Technology Workshop* (pp. 266-271). Princeton, New Jersey.
- Zaenen, A. 2002. Musings about the Impossible Electronic Dictionary. En M. H. Correard (Ed.), *Natural Language Processing: A Festschrift in Honour of B. T. S. Atkins*. EURALEX (pp. 230– 244).

**APÉNDICE 1. Los metaconceptos y la representación semántica de los roles temáticos
(tomado de Perrián y Mairal, 2010)**

| Metaconcepto | Papel | Definición |
|---------------------|--------------|--|
| #COGNITION | [Agent] | Entity that makes another entity undergo a cognitive process. |
| | Theme | Entity that undergoes a cognitive process. |
| | Referent | Entity present in the consciousness of an entity that undergoes a cognitive process. |
| #COMMUNICATION | Theme | Entity that transmits a message. |
| | Referent | Message (i.e. set of propositions) that is transmitted. |
| | Goal | Entity that receives a message. |
| #CONSTITUTION | Theme | Entity that is made up of other entities. |
| | Referent | Entity that is part of another entity. |
| #CREATION | Theme | Entity that creates another entity. |
| | Referent | Entity that is created by another entity. |
| #EMOTION | Agent | Entity that makes another entity feel an emotion. |
| | Theme | Entity that feels an emotion. |
| | [Attribute] | Entity or quality that describes an attribute of an entity when feeling an emotion. |
| #EXISTENCE | Theme | Entity that exists. |
| #IDENTIFICATION | Theme | Entity that is identified by means of another entity. |
| | [Referent] | Entity that serves to define the identity of another entity. |
| | [Attribute] | Quality ascribed to an entity. |
| #INTENTION | Theme | Entity that pursues actively a determinate aim. |
| | Referent | Something which is actively pursued by an entity. |
| #LOCATION | Theme | Entity that stays in a location. |
| | Location | Location where an entity stays. |
| #MATERIAL | Theme | Entity that, volitionally or not, performs an event. |
| | [Referent] | Entity that is directly involved in the event caused by another entity. |
| #MOTION | Agent | Entity that makes another entity move. |
| | Theme | Entity that changes its place or position. |
| | [Location] | Location in which an entity moves. |
| | [Origin] | Location from which an entity moves. |
| | [Goal] | Location to which an entity moves. |
| #PERCEPTION | Theme | Entity that perceives another entity through any of the |

| | | |
|-----------------|----------|---|
| | | senses. |
| | Referent | Entity that is perceived through any of the senses. |
| | | |
| #POSSESSION | Theme | Entity that owns another entity. |
| | Referent | Entity that is owned. |
| | | |
| #TRANSFER | Agent | Entity that transfers another entity to a third entity. |
| | Theme | Entity that is transferred. |
| | Origin | Entity from which another entity is transferred. |
| | Goal | Entity to which another entity is transferred. |
| | | |
| #TRANSFORMATION | Theme | Entity that transforms another entity. |
| | Referent | Entity that is transformed by another entity. |

**APÉNDICE 2. La interpretación semántica de los papeles temáticos de los satélites
(tomado de Perriñán y Mairal, 2010)**

| Papel | Definición |
|--------------|--|
| Beneficiary | Entity different from those of the arguments that derives benefit from the occurrence of the event. |
| Company | Entity that participates in a coordinated way with an entity of the arguments, usually Agent or Theme. |
| Comparison | Quality that is used as the basis of the comparison between two entities, usually Theme and Referent. |
| Condition | Predication that describes under which condition the event should occur. |
| Duration | Entity or quality that denotes the length of time from the beginning of the event to its end. |
| Frequency | Quality that describes how often the event occurs. |
| Instrument | Entity that is used to perform the event. |
| Manner | Entity or quality that describes the way in which the event occurs. |
| Means | Entity that, together with an Instrument, is used to perform the event. |
| Position | Quality that describes the position of Theme with respect to Location, Goal or Origin. |
| Purpose | Predication that describes the aim of the event. |
| Quantity | Entity or quality that describes the amount related to the occurrence of the event. |
| Reason | Predication that describes the cause of the event. |
| Result | Predication or entity that describes the consequence of the occurrence of the event. |
| Scene | Predication or entity that describes the situation in which the event occurs. |
| Speed | Quality that describes how fast the event is performed |