

UNIVERSITAT JAUME I

DEPARTAMENT DE TRADUCCIÓ I COMUNICACIÓ

**LAS RELACIONES CONCEPTUALES EN LA
TERMINOLOGÍA DE LOS PRODUCTOS CERÁMICOS Y SU
FORMALIZACIÓN MEDIANTE UN EDITOR DE
ONTOLOGÍAS**

Tesis doctoral

Presentada por:

María de la Nava Maroto García

Dirigida por:

Dra. M^a Amparo Alcina Caudet

Castelló, 2007

AGRADECIMIENTOS

Llegar hasta aquí ha sido una carrera de obstáculos, y si acaso estoy cerca de la meta ha sido gracias a la colaboración inestimable de un buen número de personas.

En primer lugar quisiera agradecer a la Dra. Amparo Alcina por haber aceptado la dirección de este trabajo. Su implicación en este trabajo va mucho más allá de su excelente dirección académica: he de agradecerle los ánimos, la disponibilidad absoluta y la atención que me ha dedicado a lo largo de estos años.

A los miembros del grupo de investigación TecnoLeTTra de la Universitat Jaume I, en particular a Victoria Soler y Anna Estellés por acogerme y poner a mi disposición todo su trabajo, su tiempo y su amistad.

A D. Ismael Sanz, del Departamento de Ciencia e Ingeniería de los Computadores de la Universidad Jaume I y a D^a María Jesús José Ferrandis, del Instituto Técnico de la Cerámica por ayudarme a resolver las múltiples dudas sobre el programa informático Protégé y el campo de la cerámica, respectivamente.

Estoy en deuda con mis compañeros y alumnos del Centro de Estudios Felipe II de Aranjuez, que me han prestado su comprensión, sus comentarios y los ánimos en todas las etapas del trabajo. Quisiera destacar la ayuda de Susana Cantero, Juan Pedro Rica, Cristina McLaren, Marta Peinado y las chicas de la *plaga*. Mención especial merecen la Dra. Ana Antón-Pacheco y la Dra. Mirella Marotta por la confianza que siempre han depositado en mí.

A Rosa, Asun, Rosi y Boci les agradezco que hayan estado pendientes de mí, sobre todo en los últimos meses, los más duros. A Marta Renau le debo el apoyo logístico en mis visitas a Castellón y por las interminables charlas sobre lo humano y lo divino.

A Janio, a quien cuando nos permitimos bromear llamo la Fundación Villar Sanjuán para el Desarrollo de la Terminología le doy las gracias por tratar de entenderme, por soportar mis altibajos, por los momentos robados y por todos los que sin duda están por venir.

Por último, a mis padres y a Tita, por creer siempre en mí, por preguntar "¿y por lo demás qué tal?", y por perdonar cada uno de los ratos que no he podido compartir con ellos. A Jose y a Virginia por su "servicio 24 horas" y a David y Mercedes por estar tan pendientes. Y por supuesto a mis sobrinos Nava y Julián por ser los únicos que consiguen que me olvide por completo de la tesis.

A todos ellos, y a quienes me haya podido dejar en el tintero, gracias.

Who controls the vocabulary, controls the knowledge

George Orwell

Índice

ÍNDICE.....	7
ÍNDICE DE TABLAS	13
ÍNDICE DE FIGURAS	15
1 INTRODUCCIÓN.....	19
1.1 <i>Hipótesis inicial</i>	19
1.2 <i>Objetivos</i>	23
1.3 <i>Metodología del estudio empírico</i>	25
1.4 <i>Estructura de la tesis</i>	26
PRIMERA PARTE: FUNDAMENTOS TEÓRICOS.....	29
2 LAS RELACIONES CONCEPTUALES ENTRE TÉRMINOS	33
2.1 <i>Teoría General de la Terminología</i>	34
2.1.1 Naturaleza del concepto	36
2.1.2 Características	38
2.1.2.1 Intensión y extensión	38
2.1.2.2 Clasificaciones de las características.....	39
2.1.3 Combinaciones de conceptos	43
2.1.4 Relaciones conceptuales: definición y tipología.....	44
2.1.4.1 Relaciones jerárquicas	45
2.1.4.2 Relaciones no jerárquicas.....	49
2.1.5 Profundización en la naturaleza de las relaciones conceptuales desde la TGT....	53
2.1.5.1 Weissenhofer (1995).....	54
2.1.5.2 Nuopponen (1994 y 2005)	56
2.2 <i>Nuevos enfoques de la terminología</i>	62
2.2.1 Enfoque funcionalista y dinámico de la terminología	62
2.2.1.1 Clases de conceptos: modelo conceptual estático y dinámico.....	63
2.2.1.2 Relaciones conceptuales en el enfoque funcionalista.....	71
2.2.1.3 La dinámica de la terminología (Kageura, 2002).....	74
2.2.2 Teoría Comunicativa de la Terminología.....	78
2.2.2.1 Catálogo de relaciones conceptuales de Feliu (2000 y 2004).....	82
2.2.3 Enfoque sociocognitivo de la terminología (Temmerman, 2000)	87
2.3 <i>Propiedades matemáticas de las relaciones conceptuales</i>	92
2.3.1.1 Transitividad	93
2.3.1.2 Simetría.....	94
2.3.1.3 Reflexividad.....	94
2.3.1.4 Conversividad, reciprocidad y relaciones inversas	95
2.3.1.5 Relación uno-a-uno.....	96
2.3.1.6 Recursividad de las relaciones	96
2.3.1.7 Principio de la herencia.....	96
2.3.1.8 Composición de relaciones (Otman, 1996: 60-61).....	100
2.4 <i>Las relaciones conceptuales en la Semántica Léxica</i>	100

2.4.1	Hiponimia.....	102
2.4.2	Meronimia.....	106
2.4.3	Sinonimia.....	115
2.4.4	Relaciones causales.....	117
2.5	<i>Recapitulación y conclusiones</i>	120
3	FORMALIZACIÓN DE LAS RELACIONES CONCEPTUALES EN TERMINOLOGÍA Y LINGÜÍSTICA.....	125
3.1	<i>Representación de las relaciones conceptuales en la TGT: sistemas de conceptos</i>	129
3.2	<i>Nuevos modelos para la formalización de relaciones en Terminología</i>	134
3.2.1	Bases de conocimiento terminológico: proyecto COGNITERM.....	135
3.2.1.1	Gestión de la multidimensionalidad en COGNITERM.....	137
3.2.2	Las redes semántico-terminológicas (Otman, 1996).....	139
3.2.3	Formalización mediante grafos conceptuales e hipertexto (Antia, 2000).....	145
3.2.4	Representación de las relaciones semánticas mediante esquemas relacionales (Oster, 2005).....	148
3.2.5	Aplicación de las funciones léxicas a la terminografía basada en corpus.....	151
3.3	<i>Formalización de las relaciones conceptuales en terminología por medio de ontologías</i>	159
3.3.1	ONCOTERM (Faber y Jiménez, 2002).....	160
3.3.2	El sistema informático CAOS (Madsen <i>et al</i> , 2004).....	168
3.3.2.1	Gestión de la multidimensionalidad en el sistema CAOS.....	171
3.3.3	Integración del enfoque sociocognitivo y las ontologías: Termontography (Kerremans y Temmerman, 2003).....	173
3.3.4	Proyecto GENOMA-KB.....	176
3.3.5	Ventajas e inconvenientes del empleo de ontologías en el ámbito especializado ...	178
3.4	<i>Ontologías lingüísticas y bases de datos léxicas</i>	180
3.4.1	WordNet.....	181
3.4.2	La ontología Mikrokosmos y el paradigma de la semántica ontológica.....	194
3.4.2.1	Descripción de la ontología Mikrokosmos.....	197
3.4.3	La base de datos léxica FrameNet.....	203
3.5	<i>Recapitulación y conclusiones</i>	207
4	REPRESENTACIÓN DEL CONOCIMIENTO EN LA INGENIERÍA ONTOLÓGICA.....	213
4.1	<i>Definición de ontología</i>	213
4.1.1	Técnicas de modelado y componentes básicos de las ontologías.....	215
4.1.1.1	Marcos y lógica de primer orden.....	216
4.1.1.2	Lógica descriptiva (Description Logics, DL).....	217
4.2	<i>Principios para el diseño de ontologías</i>	219
4.2.1	Metodologías y métodos para el desarrollo de ontologías.....	221
4.2.1.1	METHONTOLOGY.....	223
4.2.2	Lenguajes para la creación de ontologías.....	233
4.3	<i>Tipos de ontologías y ontologías más representativas</i>	237

4.3.1	Tipología de las ontologías	237
4.3.1.1	Ontologías según la riqueza de la estructura interna	238
4.3.1.2	Ontologías según la temática de la conceptualización	239
4.3.2	Descripción de algunas ontologías representativas	241
4.3.2.1	Ontologías de representación del conocimiento (metaontologías)	241
4.3.2.2	Ontologías de nivel superior	247
4.3.2.3	Ontologías lingüísticas	251
4.3.2.4	Ontologías de dominio	253
4.4	<i>Herramientas para el desarrollo de ontologías</i>	255
4.4.1	WebODE	257
4.4.2	OntoEdit	259
4.4.3	KAON	261
4.4.4	Protégé	263
4.4.4.1	Protégé-frames	264
4.4.4.2	Protégé-OWL	267
4.5	<i>Recapitulación y conclusiones</i>	274
2ª PARTE: ESTUDIO EMPÍRICO		277
5	DESCRIPCIÓN DEL TRABAJO EMPÍRICO	283
5.1	<i>Fases del trabajo empírico</i>	284
5.2	<i>Recursos textuales y terminológicos empleados en el trabajo empírico</i>	287
5.2.1	<i>Guía de la baldosa cerámica</i> (2003)	288
5.2.2	<i>Diccionario cerámico científico-práctico</i> (1987)	289
5.2.3	<i>Dictionary of ceramic science and engineering</i> (1994)	289
5.2.4	Base de datos terminológica <i>Cerámica</i>	290
5.2.5	Árbol de campo de la cerámica	291
5.2.6	Corpus textual <i>TXT CERAM</i>	292
5.3	<i>Descripción de las herramientas informáticas empleadas</i>	294
5.3.1	Herramienta de análisis de corpus: <i>WordSmith Tools</i>	295
5.3.2	Herramienta de gestión de bases de datos terminológicas: <i>Multiterm</i>	297
5.3.3	Editor de ontologías: <i>Protégé</i>	299
5.4	<i>Recapitulación</i>	309
6	FORMALIZACIÓN E IMPLEMENTACIÓN DE LAS RELACIONES CONCEPTUALES EN UN EDITOR DE ONTOLOGÍAS	313
6.1	<i>Catálogo de relaciones conceptuales</i>	313
6.1.1	Convenciones para la denominación de las relaciones conceptuales	314
6.1.2	Estructuración del catálogo de relaciones conceptuales	315
6.2	<i>Formalización de las relaciones conceptuales</i>	321
6.2.1	Definición formal y propiedades de las relaciones conceptuales	322
6.2.2	Clases de conceptos	325
6.2.3	Análisis de las relaciones conceptuales en el ámbito de la cerámica: organización de la información conceptual extraída	326
6.3	<i>Implementación en el editor de ontologías</i>	333

6.3.1	Implementación de los conceptos, denominaciones y relaciones conceptuales en <i>Protégé</i>	333
6.3.1.1	Implementación de los conceptos.....	334
6.3.1.2	Implementación de las denominaciones	334
6.3.1.3	Implementación de las relaciones conceptuales	335
6.3.2	Tabla representativa de la implementación de las relaciones.....	343
6.4	<i>Definición, formalización e implementación de las relaciones conceptuales del catálogo</i>	344
6.4.1	Relaciones lógicas	344
6.4.1.1	Relación superordinado-subordinado	345
6.4.1.2	Relación subordinado-subordinado	346
6.4.2	Relaciones meronímicas	347
6.4.2.1	Relación componente funcional-objeto	347
6.4.2.2	Relación miembro-colección.....	349
6.4.2.3	Relación porción-masa.....	350
6.4.2.4	Relación material-objeto	351
6.4.2.5	Relación etapa-proceso.....	352
6.4.2.6	Relación característica-actividad.....	353
6.4.2.7	Relación lugar-área	354
6.4.2.8	Relación parte-parte	355
6.4.3	Relaciones secuenciales.....	356
6.4.3.1	Relación concepto-lugar.....	357
6.4.3.2	Relación concepto-lugar al que se dirige.....	357
6.4.3.3	Relación concepto-concepto simultáneo	358
6.4.3.4	Relación concepto anterior-concepto posterior	359
6.4.4	Relaciones argumentales y circunstanciales	360
6.4.4.1	Relación proceso-agente.....	360
6.4.4.2	Relación proceso-producto.....	361
6.4.4.3	Relación proceso-paciente.....	362
6.4.4.4	Relación proceso-instrumento	363
6.4.4.5	Relación proceso-estado	364
6.4.4.6	Relación proceso-método	365
6.4.4.7	Relación causa-efecto.....	366
6.4.4.8	Relación objeto-uso.....	367
6.4.5	Otras relaciones	368
6.4.5.1	Relación fenómeno-medida.....	368
6.4.5.2	Relación objeto-característica	369
6.4.5.3	Relación asociativa.....	370
6.4.5.4	Relación de sinonimia	371
6.5	<i>Recapitulación y cuadro resumen del catálogo de relaciones conceptuales</i>	372
7	INTRODUCCIÓN DE DATOS EN EL EDITOR DE ONTOLOGÍAS Y RECUPERACIÓN DE LA INFORMACIÓN.....	385
7.1	<i>Proceso de introducción de datos en Ontoceram</i>	386

7.1.1	Protocolo para la identificación de relaciones conceptuales e introducción de datos en <i>Ontoceram</i>	397
7.2	<i>Recuperación de la información sobre las relaciones conceptuales en Ontoceram</i>	411
7.2.1	Consultas sobre relaciones conceptuales en <i>Ontoceram</i>	412
7.2.2	Realización de consultas en el editor de ontologías <i>Protégé</i>	418
7.3	<i>Propuestas para la adaptación del editor de ontologías al trabajo terminológico automatizado</i>	424
7.4	<i>Recapitulación</i>	425
8	RECAPITULACIÓN Y CONCLUSIONES	429
8.1	<i>Recapitulación de la revisión teórica</i>	430
8.2	<i>Recapitulación y aportaciones del trabajo empírico</i>	433
8.3	<i>Conclusiones y perspectivas de investigación futura</i>	439
9	BIBLIOGRAFÍA	445
	ANEXOS	469
	<i>Anexo 1: Referencias bibliográficas del corpus TXTCERAM y de los recursos textuales sobre la industria cerámica</i>	469
	<i>Anexo 2: Referencias sobre las herramientas informáticas empleadas</i>	472

Índice de tablas

Tabla 1. Relaciones morfoconceptuales propuestas por Weissenhofer (1995).	56
Tabla 2. Relaciones conceptuales lógicas en Nuopponen (2005).....	59
Tabla 3. Relaciones conceptuales ontológicas en Nuopponen (2005).....	61
Tabla 4. Principios y condiciones de la Teoría Comunicativa de la Terminología.	80
Tabla 5. Tipología de relaciones conceptuales (Feliu y Cabré, 2002).....	86
Tabla 6. Principios de la Teoría Sociocognitiva de la Terminología (Temmerman, 2000: 223).....	88
Tabla 7. Tipos de relaciones meronímicas y elementos de la relación (Winston <i>et al</i> , 1987).....	113
Tabla 8. Notaciones utilizadas para representar las relaciones semánticas entre términos en los tesauros (Currás, 2005).	128
Tabla 9. Funciones léxicas estándar que pueden aplicarse en terminología (Mel'cuk y Zholkovsky, 1988).....	153
Tabla 10. Relaciones intralingüísticas entre <i>synsets</i> en EuroWordNet (Vossen, 1999: 17).	190
Tabla 11. Glosario de términos para la ontología de viajes (Gómez Pérez <i>et al</i> (2004: 133).	225
Tabla 12. El concepto <i>Travel</i> en el diccionario de conceptos de la ontología de viajes (Gómez Pérez <i>et al</i> , 2004: 136).	228
Tabla 13. Extracto de la tabla de relaciones binarias <i>ad hoc</i> para la ontología de viajes (Gómez Pérez <i>et al</i> , 2004: 137).	229
Tabla 14. Extracto de la tabla de atributos de las instancias para la ontología de viajes (Gómez Pérez <i>et al</i> , 2004: 138).	230
Tabla 15. Extracto de la tabla de atributos de clase para la ontología de viajes (Gómez Pérez <i>et al</i> 2004: 138).....	230
Tabla 16. Extracto de la tabla de constantes para la ontología de viajes (Gómez Pérez <i>et al</i> 2004: 139).....	230
Tabla 17. Extracto de la tabla de atributos de las instancias para la ontología de viajes (Gómez Pérez <i>et al</i> 2004: 141).	232
Tabla 18. Descripción del corpus <i>TxtCeram</i> según los parámetros propuestos por Bowker y Pearson (2002).	293
Tabla 19. Conceptos relacionados con el concepto <i>baldosa cerámica</i> recogidos en el estudio preliminar.	331
Tabla 20. Cuadro resumen de la implementación de los <i>slots</i> de relación en <i>Ontoceram</i>	344
Tabla 21. Implementación de la relación subordinado-subordinado.	347
Tabla 22. Implementación de la relación componente funcional-objeto.	348
Tabla 23. Implementación de la relación miembro-colección.....	349
Tabla 24. Implementación de la relación porción-masa.	350
Tabla 25. Implementación de la relación material-objeto.	351
Tabla 26. Implementación de la relación etapa-proceso.	353
Tabla 27. Implementación de la relación característica-actividad.....	354
Tabla 28. Implementación de la relación lugar-área.	355

Tabla 29. Implementación de la relación parte-parte.	356
Tabla 30. Implementación de la relación concepto-lugar.	357
Tabla 31. Implementación de la relación concepto-lugar al que se dirige.	358
Tabla 32. Implementación de la relación concepto-concepto simultáneo.	359
Tabla 33. Implementación de la relación concepto anterior-concepto posterior.	360
Tabla 34. Implementación de la relación proceso-agente.	361
Tabla 35. Implementación de la relación proceso-producto.	362
Tabla 36. Implementación de la relación proceso-paciente.	363
Tabla 37. Implementación de la relación proceso-instrumento.	364
Tabla 38. Implementación de la relación proceso-estado.	365
Tabla 39. Implementación de la relación proceso-método.	366
Tabla 40. Implementación de la relación causa-efecto.	367
Tabla 41. Implementación de la relación objeto-uso.	368
Tabla 42. Implementación de la relación fenómeno-medida.	369
Tabla 43. Implementación de la relación objeto-característica.	370
Tabla 44. Implementación de la relación asociativa.	371
Tabla 45. Tabla resumen del catálogo de relaciones conceptuales.	382
Tabla 46. Descripción cuantitativa de la base de conocimiento <i>Ontoceram</i>	395
Tabla 47. Tabla que recoge los conceptos relacionados con <i>Pavimento de gres</i>	401
Tabla 48. Tabla completa de los conceptos relacionados con <i>pavimento de gres</i>	405
Tabla 49. Información sobre las relaciones implementadas en la base de conocimiento.	413

Índice de figuras

Fig. 1. Plantilla para la definición de las unidades de comprensión (Temmerman, 2000: 122).	90
Fig. 2. Jerarquía integrada de las relaciones parte-todo (Barrière, 2002: 95).	114
Fig. 3. Tipología de las relaciones causales propuesta por Barrière (2002: 98).	119
Fig. 4. Ejemplo del descriptor <i>ayuda humanitaria</i> procedente del tesoro Eurovoc.	129
Fig. 5. Representación gráfica del sistema conceptual de <i>instrument d'écriture</i> (ISO 704, 2000: 15)	133
Fig. 6. Fragmento de la red semántico-terminológica del concepto <i>règle</i> (Otman, 1997: 248). ...	144
Fig. 7. Grafo conceptual complejo para el concepto <i>joint resolution</i> (Antia: 2000, 169).	147
Fig. 8. Marco conceptual de referencia: representación del evento médico (Faber, 2002: 7).	162
Fig. 9. El concepto <i>abdominal-cancer</i> en la base de datos del proyecto ONCOTERM.	166
Fig. 10. Sistema de conceptos en CAOS 2 (Madsen <i>et al</i> , 2005).	172
Fig. 11. Estructura general de la base de conocimiento GENOMA-KB (Cabré <i>et al</i> , 2004).	177
Fig. 12. Los once <i>unique beginners</i> propuestos en WordNet (Miller: 1998, 30).	183
Fig. 13. Estructura superior de los conceptos en la ontología Mikrokosmos (Mahesh y Nirenburg, 1995).	201
Fig. 14. Representación gráfica de la relación <i>arrivalPlace</i> y su inversa <i>isArrivalPlaceOf</i> (Gómez Pérez <i>et al</i> , 2004: 135).	228
Fig. 15. Relación entre universales y particulares en la concepción de Guarino y Welty (2000).	248
Fig. 16. Ontología de nivel superior de Sowa (2000).	249
Fig. 17. Estructura modular de la ontología SUMO.	251
Fig. 18. Niveles superiores de la ontología GUM.	252
Fig. 19. Estructura del concepto <i>Columnist</i> en la ontología de muestra <i>Newspaper</i>	266
Fig. 20. Captura de pantalla de la clase <i>CheezyPizza</i> en la ontología <i>pizza.owl</i>	270
Fig. 21. Jerarquía de propiedades en el editor <i>Protégé-OWL</i>	271
Fig. 22. Concordancias del término <i>baldoa cerámica</i> obtenidas con el programa <i>Concord</i>	296
Fig. 23. Ficha terminológica del término <i>baldoa cerámica</i> en <i>Cerámica</i>	298
Fig. 24. La clase <i>Columnist</i> de la ontología <i>Newspaper</i>	301
Fig. 25. Ejemplo del <i>slot containing_section</i> en la ontología <i>Newspaper</i>	303
Fig. 26. <i>Instances</i> para la clase <i>Columnist</i> en la ontología <i>Newspaper</i>	305
Fig. 27. Resultado de la consulta <i>Articles whose authors are highly paid and articles are urgent with less than 50 pages</i> en la ontología <i>Newspaper</i>	307
Fig. 28. Representación gráfica de la clase <i>Columnist</i> utilizando el <i>plug-in</i> TGVizTab.	309
Fig. 29. Esquema de la relación conceptual entre <i>baldoa cerámica</i> y <i>baldoa extrudida</i>	329
Fig. 30. Nueva faceta <i>propiedades matemáticas</i> añadida para todos los <i>slots</i> de relación.	338
Fig. 31. Implementación de la relación <i>proceso-producto</i> en <i>Ontoceram</i>	339
Fig. 32. <i>Slot propiedades meronímicas</i> , añadido a las relaciones creadas con la metaclassa :MERONYMIC-RELATION-SLOT.	340

Fig. 33. Relación meronímica <i>componente funcional-objeto</i> en <i>Ontoceram</i>	341
Fig. 34. Sinónimos empleados para referirse al concepto <i>Soporte</i> recogidos en <i>Ontoceram</i>	342
Fig. 35. Estructura de la ontología Mikrokosmos implementada en <i>Ontoceram</i>	389
Fig. 36. Representación de la clase <i>Baldosa cerámica</i> en <i>Ontoceram</i>	393
Fig. 37. Instancias para la clase <i>Baldosa cerámica</i>	394
Fig. 38. Ficha terminológica de <i>pavimento de gres</i> en la base de datos <i>Cerámica</i>	400
Fig. 39. Menú inicial de la herramienta <i>Concord</i>	402
Fig. 40. Menú de selección de los textos del corpus <i>TxtCeram</i>	403
Fig. 41. Ventana de selección de la palabra o frase de búsqueda en <i>Concord</i>	403
Fig. 42. Concordancias de <i>Pavimento* cerámico</i> en el corpus <i>TxtCeram</i>	404
Fig. 43. Ventana <i>Open Project</i> del editor de ontologías <i>Protégé</i>	405
Fig. 44. Apariencia de la base de datos <i>Ontoceram</i> al abrirla en el editor de ontologías.	406
Fig. 45. Marco de la <i>Class</i> que representa al concepto <i>Pavimento de gres</i>	407
Fig. 46. Ventana del <i>slot</i> de relación <i>paciente-proceso</i> en la clase <i>Pavimento de gres</i>	408
Fig. 47. Instancias que representan las denominaciones del concepto <i>Pavimento de gres</i>	409
Fig. 48. Lista completa de las denominaciones empleadas para referirse al concepto <i>Pavimento de gres</i>	410
Fig. 49. Resultado de la consulta <i>piezas cerámicas resistentes a la helada</i>	419
Fig. 50. Resultado de la búsqueda <i>baldosas cerámicas prensadas en seco</i>	420
Fig. 51. Consulta en PAL <i>producto-acabado-baldosa-klinker</i>	422
Fig. 52. Pestaña de consultas en lenguaje axiomático PAL.....	423

1. INTRODUCCIÓN

1 Introducción

1.1 Hipótesis inicial

Desde sus inicios, los estudiosos de la terminología han hecho hincapié en la naturaleza sistemática de los términos. Como afirma Cabré (1999: 82), dicha sistematicidad se hace palpable tanto en un plano general como en un plano específico. Desde el punto de vista general, los términos son sistemáticos en relación al sistema de la lengua, del que forman parte. Desde el punto de vista específico, los términos de una disciplina mantienen relaciones conceptuales entre sí, conformando un sistema en el que no se encuentran aislados, sino que dependen unos de otros.

La terminología posee una doble función: de una parte, organiza el conocimiento, es decir, participa en su estructuración mental, y de otra, posibilita la comunicación especializada (Cabré, 1999: 80).

Esto nos lleva a pensar que los términos de una disciplina desempeñan un papel estructurador del conocimiento, de modo que, si somos capaces de hacer explícitas las relaciones que se establecen entre ellos, estaremos en disposición de representar y formalizar el conocimiento de dicha área de especialidad.

La terminografía, vertiente aplicada de la terminología, se ha dedicado tradicionalmente a buscar modelos para hacer explícitas las relaciones existentes entre conceptos a través de los sistemas conceptuales y sus representaciones gráficas. Sin embargo, las representaciones tradicionales resultan insuficientes, puesto que apenas tienen en cuenta la multidimensionalidad de los sistemas de conceptos, es decir, los distintos puntos de vista desde los que éstos pueden observarse.

Otra insuficiencia de los sistemas de conceptos tradicionales radica en su falta de formalización. Las normas internacionales sobre terminología dan pautas sobre cómo estructurar las relaciones genéricas y partitivas, pero adolecen de falta de directrices para representar claramente otros tipos de relación, como las relaciones causa-efecto o las argumentales o asociativas. Los sistemas conceptuales tradicionales tampoco permiten reflejar la complejidad de algunas relaciones, como las partitivas. Además, en la mayoría de los casos las relaciones

se hacen explícitas a través de definiciones expresadas en lenguaje natural, lo que dificulta el empleo de sistemas informáticos para la gestión del conocimiento.

Por ello, desde hace algunos años se han venido realizando esfuerzos para representar la sistematicidad de los términos teniendo en cuenta la multidimensionalidad y los distintos tipos de relaciones, tales como la propuesta de COGNITERM (Meyer *et al*, 1997) o la propuesta de redes semántico-terminológicas de Otman (1996).

Más recientemente, la terminología ha buscado respaldo en otras disciplinas, como la ingeniería del conocimiento, que trata de estructurar la realidad especializada para utilizar las representaciones del conocimiento en aplicaciones diversas. Así, proyectos como ONCOTERM (Faber y Jiménez, 2002) o la propuesta termontográfica de Temmerman y Kerremans (2003), entre otros, han empezado a considerar el empleo de las ontologías para la representación de conceptos.

Temmerman y Kerremans (2003) identifican tres grandes cambios que están revolucionando tanto la concepción de la unidad terminológica como los métodos de análisis y almacenamiento de la información sobre los términos.

El primer cambio hace referencia a la **gestión informatizada de la terminología**. Este cambio se ha visto propiciado por dos aspectos fundamentales:

1. El desarrollo de herramientas que permiten almacenar y analizar grandes corpus de textos especializados, a partir de los cuales podemos extraer información relevante.
2. La aparición de gran cantidad de programas que posibilitan la gestión de la terminología mediante bases de datos, fácilmente accesibles para los terminógrafos.

El segundo cambio, de índole más teórica, está relacionado con la concepción de la teoría de la terminología desde una **perspectiva lingüística**, que tiene en cuenta aspectos de la comunicación que habían permanecido relegados a un segundo plano en la concepción clásica de esta disciplina. Este cambio tiene su origen en aportaciones como la realizada por Sager (1990) a principios de la década de los noventa, y ha alcanzado mayor desarrollo en nuevos enfoques como la teoría comunicativa de la terminología (Cabré, 1999) o el enfoque

sociocognitivo (Temmerman, 2000), que reivindican la naturaleza lingüística del término.

El tercer cambio, y el más reciente, es el producido hacia la organización de la información terminológica mediante **ontologías**, como evolución de las bases de conocimiento terminológico (*terminological knowledge bases*) concebidas por Meyer *et al* (1992). Desde esta nueva perspectiva, se reafirma la importancia de las relaciones semánticas entre conceptos.

La presente investigación se ubica en este panorama, y pretende servir de él y contribuir al desarrollo de una nueva terminología cuya labor se centra en hacer explícita la estructura conceptual interna de los ámbitos especializados, y que recurre a las nuevas herramientas informáticas desarrolladas para el almacenamiento y recuperación del conocimiento.

Esta tesis doctoral se sitúa en el marco de dos proyectos de investigación desarrollados bajo la dirección de la Dra. Amparo Alcina en el Departamento de Traducción y Comunicación de la Universidad Jaume I de Castellón. Tales proyectos son “TXTCERAM. Extracción semiautomática y análisis conceptual formal de términos de la cerámica a partir de un corpus electrónico. Su eficacia y utilidad en la mediación lingüística”, financiado por la Generalitat Valenciana (código del proyecto: GV05/260), y “ONTODIC. Metodología y tecnologías para la elaboración de diccionarios onomasiológicos basados en ontologías. Recursos terminológicos para la e-traducción”, financiado por el Ministerio de Ciencia y Tecnología (código del proyecto: TSI2006-01911). El proyecto TXTCERAM tiene como objetivo la creación de un corpus electrónico de textos de especialidad del ámbito de la cerámica en el cual se pueda probar la eficacia de algunas herramientas informáticas para el diseño de un sistema integral de terminología asistida orientado a la elaboración y consulta de terminologías. Por su parte, el proyecto ONTODIC tiene como objetivo proponer una metodología sistemática para la elaboración de diccionarios terminológicos onomasiológicos empleando, entre otras herramientas, un editor de ontologías (Alcina, en prensa). Estos dos proyectos son herederos de un tercero, denominado “Descripción terminológica de una rama profesional de la industria cerámica. Elaboración de un diccionario multilingüe”, financiado por Caixa de Castelló/Bancaja en el período 1998-2000 y por la Consellería de Educación y Ciencia de la Generalitat Valenciana entre 2001

y 2003 bajo la dirección de la Dra. Pilar Civera (código del proyecto: GV00-143-9). La elección del ámbito temático de la industria cerámica se justifica por la importancia de este sector económico en la provincia de Castellón.

El punto de partida del presente trabajo es la observación de que, por regla general, en las terminologías existe una escasa formalización de las relaciones que los conceptos establecen entre sí. El modo tradicional de hacer explícita la información conceptual es la definición en lenguaje natural. En ella, la información se encuentra dispersa, lo cual dificulta su recuperación, puesto que la falta de formalización limita el acceso a ella por medio de sistemas informáticos. A pesar de que el enfoque del trabajo terminológico se suele considerar onomasiológico, es decir, que parte del concepto para llegar a la denominación, el proceso de recuperación de la información almacenada en las bases de datos terminológicas convencionales se realiza a partir de la denominación, o dicho de otro modo, de forma semasiológica.

Para consultar las obras terminográficas de un modo verdaderamente onomasiológico (Sierra y McNaught, 2000) es necesario almacenar la información conceptual de una manera altamente estructurada. Esto se puede conseguir empleando herramientas que hagan explícita la organización conceptual de las áreas de especialidad. Creemos que la formalización del conocimiento con ayuda de un editor de ontologías puede servir de apoyo en las herramientas terminológicas. Una ontología se define como un conjunto de conceptos organizados de manera jerárquica, representados en un sistema informático, que puede servir de soporte para diversas aplicaciones (Moreno Ortiz, 2002: 31).

Hasta ahora, los proyectos que desde la terminología han aprovechado las ontologías para la estructuración conceptual han diseñado sus propias herramientas de gestión, teniendo en cuenta las especificidades del trabajo terminológico. Este es el caso, por ejemplo, del proyecto ONCOTERM (Faber y Jiménez, 2002), para el que se desarrolló el editor de ontologías ONTOTERM®, orientado específicamente al trabajo terminográfico. Sin embargo, en el presente trabajo de investigación nos planteamos como principal objetivo aprovechar los editores de ontologías que la ingeniería del conocimiento pone a nuestro alcance para la formalización de las relaciones conceptuales entre términos. Así, en lugar de desarrollar una nueva herramienta, exploraremos las posibilidades de uno de los editores de ontologías existentes (en nuestro caso el editor *Protégé*) para la

estructuración de la terminología de un dominio. En caso necesario, propondremos adaptaciones de esta herramienta, ampliamente utilizada en la ingeniería del conocimiento, para el trabajo terminológico.

1.2 Objetivos

Junto al objetivo general de este trabajo que acabamos de apuntar, es decir, la aplicación de los editores de ontologías para la formalización de las relaciones entre los conceptos de un área de especialidad, en nuestro caso la industria cerámica, nos planteamos los siguientes objetivos específicos:

1. En primer lugar, se desarrollará un catálogo de relaciones conceptuales completo aplicable a cualquier ámbito temático y no sólo al dominio concreto de la industria cerámica.
2. A partir de este catálogo bien definido, se analizará qué tipo de información sobre las relaciones conceptuales habrá de formalizarse y se propondrá un modo de llevar a cabo esta formalización.
3. Una vez formalizada la información sobre las relaciones conceptuales de nuestro catálogo, afrontaremos su implementación en el editor de ontologías *Protégé*, teniendo en cuenta las características específicas de esta herramienta informática.
4. Con el fin de probar la utilidad de nuestra propuesta de implementación, la información sobre los conceptos y relaciones conceptuales correspondientes a la rama del producto cerámico acabado se introducirá en una base de datos conceptual creada con el editor de ontologías *Protégé*.
5. Por último, llevaremos a cabo una primera valoración de la utilidad de *Protégé* como herramienta para la consulta onomasiológica de los conceptos y relaciones conceptuales sobre los productos cerámicos acabados, lo cual nos permitirá proponer adaptaciones del editor de ontologías para el trabajo terminológico.

El destinatario final de nuestra base de datos conceptual sobre los productos cerámicos acabados es un usuario potencial que quiere obtener la información contenida en la ontología a partir de los elementos de su significado, tales como las características y relaciones conceptuales. Este usuario podría ser un traductor, un terminólogo, un redactor de textos sobre cerámica o incluso un experto en la materia, por lo que las necesidades a las que la base de conocimiento pretende dar respuesta serán necesariamente variadas.

Otman (1992, citado en Otman, 1996: 98-99) ofrece una tipología de las posibles informaciones léxicas y terminológicas que puede necesitar el usuario experto. Distingue entre varias posibilidades, dependiendo del número de conceptos que se consideren en la consulta.

El tipo de consultas que podría plantear el usuario de nuestra ontología terminológica podrían ser similares a las siguientes:

- ¿Con qué conceptos está relacionado un concepto A? ¿Cuál es la naturaleza de dicha relación?
- ¿Qué características comparten el concepto A y al concepto B?
- ¿Qué conceptos son subordinados de un concepto determinado?

Para dar respuesta a estas consultas, es necesario que la información sobre las relaciones conceptuales esté formalizada de manera explícita. Por ejemplo, si queremos responder a la primera pregunta que acabamos de plantear, “¿con qué conceptos está relacionado un concepto A y cuál es la naturaleza de dicha relación?”, la información almacenada en el editor de ontologías debe estar formalizada de forma que se indique el tipo de relación que vincula a estos conceptos y sus características para que sea recuperable. Pongamos un ejemplo más concreto: si queremos consultar los conceptos relacionados con el producto cerámico acabado *baldosa cerámica* mediante la relación *objeto-componente funcional*, en la ontología debemos representar este concepto y sus conceptos relacionados, y hemos de establecer de forma explícita una relación que vincule a *baldosa cerámica* con sus componentes (*soporte, esmalte, arista*, entre otros) y que sea recuperable en la base de datos conceptual. Esta relación tendrá unas propiedades concretas que facilitarán que editor de ontologías pueda dar respuesta a esta consulta.

1.3 Metodología del estudio empírico

Con el fin lograr los objetivos específicos estipulados más arriba, se llevará a cabo un estudio empírico que consta de las siguientes etapas.

En primer lugar, se elaborará un catálogo de relaciones conceptuales exhaustivo a partir de la revisión de las propuestas de clasificación de las relaciones presentadas en la parte teórica. En este catálogo se describirán en lenguaje natural las relaciones conceptuales que habrán de formalizarse para su representación en el editor de ontologías.

Una vez establecido el catálogo de relaciones conceptuales, se procederá a su formalización en función de elementos como las clases conceptuales implicadas en la relación y las propiedades de las relaciones (transitividad, simetría, existencia de relaciones inversas y recursividad). El resultado de esta etapa será un catálogo de relaciones conceptuales formalizadas y listas para su implementación en el editor de ontologías.

La tercera fase consistirá en la implementación de las relaciones conceptuales formalizadas en el editor de ontologías *Protégé*. Se describirá el modo en el que se han hecho explícitas las relaciones conceptuales en el editor, teniendo en cuenta los componentes y características del mismo.

Una vez implementadas las relaciones en la base de datos conceptual elaborada con el editor de ontologías, se procederá a introducir los conceptos sobre los productos cerámicos acabados y sus conceptos relacionados. Esta fase nos permitirá valorar la utilidad de nuestro modelo de implementación para representar las relaciones conceptuales en el editor de ontologías.

A continuación, se elaborarán una serie de consultas tipo para comprobar la utilidad de la base de datos conceptual almacenada mediante el editor de ontologías para la representación y recuperación de la información sobre relaciones conceptuales.

Los resultados de nuestro estudio empírico serán los siguientes:

1. Un catálogo exhaustivo de relaciones conceptuales.
2. Una propuesta de formalización de las relaciones conceptuales mediante la explicitación de las clases conceptuales susceptibles de intervenir en cada relación y las propiedades de cada una de ellas.

3. Una propuesta de implementación de las relaciones conceptuales en el editor de ontologías *Protégé*.
4. Una base de conocimiento en la que se almacenen los conceptos sobre los productos cerámicos acabados y sus conceptos relacionados, en la cual las relaciones conceptuales estén recogidas de forma explícita.
5. Un protocolo para la introducción de conceptos y relaciones en la base de datos conceptual elaborada en el editor de ontologías *Protégé*.
6. Un primer listado de consultas que podrían plantearse en dicha base de conocimiento, mediante las cuales podremos evaluar la recuperación de la información sobre los conceptos y sus relaciones mediante el editor de ontologías.
7. Una serie de propuestas encaminadas a mejorar la utilidad del editor de ontologías para su uso en terminografía.

1.4 Estructura de la tesis

La tesis se divide en dos partes: la primera, que comprende los capítulos 2, 3 y 4, está dedicada a los fundamentos teóricos, y la segunda, que abarca los capítulos 5, 6 y 7, corresponde a la presentación del trabajo empírico y sus resultados.

En el capítulo 2 se presentarán y revisarán las propuestas de clasificación de las relaciones conceptuales entre términos que se han hecho desde la Terminología. Si bien la revisión teórica se centrará en las aportaciones realizadas desde esta disciplina, no se pasarán por alto otras realizadas desde la Lingüística y la Psicología, que pueden ser de utilidad para la comprensión y caracterización de las relaciones conceptuales entre términos.

En el capítulo 3 se revisarán las propuestas de representación y formalización de las relaciones conceptuales en la Terminología y la Lingüística. Veremos cuáles son los modelos tradicionales de representación, y a continuación presentaremos las aportaciones que, en los últimos años, han buscado una mejora de los sistemas de representación del conocimiento especializado.

En el capítulo 4 se describirá qué se entiende por ontología en la ingeniería del conocimiento y cuáles son sus principales componentes. A continuación presentaremos algunas ontologías relevantes y los lenguajes ontológicos más representativos. Se describirán y compararán las principales herramientas para el

desarrollo de ontologías. Por último, valoraremos diferentes editores de ontologías con el fin de seleccionar uno de ellos para nuestro trabajo.

El estudio empírico, que comprende los capítulos 5, 6 y 7, se inicia en el capítulo 5, en el que se detallará la metodología del estudio así como los recursos empleados. Se presentarán asimismo en este capítulo las herramientas informáticas de las que nos hemos servido, haciendo especial hincapié en la descripción del editor de ontologías seleccionado en el capítulo 4.

En el capítulo 6 ofreceremos nuestro catálogo de relaciones conceptuales y describiremos nuestra propuesta de formalización. Cada relación se describe en función de las clases conceptuales entre las que se establece y de sus propiedades lógicas. En este capítulo se propondrá también un modo de implementar las relaciones conceptuales empleando el editor de ontologías *Protégé*.

El capítulo 7 se dedicará a la descripción del proceso de introducción de datos en la base de conocimiento elaborada con el editor de ontologías. Analizaremos las dificultades surgidas durante el proceso de introducción de datos y explicaremos las soluciones adoptadas en el marco de nuestro trabajo. Se hará una primera valoración de la utilidad del editor de ontologías como herramienta para el almacenamiento y recuperación de los datos sobre relaciones conceptuales. De este modo seremos capaces de proponer algunas adaptaciones del editor de ontologías, de cara a su utilización para la elaboración y consulta de ontologías terminológicas.

En el capítulo 8 se presentarán las conclusiones del trabajo y las futuras vías de investigación que abre esta tesis doctoral.

PRIMERA PARTE: FUNDAMENTOS TEÓRICOS

**CAPÍTULO 2: LAS RELACIONES CONCEPTUALES
ENTRE TÉRMINOS**

2 Las relaciones conceptuales entre términos

El objetivo de este capítulo es la revisión de distintos enfoques sobre las relaciones conceptuales propuestos desde la teoría de la terminología y desde disciplinas afines, como la Semántica Léxica. Este análisis nos servirá como punto de partida teórico para el estudio de las relaciones conceptuales entre términos desde una perspectiva terminológica próxima a la Lingüística.

La terminología, surgida en el primer tercio del siglo XX con la intención de facilitar la comunicación especializada, tiene su máximo exponente en Wüster y sus seguidores de la Escuela de Viena, que a lo largo del siglo pasado sentaron las bases de la disciplina teórica y formularon los principios que rigen el trabajo terminológico práctico en lo que se ha dado en llamar Teoría General de la Terminología (en adelante TGT). En el apartado 2.1 se presentan de manera sucinta dichos principios básicos, haciendo especial hincapié en lo que se refiere al análisis de conceptos y al establecimiento de relaciones entre ellos. Definiremos qué se entiende por concepto (2.1.1), qué son las características (2.1.2), cómo se clasifican y cómo pueden combinarse conceptos preexistentes para generar otros nuevos (2.1.3). Posteriormente veremos qué se entiende por relación conceptual desde la TGT, cuáles son las relaciones conceptuales relevantes para este enfoque y las tipologías de relaciones conceptuales clásicas (2.1.4) y algunas más recientes que profundizan en las relaciones aplicando el paradigma de la TGT (2.1.5).

En el apartado 2.2 presentaremos algunas propuestas teóricas que, siempre en el seno de la terminología, cuestionan y se replantean los principios de la teoría clásica con el fin de acercarla al ámbito de la Lingüística y las Ciencias Cognitivas. Nos detendremos en aquellos que ofrecen nuevas propuestas o profundizan en la clasificación de las relaciones conceptuales. En el apartado 2.2.1 se recoge el enfoque funcionalista y dinámico propuesto por Sager (1990) y Kageura (2002). El apartado 2.2.2 está dedicado a la presentación de la Teoría Comunicativa de la Terminología propuesta por Cabré (1999), así como del catálogo de relaciones conceptuales que en el marco de esta teoría propone Feliu (2000 y 2004). Por último, en el apartado 2.2.3 nos fijaremos en la visión prototípica del concepto propuesta por Temmerman (2000) en su enfoque sociocognitivo.

El apartado 2.3 definiremos las propiedades matemáticas de las relaciones, ya que consideramos que son un aspecto valioso para su formalización. En el apartado 2.4 estudiaremos en mayor detalle la complejidad de algunas relaciones, como la hponimia (2.4.1), la meronimia (2.4.2), la sinonimia (2.4.3) y las relaciones de tipo causal (2.4.4).

Por último, en el apartado 2.5 presentamos las conclusiones de este capítulo, con el que pretendemos ofrecer un panorama amplio de la concepción clásica sobre las relaciones conceptuales y de las aportaciones y revisiones propuestas desde los nuevos enfoques teóricos y desde otras disciplinas que, a nuestro juicio, completan nuestra concepción de las relaciones conceptuales entre términos.

2.1 Teoría General de la Terminología

La terminología surge como disciplina independiente en el primer tercio del siglo XX, en un contexto en el que la armonización internacional en la ciencia y la técnica se hacía necesaria en aras de la comunicación y del intercambio internacionales, en un momento de desarrollo de la ciencia y de la técnica. En este contexto, Eugen Wüster publica su tesis doctoral *The Machine Tool* (1968) y sienta las bases y los principios metodológicos que tradicionalmente han regido el trabajo terminológico. El objetivo prioritario inicialmente es la normalización terminológica, y por ello el carácter de la TGT es prescriptivo. Wüster propone una terminología llevada a cabo por especialistas en cada materia, en la que el centro y origen de todo el trabajo terminológico es el concepto, a partir del cual se establecen relaciones biunívocas entre cada concepto y su denominación. Esto es lo que se conoce como proceso onomasiológico.

Felber (1984: 98) identifica tres características específicas de la teoría general de la terminología. En primer lugar, cualquier trabajo terminológico tiene que tomar como punto de partida los conceptos, y su objetivo ha de ser la delimitación estricta de estos. Asimismo, defiende que la “esfera de los conceptos” es independiente de la “esfera de los términos”, es decir, que las vertientes conceptual y denominativa están separadas, y la terminología debe ocuparse únicamente de la primera de ellas.

En segundo lugar, Felber mantiene que lo único relevante para el terminólogo han de ser los términos en tanto en cuanto se refieren a los conceptos

(es decir, las terminologías, las colecciones de términos). Esto supone que la dimensión lingüística del término (sintáctica, gramatical, fonética) pasa a un segundo plano.

Por último, la visión que la terminología tiene de la lengua es una visión sincrónica, es decir, la terminología sólo se interesa por los significados actuales de los términos y el sistema conceptos en el que se ubican en un momento concreto.

Junto a estas precisiones sobre la teoría de la terminología, Felber (1984: 98-99) identifica tres peculiaridades de la naturaleza de las terminologías (entendidas como el conjunto de términos que se emplea en una disciplina). Primero, se trata de creaciones deliberadas, frente a lo que sucede en la lengua general, en la que el uso rige la evolución. A continuación, mantiene que para la normalización de términos individuales es necesario que existan directrices translingüísticas unificadas. Por último, defiende la preferencia por la forma escrita de los términos frente a su forma fónica.

Presentamos a continuación las definiciones de los conceptos básicos que manejaremos en el presente trabajo, tal y como se conciben desde la TGT. En el primer apartado nos centramos en la naturaleza del concepto como elemento central de todo trabajo terminológico. En el segundo apartado, revisaremos la definición de característica y presentaremos las tipologías de características propuestas por la TGT. A continuación veremos los métodos que permiten la formación de nuevos conceptos. Por último, haremos un repaso de las definiciones y tipologías de relación conceptual propuestas desde la TGT.

Antes de continuar, conviene hacer una precisión sobre la naturaleza polisémica de la palabra *término*. La Teoría General de la Terminología define *término* en dos sentidos. Por un lado, la norma alemana DIN 2342 (citada en Arntz y Picht, 1989), lo define como la suma de concepto y denominación. Por otro, la norma ISO 704 (2000: 26) mantiene que un término es «une désignation composée d'un ou plusieurs mots et représentant un concept général dans une langue de spécialité», es decir la designación de un concepto por medio de una expresión lingüística. Este segundo sentido es el más extendido entre los teóricos de la terminología.

Sobre el uso de *denominación* también existe cierta controversia. Algunos autores, como Depecker (2000), prefieren utilizar *designación*, pues consideran

que *denominación* da a entender que hablamos sólo de nombres y no de otras clases de palabra¹.

2.1.1 Naturaleza del concepto

Wüster, considerado el padre de la disciplina terminológica, (1998: 39) ofrece la siguiente definición de concepto:

Todo concepto, con excepción de los conceptos de objetos individuales, corresponde a los elementos comunes que los seres humanos perciben en un gran número de objetos y que utilizan como medio de clasificación mental (para *entender*) y, por consiguiente, también para comunicarse. Por lo tanto, el concepto es un *elemento del pensamiento*.

(Wüster, 1998: 39)

Según esta definición, el concepto se interpreta en función de los “elementos comunes” de los objetos que percibimos, y destaca su utilidad para la comprensión y la comunicación. Los conceptos se sitúan, por tanto, en el plano del pensamiento.

La norma UNE 1-070-79 (correspondiente a la norma ISO 1087) añade que los conceptos pueden representar mentalmente no sólo entidades, sino también cualidades, acciones e incluso localizaciones, situaciones o relaciones:

Las nociones pueden ser la representación mental no solamente de seres o de cosas (expresadas por sustantivos), sino igualmente, en un sentido más amplio, de cualidades (expresadas por adjetivos o sustantivos), acciones (expresadas por verbos o sustantivos), y aún localizaciones, situaciones o relaciones (expresadas por adverbios, preposiciones, conjunciones o sustantivos).

(UNE 1-070-79, 1979: 3)

Por su parte, Dahlberg (1982), citada por Picht y Draskau (1985: 37) ofrece un punto de vista más cercano a la lógica formal, y define *concepto* como

¹ En nuestro trabajo, emplearemos la palabra *término* en los dos sentidos indistintamente, para referirnos sólo a la denominación o para referirnos a la suma de concepto y denominación. Será el contexto el que determine a qué sentido nos estamos refiriendo en cada momento. En cuanto al uso de *denominación*, preferimos mantenerlo porque consideramos que está suficientemente arraigado entre los especialistas, y no suele dar lugar a confusión sobre la categoría gramatical de los términos.

la síntesis, fijada por medio de una denominación, de los distintos predicados sobre un objeto determinado.

Para Arntz y Picht (1989/1995: 58) el término se define como una unidad constituida por un concepto y su denominación. El concepto se define como “una unidad de pensamiento que abarca las características comunes asignadas a objetos”. Esta definición, al igual que la de Wüster, destaca el hecho de que el concepto pertenece al ámbito del pensamiento, y llama la atención sobre las características como rasgos distintivos de los objetos mediante cuya abstracción llegamos al concepto².

Por su parte, la norma ISO 704 (2000: 2) pone el acento sobre el hecho de que a la terminología le interesan únicamente los conceptos en tanto que representaciones mentales de los objetos en un contexto o en un ámbito de especialidad. Es decir, los conceptos que interesan a la terminología son únicamente aquellos que representan a los objetos pertenecientes a la realidad especializada.

A la luz de las definiciones que acabamos de presentar, y de otras similares como las de Cabré (1993) y Pavel y Nolet (2001), podemos resumir aquellos aspectos que coinciden en destacar todos los autores:

- a) Los conceptos pertenecen al ámbito del pensamiento, y no al de los objetos de la realidad, si bien surgen a partir de ella mediante la abstracción de sus características.
- b) Los conceptos permiten estructurar el conocimiento sobre los objetos.
- c) Los conceptos abarcan no sólo a objetos o entidades materiales, sino también a cualidades, acciones y relaciones.
- d) La terminología se interesa únicamente por aquellos conceptos relevantes para los ámbitos de especialidad, y no por los de la lengua general.
- e) Los conceptos desempeñan un papel en la comunicación.

² Antia (2000: 93) define abstracción como un proceso de reducción y generalización que implica la identificación de las propiedades compartidas por un conjunto de objetos, que constituyen las características del mismo.

En definitiva, para la TGT un concepto es un elemento del pensamiento generado a partir de la abstracción de sus propiedades, que permite la estructuración del conocimiento sobre los objetos materiales e inmateriales relativos a un área de especialidad, así como la comunicación del conocimiento especializado en dicha área.

Una vez precisado qué se entiende por concepto en la TGT, haremos referencia a una de las principales tareas a las que debe hacer frente la terminología: el análisis de conceptos, que se define como el proceso de descubrir y representar la estructura conceptual que subyace a los términos de un dominio (Meyer y Mackintosh, 1996: 1).

En esta labor de análisis desempeñan un papel fundamental las características. Por ello, dedicaremos el siguiente apartado a exponer las distintas definiciones y clasificaciones de las características propuestas por la Teoría General de la Terminología.

2.1.2 Características

Las características son elementos del pensamiento que representan una propiedad de un objeto que sirve para la formación y delimitación de un concepto. La norma ISO 704 (2000: 3) destaca la importancia de las características como elemento fundamental en la formación de conceptos. Las características se combinan de una manera única para formar conceptos, cuya representación en un lenguaje de especialidad se realiza generalmente por medio de una denominación (es decir, un término, una apelación o un símbolo).

Como nos recuerdan Arntz y Picht (1989/1995: 78), las características son a su vez conceptos, cuya importancia radica en su capacidad para determinar la intensidad de otros conceptos, para formar términos y estructurar sistemas de conceptos, y para determinar la equivalencia entre términos.

2.1.2.1 Intensión y extensión

Antes de continuar con nuestra descripción de las características, conviene definir qué se entiende por **intensión** (también llamada **comprensión**) y **extensión** de un concepto, ya que, como afirma Depecker (2000: 94), son los dos ejes según los cuales se puede afrontar el análisis de conceptos por medio de sus características. La norma ISO 704 (2000: 4) define la intensidad de conceptos

como el conjunto de características que constituyen la unidad del concepto y que permiten delimitarlo. La extensión, por su parte, se define como el conjunto de los objetos considerados como conjunto y conceptualizados por el concepto. Wüster (1998: 40) entiende la extensión conceptual de dos modos. Por un lado, la define como el conjunto de todos los conceptos subordinados que comparten el mismo grado de abstracción. Por otro, hace referencia al conjunto de todos los objetos individuales que abarca este concepto.

Picht y Draskau (1985: 41) distinguen también entre dos significados de extensión, que llaman **extensión por semejanza** y **extensión por composición**. La primera designa al conjunto de todos los objetos individuales a los que se refiere un concepto pertenecientes al mismo nivel de abstracción. La segunda hace referencia al conjunto de todas las partes de un todo, consideradas por separado. No obstante, consideran que la extensión por composición no es en realidad un caso de extensión en sentido estricto, sino que se refiere a la relación entre un todo y sus partes.

Conviene diferenciar el significado de los conceptos de intensión y extensión conceptual tal y como se entienden en terminología del que se les asigna en Lingüística, donde se suele entender por intensión el conjunto de las propiedades esenciales que un objeto o clase de objetos debe tener para que una palabra concreta pueda aplicársele, mientras que la extensión se refiere a la clase de objetos a la cual se puede asignar correctamente una palabra (Lyons, 1977: 158).

2.1.2.2 Clasificaciones de las características

En la bibliografía clásica sobre terminología (Felber, 1984; Picht y Draskau, 1985; Cabré, 1993; Arntz y Picht, 1989; ISO 704, 2000) encontramos distintas propuestas de clasificación de las características, atendiendo a los siguientes criterios:

- a) La relación que mantienen las características con el objeto al que representan
- b) Su relevancia en la conformación del concepto
- c) La relación entre características
- d) La complejidad de las características.

A continuación ofrecemos nuestra propia síntesis de todas estas propuestas de clasificación surgidas de la TGT, atendiendo a los criterios anteriormente expuestos.

a) Según la relación que mantienen con el objeto que representan

Atendiendo a la relación que tienen las características con el objeto al que representan, se distingue entre **características intrínsecas o inherentes** y **características extrínsecas o relacionales** (Wüster, 1998; Picht y Draskau, 1985 y Arntz y Picht, 1989).

Las primeras se refieren al objeto en sí mismo, no en relación con otros. Una característica intrínseca puede observarse mediante un examen del objeto y no requiere conocimientos sobre su origen o uso (Wüster, 1998: 46). Entre ellas destacan la forma, el tamaño, el color y el material.

Las características extrínsecas se conciben a partir de la relación de unos objetos con otros. Hacen referencia al origen, la potencia o rendimiento, la situación, o el inventor, entre otras. Dentro de ellas, suele distinguirse entre **características de uso** y **características de procedencia u origen**. Las primeras hacen referencia a la finalidad con la que se emplea un objeto, y entre ellas se incluye la aplicación, la integración espacial o el funcionamiento. Las características de procedencia u origen indican dónde, a través de quién o cómo comienza a existir un objeto, o comienza a ser conocido. Entre ellas se encuentran el productor, el inventor, el proveedor, el método de fabricación, el país y la época de origen, entre otras.

En esta misma línea, Picht y Draskau (1985) y Arntz y Picht (1989) proponen la distinción entre **características de estado** (también llamadas características propias), que coinciden a grandes rasgos con las características intrínsecas, **características relacionales** (de procedencia, de uso, de comparación, de valoración o de situación en el espacio) y **características funcionales** (potencia o rendimiento, empleo). Estos dos son tipos de características extrínsecas.

b) Según la relevancia de las características para la conformación de un concepto desde la perspectiva de un ámbito de especialidad concreto

Otra distinción entre características puede establecerse en función de la importancia relativa de las mismas para la definición del concepto desde una perspectiva concreta. La norma ISO 704 (2000: 5) distingue entre **características**

esenciales y **no esenciales**, también llamadas por algunos autores complementarias (Cabré, 1993) o accidentales (Dahlberg, 1982). Se considera que una característica es esencial si es necesaria para definir un concepto en un dominio concreto, es decir, si su ausencia modifica el concepto de manera sustancial. Esta distinción depende de los objetivos de cada trabajo terminológico concreto. Así, lo que se considera esencial desde un punto de vista determinado, podría no serlo si la clasificación se hace desde una perspectiva distinta o para un ámbito temático distinto. Dentro de las características esenciales, la norma ISO 704 (2000) hace referencia a las **características distintivas**, que son aquellas características esenciales que permiten situar al concepto con respecto a otros conceptos del dominio. Wüster (1998: 46) y Felber (1984: 117) denominan a este tipo de características **clasificadoras**.

Por ejemplo, en la definición de del concepto *oxígeno*, las características que se consideran esenciales desde la perspectiva de la química (su peso atómico, el tipo de enlaces químicos que establece, etc.) no coinciden con las que serían esenciales para la definición de este mismo concepto desde la perspectiva de la asistencia clínica (en qué casos se debe aplicar el oxígeno, en qué proporciones, etc.).

c) Según su relación con otras características

Otra posible distinción, propuesta por Wüster (1998: 51) y recogida también por Felber (1984), Picht y Draskau (1985) y Arntz y Picht (1989), es la que se establece entre **características equivalentes** y **características no equivalentes**.

Se dice que dos características son equivalentes si, aunque sean distintas, pueden sustituirse la una por la otra en un contexto determinado sin que ello modifique el significado global del término en cuestión. Por ejemplo, al referirnos a un triángulo, las características equilátero y equiangular son intercambiables. Dos características no equivalentes son las que no se pueden sustituir mutuamente sin cambiar de manera sustancial el significado del concepto.

También se suele distinguir entre **características dependientes** e **independientes**. Las primeras desempeñan una labor en los sistemas conceptuales lógicos, ya que en cada nivel de clasificación se añade una característica que restringe el concepto superordinado. Toda característica dependiente requiere que una característica superordinada la preceda en una serie vertical de conceptos. La

característica que se añade en el nivel de abstracción inferior presupone la presencia de otra característica en el nivel superior. Las características independientes son aquellas que no se condicionan mutuamente, es decir, aquellas en las que el orden y la forma de combinarlas carece de importancia.

Felber (1984: 119) incluye las características dependientes e independientes dentro de las características no equivalentes.

Un ejemplo, extraído de Felber (1984: 120), sería el concepto *grinding machine*, que depende jerárquicamente del concepto *machine tool*, ya que la característica específica *grinding* está subordinada a la característica general *mechanical machining with a tool*.

En el caso de que en un trabajo terminológico se consideren características equivalentes, se recomienda establecer un orden de preferencia entre ellas. Las características intrínsecas prevalecen sobre las extrínsecas, y dentro de estas últimas son prioritarias las características de uso con respecto a las de origen. Este orden de preferencia tiene consecuencias tanto para la definición como para la denominación de conceptos de cara a la normalización.

Por su parte, las características independientes pueden sucederse mutuamente en diferentes niveles de una serie vertical de conceptos, y pueden combinarse de forma arbitraria. Un ejemplo serían las características *transatlántico* y *charter* en el concepto *vuelo charter transtalántico*, ya que el hecho de que un vuelo sea charter no influye en el hecho de que sea o no transatlántico, y viceversa.

d) Según su complejidad

Atendiendo a la complejidad de las características, algunos autores, como Felber (1984) y también Arntz y Picht (1989), distinguen entre **características simples** y **complejas**, dependiendo del número de propiedades que se vean implicadas. Las características simples se refieren a una única propiedad de los objetos, como por ejemplo el hecho de tener forma circular. Las características complejas pueden ser una combinación de dos o más características, como por ejemplo “disposición unilateral mortis causa” aplicado a un testamento. Arntz y Picht (1989) mantienen que en el trabajo terminológico no siempre es recomendable descomponer las características complejas.

A continuación ofrecemos un resumen de las tipologías de características presentadas en este apartado.

Resumen sobre las tipologías de características:

a) Según la relación que mantienen con el objeto al que representan:

- a. Características intrínsecas
- b. Características extrínsecas:
 - i. Características relacionales
 - ii. Características funcionales

b) Según su relevancia para la conformación de un concepto en un ámbito de especialidad concreto:

- a. Características esenciales
- b. Características no esenciales

c) Según su relación con otras características

- a. Características equivalentes
- b. Características no equivalentes
 - i. Características dependientes
 - ii. Características independientes

d) Según su complejidad:

- a. Características simples
- b. Características complejas

2.1.3 Combinaciones de conceptos

Generalmente se identifican cuatro métodos fundamentales para la formación de nuevos conceptos mediante la combinación de conceptos existentes: determinación, conjunción, disyunción e integración. Los tres primeros establecen relaciones lógicas entre el concepto existente y el nuevo concepto creado, y la integración establece una relación ontológica entre el concepto existente y el nuevo (ver apartado sobre relaciones, 2.4). Las relaciones lógicas provienen del propio contenido de los conceptos, es decir, de sus características, mientras que en las relaciones ontológicas lo que se compara son las características que intervienen en la realización de los conceptos (Wüster, 1996a: 115)

La **determinación** se produce cuando se enriquece de manera explícita la intensión de un concepto existente por la adición de nuevas características. La intensión del concepto previo está incluida en la del nuevo concepto, que es más específico. El nuevo concepto es un subordinado del existente. Un ejemplo,

extraído de Clavería y Torruella (1993: 340), que ilustra la formación de nuevos conceptos por determinación sería la formación de *puerta blindada* a partir de *puerta*. La intensión de *puerta* se enriquece mediante la adición de nuevas características (materiales más seguros), mientras que la extensión del nuevo concepto comprende un número menor de objetos que la del concepto inicial.

La **conjunción** consiste en la fusión de las intensiones de dos conceptos básicos. El concepto resultante procede de la suma de los conceptos constituyentes, que contribuyen con sus características a la intensión del nuevo concepto a partes iguales. En este caso, la extensión del nuevo concepto también se ve reducida, pero la intensión se ve enriquecida con las características de los conceptos preexistentes. El nuevo concepto es un subordinado común de los dos conceptos constituyentes. Un ejemplo, procedente también de Clavería y Torruella (1993: 343) sería el concepto resultante de la conjunción de *bar* y *restaurante* para formar el nuevo concepto *bar-restaurante*.

La **disyunción** genera un nuevo concepto a partir de la fusión de las extensiones de los conceptos preexistentes. Los conceptos básicos establecen entre sí una relación de oposición o complementariedad y son excluyentes. Un ejemplo sería la formación del concepto *billete de ida y vuelta* a partir de *billete de ida* y *billete de vuelta*. En la disyunción los conceptos constituyentes se fusionan en un nuevo concepto superordinado común.

Clavería y Torruella (1993: 343) agrupan las combinaciones de conceptos por conjunción y disyunción bajo el epígrafe *yuxtaposición de conceptos*. En la yuxtaposición los dos conceptos previos contribuyen en pie de igualdad a la formación del nuevo concepto, sin que uno sea más relevante que otro.

Por último, la **integración** se produce por la fusión de los conceptos constituyentes y consiste en la unión de los objetos a los que se refieren los conceptos previos. Un ejemplo extraído de Picht y Draskau (1985: 44) es la integración de los conceptos *hombre* y *mujer* para dar lugar al concepto *pareja*.

2.1.4 Relaciones conceptuales: definición y tipología

Los conceptos no existen como unidades de pensamiento aisladas, sino que siempre aparecen en relación con otros conceptos, aunque esta relación no siempre se exprese de forma explícita. En terminología se suele hablar de **relaciones conceptuales** para referirse a los vínculos que se establecen entre

conceptos dentro de un ámbito de especialidad. Las relaciones conceptuales son fundamentales para establecer sistemas que reflejan la estructuración del conocimiento especializado.

A pesar de que se destaca la importancia de las relaciones conceptuales, no hemos encontrado en las normas ni en los manuales básicos de la TGT ninguna definición explícita de lo que se entiende por relación conceptual. Simplemente se destaca su importancia y se pasa a clasificarlas.

Tanto la norma ISO 704 (2000) como casi todos los manuales básicos de terminología consultados (Arntz y Picht, 1989; Cabré, 1993) establecen una primera distinción básica entre **relaciones jerárquicas** y **no jerárquicas**. Wüster (1998: 58) y Felber (1984), sin embargo, prefieren utilizar la expresión relación jerárquica en sentido amplio, y prefieren hablar de relaciones lógicas, ontológicas y de efecto. En nuestra presentación de las relaciones mantenemos, no obstante, la distinción básica entre relaciones jerárquicas y no jerárquicas, puesto que su uso está bastante extendido entre los teóricos de la TGT.

A continuación haremos un repaso por las distintas clasificaciones de las relaciones conceptuales propuestas desde la TGT. Dado que muchas de las propuestas teóricas coinciden a grandes rasgos en su presentación de las relaciones conceptuales, las describiremos de forma conjunta, indicando las peculiaridades de cada tipología cuando sea necesario.

2.1.4.1 Relaciones jerárquicas

Las relaciones jerárquicas son aquellas en las que los conceptos están organizados en niveles de abstracción en los cuales un **concepto superordinado** se subdivide en al menos un **concepto subordinado**. Los conceptos superordinados son más genéricos y abarcan a los conceptos subordinados, que son más específicos. Dentro de un mismo nivel se habla de **conceptos coordinados** que responden a un mismo criterio de subdivisión, y que juntos constituyen una dimensión. Un concepto superordinado puede poseer varias dimensiones, en cuyo caso se dice que el sistema conceptual resultante es multidimensional (ISO 704: 5)³.

³ El concepto de multidimensionalidad es tratado en mayor profundidad por otros autores, como Meyer *et al* (1992), tal como veremos más adelante.

Dentro de las relaciones jerárquicas se distingue entre **relaciones lógicas** (también llamadas genéricas o de abstracción) y **relaciones ontológicas** (también denominadas partitivas, parte-todo o meronimia)⁴.

2.1.4.1.1 Relaciones lógicas

Las relaciones lógicas se basan en la semejanza entre conceptos, es decir, en el hecho de que los conceptos posean o no características comunes. En estas relaciones podemos identificar un concepto superordinado cuya comprensión se ve enriquecida por la adición de características distintivas suplementarias (determinación), lo que hace que el concepto resultante se encuentre en un nivel inferior de clasificación. El concepto superordinado se denomina **concepto genérico** y el subordinado **concepto específico**.

Wüster (1998:40) considera que las relaciones lógicas son relaciones directas que se basan en el grado y en el tipo de semejanza. Se establece este tipo de relación cuando dos conceptos comparten una parte de su intensión, aunque sea mínima.

Por tanto, cabe hablar de **subordinación lógica** y de **coordinación lógica**, que dan lugar, respectivamente, a relaciones verticales y relaciones horizontales entre conceptos.

Wüster (1998) y Felber (1984) se refieren también a las **relaciones lógicas diagonales**, que se establecen entre dos subconceptos del mismo superordinado que no están vinculados por subordinación ni coordinación lógicas, sino que responden a la clasificación según distintas dimensiones. Así, tanto *caracol* como *mamífero* son animales, pero su subordinación al concepto *animal* se hace por dimensiones distintas. Entre ellos se establece una relación lógica diagonal.

Asimismo, cuando la relación lógica vincula tres o más conceptos se habla de series lógicas verticales, horizontales o diagonales. En este caso, Wüster (1998) distingue dos posibilidades, dependiendo de si la subordinación y coordinación lógicas se producen simultáneamente o no. Si sólo se produce subordinación o

⁴ Esta dispersión denominativa para referirse a las relaciones contraviene la pretendida precisión de los términos que la terminología pretende conseguir en los lenguajes especializados. En nuestro trabajo hemos optado por las denominaciones *relaciones lógicas* y *relaciones ontológicas* porque consideramos que son más sistemáticas. Sin embargo, las denominaciones conviven –aparentemente sin muchos conflictos– en los distintos manuales.

coordinación, hablaremos de *series lógicas verticales y horizontales*, respectivamente. En el caso de que dos de los tres o más elementos estén coordinados, mientras que el tercero sea superordinado o subordinado de los otros dos, los dos conceptos que están coordinados se consideran **conceptos de base**, de cuya combinación surgirá el tercer concepto.

El siguiente ejemplo, extraído de Arntz y Picht (1989/1995: 110) ilustra la relación genérica: el concepto *motor* es el concepto superordinado de los conceptos *motor de combustión*, *motor eléctrico* y *motor eólico*. Todos estos conceptos mantienen una relación vertical con el concepto *motor*, del que surgen por determinación. Entre los conceptos subordinados se establece una relación horizontal, y se dice que son **cohipónimos**. En este caso, los tres conceptos subordinados aparecen por adición al concepto *motor* de una característica que indica la fuente de energía empleada para su funcionamiento.

2.1.4.1.2 Relaciones ontológicas

Las relaciones ontológicas se basan en la relación existente entre un **concepto incluyente** y las partes que lo componen, que se denominan **conceptos específicos parciales** o **conceptos partitivos**. Al igual que en las relaciones lógicas, en las relaciones partitivas también podemos hablar de **subordinación partitiva** (relación partitiva vertical), **coordinación partitiva** (relación partitiva horizontal) y **relación partitiva diagonal**. Otros autores, como Arntz y Picht (1989), denominan a estas relaciones **relaciones partitivas** o **relaciones parte-todo**, mientras que en las clasificaciones procedentes de la Lingüística, como las de Lyons (1977) y Cruse (1986), se prefiere denominarlas relaciones meronímicas.

Si la relación partitiva se establece entre tres o más conceptos, podemos establecer series partitivas horizontales o verticales.

La norma ISO 704 (2000: 10) mantiene que las partes constitutivas de un todo pueden tener la misma naturaleza (por ejemplo un átomo de una molécula de oxígeno) o pueden ser muy diferentes entre sí. Del mismo modo, las partes pueden ser **obligatorias** o **facultativas**, es decir, esenciales o no. Algunas partes pueden ser, además de esenciales, **distintivas**, ya que permiten la distinción del todo frente a otros conceptos incluyentes similares. Algunas partes pueden ser **múltiples**, como por ejemplo las páginas dentro de un libro, y otras variables

dentro de una gama. La norma propone el ejemplo en francés *stylo*, que puede tener un cartucho, una reserva o una recarga de tinta.

El siguiente ejemplo, extraído de Arntz y Picht (1989/1995: 121), ilustra la relación partitiva. El concepto *bicicleta* se refiere a un objeto constituido por varias partes, como pueden ser la *rueda*, el *manillar* o el *sillín*, que son conceptos partitivos. Estos a su vez establecen entre sí una serie horizontal de conceptos.

Al hablar de las relaciones ontológicas, Wüster (1998) ofrece un punto de vista ligeramente distinto, presentándolas como relaciones existentes entre objetos individuales que se hallan dentro de los mismos conceptos y que se caracterizan por la contigüidad (espacial, temporal, causal, etc.). Se trata en primer lugar de relaciones referenciales, que son relaciones entre conceptos sólo de forma indirecta.

Wüster (1998: 44) considera que las relaciones partitivas –que limita a la contigüidad en el espacio– son únicamente el tipo más importante de relaciones ontológicas, y mantiene que existen otros tipos de relación ontológica basadas en otros tipos de contigüidad. Define las relaciones ontológicas como relaciones indirectas que se dan entre los individuos representados por los conceptos, y se basan en el contacto de estos individuos en el espacio o en el tiempo (contigüidad), y en sus conexiones causales. Dentro de las relaciones partitivas, Wüster distingue entre las que afectan a dos conceptos y las que afectan a tres o más. Dentro de las primeras, se distinguen las **relaciones de inclusión (subordinación partitiva)** y la **coordinación partitiva**. La primera se establece entre un objeto individual (un todo) y sus partes específicas. En este caso, un objeto individual incluye a otro. Se habla entonces de concepto parcial y concepto inclusivo o concepto partitivo superordinado. La coordinación partitiva se da entre dos objetos individuales que son partes de un mismo conjunto. Dichos objetos individuales reciben el nombre de elementos co-constituyentes.

Cuando la relación partitiva se establece entre tres o más conceptos, Wüster distingue tres posibilidades: **relación partitiva horizontal**, **relación partitiva vertical** y **combinación partitiva**. La primera se da cuando hay coordinación y los conceptos conforman una serie partitiva horizontal, mientras que la segunda se da cuando se produce una subordinación. Wüster (1996a: 119-129) da como ejemplo de relación partitiva horizontal la que se establece entre

cabeza, tronco y extremidades con relación a *cuerpo*. Entre *cuerpo* y sus partes se establece una relación partitiva vertical.

La combinación partitiva se produce cuando se establece una relación triangular entre las partes de un conjunto. Al igual que en el caso de la combinación lógica, Wüster menciona aquí un cuarto modo de formación de nuevos conceptos mediante una combinación de conceptos. Se trata en este caso de la **integración**, en la que se combinan objetos individuales incluidos en el concepto partitivo superordinado. Por ejemplo, una *pared* se construye integrando *ladrillos* (Wüster, 1998: 46).

Las normas internacionales y los manuales clásicos de terminología, como el de Arntz y Picht (1989/1995: 121) se limitan a indicar que las relaciones partitivas permiten relacionar el todo y sus partes. Sin embargo, como veremos en el apartado 2.4.2, estudios procedentes de otras disciplinas, como la Semántica Léxica (Iris *et al*, 1988) o la Psicología (Winston *et al*, 1987), amplían esta clasificación, distinguiendo varios tipos de relaciones meronímicas en función de la naturaleza de la relación que se establezca entre el concepto incluyente y el/los conceptos específicos parciales.

Las relaciones jerárquicas han sido las más estudiadas dentro de la teoría clásica de la terminología, puesto que son las más útiles a la hora de estructurar el conocimiento y dar lugar a sistemas (véase el apartado 3.1 sobre sistemas de conceptos).

Sin embargo, incluso desde la TGT se aprecia que las relaciones jerárquicas no son las únicas que permiten estructurar el conocimiento especializado. Por ello las normas y los manuales básicos consideran otro tipo de relaciones, denominadas relaciones no jerárquicas.

2.1.4.2 Relaciones no jerárquicas

Aunque, como hemos visto, existe un alto grado de consenso en la literatura al hablar de las relaciones jerárquicas, esto no ocurre en el caso de las no jerárquicas. Esta denominación se emplea en algunos casos como una especie de “cajón de sastre” en el que se agrupan relaciones de diversa naturaleza. A continuación tratamos de sistematizar los principales tipos de relación no jerárquica a partir de las normas y los manuales básicos consultados.

Entre las **relaciones no jerárquicas** se suele dar especial importancia a las relaciones **secuenciales** (temporales o espaciales) y a las **relaciones asociativas**, a las que en ocasiones se denomina pragmáticas o temáticas. En el caso de las relaciones secuenciales no existe consenso sobre si se deben incluir entre las relaciones jerárquicas o entre las no jerárquicas. Mientras que Cabré (1993) las considera relaciones ontológicas, y por tanto jerárquicas, Arntz y Picht (1989) las ubican claramente entre las relaciones no jerárquicas, y Wright y Budin (1997) las sitúan a medio camino entre ambas. Mientras que Cabré las reduce a las relaciones de coordinación y la relación causa-efecto, Arntz y Picht, siguiendo la norma alemana, las subdividen en siete tipos diferentes (cronológicas, causales, genéticas, de producción, de transmisión, instrumentales y funcionales). Para superar esta falta de uniformidad, a continuación describimos los tipos de relación propuestos por los principales manuales.

La norma ISO (704, 2000: 13) agrupa todas las relaciones no jerárquicas bajo el nombre de **relaciones asociativas**, que se definen como relaciones basadas en una vinculación temática entre conceptos establecida a partir de la experiencia. Dentro de ellas, habla de relaciones en las que se puede identificar una proximidad espacial o temporal.

A continuación ofrece una lista no cerrada de posibles relaciones, que dependerán del ámbito temático en el que estemos trabajando. Las relaciones asociativas que se mencionan (sin ofrecer ejemplos concretos), aunque no se definen expresamente, en la norma ISO 704 (2000: 13) son las siguientes:

- a) Relación materia prima-producto
- b) Relación acción-equipo/herramienta
- c) Relación tamaño-unidad
- d) Relación material-propiedad
- e) Relación material-estado
- f) Relación materia/sustancia-propiedad
- g) Relación elemento concreto-propiedad
- h) Relación elemento concreto-forma
- i) Relación acción-meta
- j) Relación acción-lugar/emplazamiento
- k) Relación acción-actor
- l) Relaciones de los acontecimientos en el tiempo: proceso-tiempo

m) Relación causa-efecto

Por su parte, Arntz y Picht (1989), siguiendo la norma DIN 2330 (1979), distinguen dos tipos de relación no jerárquica, las **relaciones secuenciales** y las **relaciones pragmáticas**. Entre las primeras se pueden enumerar las siguientes:

- a) Relaciones cronológicas
 - i. Relación cronológica simultánea
 - ii. Relación cronológica consecutiva
- b) Relaciones causales (causa-efecto)
- c) Relaciones genéticas (productor-producto)
- d) Relaciones de producción (material-producto)
- e) Relaciones de transmisión (emisor-receptor)
- f) Relaciones instrumentales (instrumento-utilización del instrumento)
- g) Relaciones funcionales (argumento-función)

Las **relaciones pragmáticas** se definen como relaciones menos estrechas, basadas en la proximidad temática entre conceptos, por lo que no pueden incluirse en el grupo de las relaciones jerárquicas ni en el de las secuenciales. El ejemplo propuesto por Arntz y Picht es la relación que se establece entre el concepto *disolución del matrimonio* y otros relacionados temáticamente con él, como *juicio, divorcio o enfermedad mental*.

Wright y Budin (1997) también hablan de relaciones secuenciales, si bien las subdividen únicamente en relación secuencial espacial y relación secuencial temporal. Cabré (1993) opta por incluir las relaciones secuenciales dentro de las relaciones ontológicas, y dentro de ellas se refiere únicamente a la relación causa-efecto.

A pesar de esta diversidad de propuestas, todos los autores coinciden en señalar que la naturaleza de las relaciones no jerárquicas depende necesariamente del ámbito temático, y su utilidad para la elaboración de sistemas conceptuales se establecerá en cada caso.

Para tratar de sistematizar los tipos de relaciones conceptuales propuestos desde la TGT, a continuación ofrecemos un esquema-resumen que recoge la clasificación de las relaciones comúnmente aceptada:

Relaciones conceptuales en la teoría general de la terminología

- a) Relaciones jerárquicas:

- i. Relaciones lógicas (genéricas o de abstracción):
 - 1. Subordinación lógica
 - 2. Coordinación lógica
 - 3. Relaciones lógicas diagonales
 - ii. Relaciones ontológicas (parte-todo, partitivas, meronímicas):
 - 1. Subordinación partitiva
 - 2. Coordinación partitiva
 - 3. Relaciones partitivas diagonales
- b) Relaciones no jerárquicas:
- i. Relaciones secuenciales:
 - 1. Relaciones temporales (cronológicas)
 - a. Relación cronológica simultánea
 - b. Relación cronológica consecutiva
 - 2. Relaciones espaciales
 - 3. Relaciones causales (causa-efecto)
 - 4. Relaciones genéticas (productor-producto)
 - 5. Relaciones de producción (material-producto)
 - 6. Relaciones de transmisión (emisor-receptor)
 - 7. Relaciones instrumentales (instrumento-utilización del instrumento)
 - 8. Relaciones funcionales (argumento-función)
 - ii. Relaciones asociativas (pragmáticas o temáticas)

Valoración

Dentro de la teoría clásica de la terminología, los estudios sobre relaciones conceptuales se han centrado en la descripción de las relaciones jerárquicas, que suelen considerarse las más relevantes para la estructuración del conocimiento especializado. Los manuales básicos y las normas sobre terminología coinciden en cuanto a la subdivisión de las relaciones jerárquicas en genéricas y partitivas, y

difieren en la exhaustividad con la que enumeran las relaciones no jerárquicas. En el mejor de los casos, se limitan a mencionarlas y a aportar algún ejemplo que las ilustre. No obstante, todas las clasificaciones identifican otros tipos de relación que se consideran funcionales para la representación de los conceptos de un área de especialidad, pero sin estudiarlas de forma exhaustiva ni sistemática.

La consideración de las relaciones genéricas y partitivas en los ámbitos de especialidad puede resultar de utilidad en ámbitos muy estructurados; sin embargo, resulta insuficiente para dar cuenta de fenómenos como la multidimensionalidad de las áreas de conocimiento, es decir, los diferentes puntos de vista desde los que se pueden clasificar los conceptos.

Sorprende asimismo el hecho de que la TGT no ofrezca una definición explícita de lo que se entiende por relación conceptual, a pesar de reconocer su relevancia para la estructuración del conocimiento especializado.

Como veremos más adelante, para encontrar una definición formal de relación conceptual en terminología tenemos que remitirnos a trabajos más recientes que se apoyan en un enfoque de las relaciones conceptuales más cercano a la Lógica y a la Lingüística, como Otman (1996), Feliu (2004) u Oster (2003), que tienen en cuenta las aportaciones de otras disciplinas al estudio de las relaciones. Al igual que estos autores, consideramos que, para dar cuenta de la complejidad de los vínculos que se establecen entre los conceptos de un área de especialidad, es necesario definir claramente qué se entiende por cada tipo de relación, indicando cuáles son sus propiedades lógicas, entre qué clases de conceptos puede establecerse, y cómo puede ser representada de manera formal.

A pesar de las limitaciones que –como acabamos de ver– tienen los planteamientos sobre las relaciones propuestos por la Teoría General de la Terminología, no cabe duda de que la TGT sienta las bases para el trabajo terminológico, y por ello consideramos que cualquier estudio de este tipo tiene que tomar como punto de partida estos principios y conceptos básicos, aunque posteriormente sean revisados.

2.1.5 Profundización en la naturaleza de las relaciones conceptuales desde la TGT

Hasta ahora hemos visto la tipología de las relaciones conceptuales presentada en los manuales básicos y en las normas de armonización

terminológica que recogen los principios de la Teoría General de la Terminología. Dentro de esta corriente teórica, en los últimos años se han realizado estudios más completos sobre las las relaciones conceptuales, y que describen un número de relaciones más elevado. En los siguientes apartados resumimos dos de estas aportaciones, que presentamos aquí por su representatividad y exhaustividad. Se trata de las tipologías de relaciones de Weissenhofer (1995) y Nuopponen (1994 y 2005).

2.1.5.1 Weissenhofer (1995)

Aunque mantiene los principios de la TGT, Weissenhofer (1995) defiende que la conceptología, definida como el estudio de la naturaleza de los conceptos y las relaciones que se establecen entre ellos, debería tener en cuenta aspectos lingüísticos procedentes de la semántica y de la formación de palabras, tales como la naturaleza prototípica de los conceptos y las teorías psicolingüísticas sobre el desarrollo morfológico.

A partir del análisis de un corpus sobre béisbol en lengua inglesa, propone un sistema de clasificación morfológica de dicha terminología en función de las **relaciones morfoconceptuales** que se establecen entre los constituyentes de los términos complejos. Así, Weissenhofer presenta un sistema de clasificación exhaustivo para los términos compuestos, basado en las relaciones existentes entre los conceptos a los que se pueden asociar los constituyentes de los términos.

Weissenhofer revisa el modelo conceptual propuesto por Wüster y lo modifica con el fin de tener en cuenta la prototipicidad de algunos conceptos⁵. Esto le permite aplicar dicho modelo conceptual no sólo a los términos representados por un concepto, sino también a los morfemas que forman parte de los compuestos terminológicos. Concluye que dichos morfemas remiten a su propia representación conceptual, es decir, que los morfemas también poseen una representación conceptual y establecen relaciones entre sí. Ampliando el catálogo de relaciones conceptuales de la teoría clásica de la terminología, Weissenhofer

⁵ El concepto de prototipicidad proviene de la psicología cognitiva. Rosch (1978) considera que las categorías conceptuales poseen una serie de ejemplares que se consideran prototípicos (es decir, más representativos de la categoría) y otra de ejemplares menos representativos de la misma, que se consideran menos prototípicos.

considera que es posible dar cuenta de las relaciones morfoconceptuales presentes entre los formantes de los términos complejos.

Weissenhofer identifica dieciséis tipos de relaciones morfoconceptuales que aparecen en los términos complejos del béisbol, y cuantifica su importancia relativa en ese ámbito. A continuación ofrecemos un cuadro que resume y define las dieciséis relaciones morfoconceptuales propuestas por Weissenhofer⁶:

RELACIÓN	PARÁFRASIS O DEFINICIÓN	EJEMPLO
Composition (material)	El concepto B está hecho o se compone de A	<i>leather glove</i>
Resemblance (semejanza)	El concepto B se parece al concepto A	<i>balloon ball</i>
Meronymy (meronimia)	El objeto designado por el concepto B es parte del designado por el concepto A (todo-parte)	<i>baseball cover</i>
	El concepto A designa un objeto material o inmaterial que forma parte del designado por un concepto B inanimado (parte-todo)	<i>cowhide ball</i>
Quantity (cantidad)	El concepto A indica el valor cuantitativo de B	<i>45-foot line</i>
Place (lugar)	Un concepto designa el lugar donde se produce una acción designada por el otro concepto o donde se ubica una persona u objeto	<i>center field</i>
Time (tiempo)	Un concepto indica el tiempo durante el cual se produce la actividad designada por el otro concepto	<i>Sunday ball</i>
Origin (origen)	El primer elemento del compuesto tiene la función de dar nombre al segundo	<i>Hank Aaron trophy</i>
Purpose (finalidad)	El concepto B designa un objeto material o inmaterial o una actividad cuya finalidad es aquella que designa el concepto A	<i>go signal</i>
Agent (agente)	El segundo constituyente designa a un objeto animado o inanimado que puede ser interpretado como el agente que lleva a cabo la acción o actividad designada por el concepto A	<i>back-up player</i>
Equation (ecuación)	El concepto A es el concepto B o al revés	<i>sacrifice hit</i>
Causation (causa)	El objeto o acción designado por uno de los conceptos causa, lleva a cabo o tiene como consecuencia aquello que es designado por el otro concepto	<i>baseball finger</i>
Instrument (instrumento)	El objeto designado por uno de los conceptos se utiliza como instrumento utilizado para o en la actividad designada por el otro	<i>battling glove</i>
Negation	El concepto A expresa la ausencia de aquello que es	<i>no-hit game</i>

⁶ La tabla tal y como se presenta aquí está tomada de Oster (2005: 30-31), ya que esta autora nos ofrece una versión traducida de la propuesta de Weissenhofer (1995: 100-182). Hemos optado por no introducir el porcentaje de aparición de cada tipo de relación, puesto que se refiere únicamente al ámbito del béisbol y no nos parece relevante para nuestro estudio.

(negación)	designado por el concepto B	
Constitute (constitución)	El objeto o actividad designado por uno de los conceptos constituye aquel que es designado por el otro concepto	<i>hit-and-run play</i>
Possession (posesión)	El objeto designado por el concepto B tiene o posee aquello que es designado por el concepto A o al revés	<i>batter's glove</i>
Property (propiedad)	El concepto B es aquello que es designado por el concepto A, este designa un material específico o una propiedad inmaterial (color, temperatura, potencia, etc.)	<i>fastball</i>

Tabla 1. Relaciones morfoconceptuales propuestas por Weissenhofer (1995).

Una de las conclusiones del trabajo de Weissenhofer es que las relaciones morfoconceptuales no se dan únicamente en los términos compuestos, sino que también se observan en términos derivados (por prefijación y sufijación), en términos formados por conversión (recategorización gramatical) y en combinaciones que no suelen considerarse en la formación de términos típica, como las frases verbales y las combinaciones entre adjetivo/participio y nombre.

Valoración

La propuesta de Weissenhofer nos parece interesante en la medida en que propone un puente entre la Terminología clásica y la Lingüística. Si bien su estructuración de las relaciones morfoconceptuales se emplea para clasificar los términos complejos en función de las relaciones que se establecen entre sus componentes, y no para hacer explícitas las relaciones entre conceptos representados por denominaciones independientes, su propuesta es interesante, puesto que los tipos de relación que identifica también se observan entre conceptos representados por términos simples.

Sin embargo, encontramos algunas deficiencias en su propuesta, como el hecho de que no se exprese en todos los casos cuáles son las clases conceptuales relacionadas mediante cada tipo de relación. Asimismo, el modo de llegar a esta clasificación de las relaciones no queda explicado de manera clara en su trabajo, y únicamente ofrece una lista de autores en los que se basa y sus propias adiciones a la lista derivadas de su trabajo con el corpus de textos sobre el béisbol.

2.1.5.2 Nuopponen (1994 y 2005)

En su tesis doctoral de 1994, Nuopponen presenta una clasificación exhaustiva de las relaciones conceptuales tomando como punto de partida las propuestas de Wüster y aplicando métodos terminológicos para su descripción. Su

catálogo de relaciones se concibe como un menú a partir del cual se pueden seleccionar las relaciones necesarias para elaborar sistemas conceptuales en un proyecto en particular. Para Nuopponen, las relaciones se conciben como conceptos de relación que son el resultado de la abstracción, cuyo referente son las relaciones entre entidades individuales. Basándose en la suposición de que el análisis de conceptos se produce en tres niveles (cualitativo, cuantitativo y análisis orientado al sistema), Nuopponen (1994) propone tres tipos de relaciones:

- a) **Relaciones conceptuales cualitativas**, que determinan el tipo de sistema conceptual al que dan lugar,
- b) **Relaciones conceptuales cuantitativas**, que se ocupan de la intensión y extensión de los conceptos,
- c) **Relaciones conceptuales sistémicas**, que son relaciones formales entre los conceptos en un sistema.

Mediante la combinación de estos tres tipos de relaciones conceptuales, Nuopponen elabora un catálogo de relaciones conceptuales muy exhaustivo. En 2005, Nuopponen revisa dicha clasificación para tener en cuenta los últimos avances realizados en la investigación en terminología y en las disciplinas que se ocupan de la organización y recuperación del conocimiento. Nuopponen defiende que los sistemas de gestión de la terminología avanzados, las aplicaciones para la web semántica y para el modelado de conceptos requieren una clasificación más exhaustiva que la propuesta en las normas terminológicas.

En su propuesta distingue entre dos grandes grupos de relaciones conceptuales, las **relaciones lógicas**, basadas en el grado de abstracción de los conceptos, y las **relaciones ontológicas**, que son relaciones entre objetos. Al igual que Wüster, considera que la principal diferencia entre ellas estriba en el hecho de que las primeras son relaciones directas entre conceptos, mientras que las ontológicas surgen de forma indirecta entre los conceptos.

Nuopponen divide las **relaciones lógicas** atendiendo a dos criterios distintos: la posición relativa de los conceptos en el sistema conceptual y la comparación entre la intensión y la extensión de los conceptos. El primer criterio da como resultado la división de las relaciones lógicas en **relaciones de superordenación, subordinación, coordinación y relaciones diagonales**. El segundo permite distinguir entre **relaciones intensionales y extensionales**, que a su vez pueden dividirse en **relaciones de identidad, inclusión, intersección y**

disyunción. Estas relaciones permiten analizar y describir la multidimensionalidad, que la autora entiende como la subdivisión de los conceptos atendiendo a distintos criterios en un mismo nivel de abstracción.

En cuanto a las **relaciones ontológicas**, Nuopponen afirma que los conceptos poseen otras dimensiones no expresadas mediante las relaciones lógicas y los sistemas conceptuales. El número y el tipo de dichas dimensiones depende de la categoría del concepto, es decir, de si se trata de una entidad, actividad, proceso, método o propiedad, y del ámbito temático en el que se lleve a cabo el estudio de las relaciones.

Las relaciones ontológicas se dividen en **relaciones de contigüidad** entre fenómenos concretos y abstractos (espaciales o temporales) y las relaciones que poseen un componente causal (denominadas **relaciones conceptuales de influencia**), si bien la separación entre ambos tipos de relación no siempre es clara.

Las relaciones de contigüidad se subdividen en **relaciones partitivas**, **relaciones de mejora** (*enhancement relations*, entre dos entidades, una de las cuales puede asociarse a otra sin ser una parte de ella), **relaciones locativas**, **relación de material constitutivo** (relación entre una entidad y el material de que se compone), **relación de propiedad** (*ownership relations*, entre una entidad y sus propiedades), **relación de pertenencia**, **relación de rango** (ordenación según un tipo de propiedad) y **relación temporal**.

En las **relaciones temporales de influencia** existe un elemento causal que participa en la relación entre entidades. Se subdividen en **relaciones causales**, **relaciones de desarrollo**, **relaciones funcionales** y **relaciones interaccionales**. Nuopponen (1994) distingue varios tipos de causas, que pueden darse de forma simultánea o alternativa, así como diversos efectos posibles. Las relaciones de desarrollo se basan en las fases del proceso de desarrollo de un individuo, una especie o una familia, y dan lugar, respectivamente, a **relaciones ontogenéticas** (*niño-adulto*), **filogenéticas**, **genealógicas** y **de desarrollo material** (*agua-hielo*). En la categoría de las relaciones de desarrollo se incluye la de cambio de rol de un mismo individuo (*acusado-culpable*).

Las **relaciones funcionales**, que también son relaciones de influencia, se dividen en **relaciones de actividad**, **de origen** e **instrumentales**. Las primeras se basan en una conexión entre el concepto de actividad y los fenómenos

involucrados en ella (*agente, objeto/paciente, herramienta y ubicación*). Las relaciones de origen se establecen entre un concepto referido a una entidad concreta o abstracta y el origen de dicha entidad. Las relaciones instrumentales, por su parte, forman un tercer tipo de relaciones funcionales, coincidentes en parte con las relaciones de actividad.

Por último, las relaciones interaccionales se basan en las relaciones que se dan entre los fenómenos producidos entre los referentes. Nuopponen las divide en **relaciones de transmisión, de dependencia, de correlación y de representación**. En las relaciones de transmisión se utilizan los conceptos de emisor y receptor para referirse tanto a los objetos como a los lugares implicados en la transmisión. Las otras tres relaciones interaccionales son menos comunes, y dependen en gran medida del ámbito temático.

A continuación ofrecemos una tabla que resume su clasificación. En los casos en los que Nuopponen facilita ejemplos, estos se han incluido entre paréntesis y en cursiva, con el fin de ilustrar cada relación.

Relaciones conceptuales lógicas	Superordenación (<i>personal computer – portable computer</i>)	Superordenación directa Superordenación indirecta (<i>computer – palmtop</i>)
	Subordinación (<i>handheld computer – portable computer</i>)	Subordinación directa Subordinación indirecta
	Coordinación	Coordinación directa Coordinación indirecta (<i>portable computer – table top</i>)
	Relación diagonal (<i>minicomputer – palmtop</i>)	
	Relación intensional	Identidad intensional Inclusión intensional Intersección intensional Disyunción intensional
	Relación extensional	Identidad extensional Inclusión extensional Intersección extensional Disyunción extensional

Tabla 2. Relaciones conceptuales lógicas en Nuopponen (2005).

Las relaciones conceptuales en la terminología de los productos cerámicos acabados

Relaciones conceptuales ontológicas	Relaciones conceptuales de contigüidad	Relación partitiva (<i>house – roof, door</i>)	Relación de composición	Superordinación partitiva	a) Superordenación canónica (<i>car – tyre</i>) b) Superordenación facultativa (<i>hotel – restaurant</i>)	
				Subordinación partitiva	a) Subordinación canónica (<i>finger – hand</i>) b) Subordinación facultativa (<i>tree – forest</i>)	
				Coordinación partitiva	a) Coordinación partitiva directa b) Coordinación partitiva indirecta	
		Relación de partición (<i>bread – slice</i>)				
		Relación de conjunto	a) Relación conjunto-elemento (<i>firm – employer</i>) b) Relación elemento-elemento			
		Relación de mejora (<i>car – trailer; computer – CD-ROM</i>)				
		Relación locativa (<i>fish – lake; passenger – airport terminal; tea – teapot</i>)				
		Relación de componente material (<i>olive – oil; egg – cholesterol</i>)				
		Relación de propiedad (<i>silicone – heat-resistant</i>)				
		Relación de rango (<i>business class – tourist class</i>)	a) Relación de orden (<i>president – vice president</i>) b) Relación de equivalencia			
	Relación temporal (<i>pre wash – whashing</i>)	a) Relación de evento b) Relación de sucesión (<i>starters – main course</i>) c) Relación simultánea d) Relación consecutiva				
	Relaciones conceptuales de influencia	Relaciones causales	Secuencia causal	a) Causa de producción-efecto b) Causa explicativa-efecto c) Causa-efecto resultante d) Causa-estado resultante e) Causa-producto resultante		
			Coordinación causal	a) Multicausalidad b) Relación de efecto múltiple		
		Relaciones de desarrollo (<i>frogs egg – tadpole</i>)	a) Relación filogenética b) Relación ontogenética (<i>child – adult</i>) c) Relación genealógica (<i>father – son</i>) d) Relación de desarrollo material (<i>water – ice</i>) e) Cambio de rol (<i>accused – guilty</i>)			
		Relaciones funcionales	Relaciones de actividad	a) Relación de agente b) Relación de objeto c) Relación de herramienta d) Relación de ubicación e) Relación de acción temporal f) Relación teleológica (actividad-finalidad) g) Relación resultativa		
Relaciones de origen			a) Relación de precursor b) Relación producto-instrumento c) Relación resultativa d) Relación de ingrediente e) Relación de lugar de origen			

			f) Relación de momento de origen g) Producto-finalidad de la creación	
		Relaciones instrumentales	a) Agente instrumento b) Relación funcional (entidad-modo de funcionamiento) c) Relación producto-instrumento (actividad-herramienta) d) relación locativa	
		Relaciones interaccionales	Relación de transmisión	
			Relación de transmisión directa (emisor-receptor)	
			Relación de transmisión secuencial	a) Emisor-intermediario b) Intermediario-receptor
			Relación de fuente	a) Emisor/lugar de partida-objeto b) Intermediario-objeto
			Relación de objetivo	a) Objeto-receptor/destino b) Objeto-intermediario
			Transmisión herramienta-objeto	
			Relación de dependencia (<i>employer – employee</i>)	
		Relación de correlación (<i>height – weight</i>)		
		Relación representacional (<i>concept – term; people – parliament</i>)		

Tabla 3. Relaciones conceptuales ontológicas en Nuopponen (2005).

Valoración

Esta clasificación posee la ventaja de cubrir prácticamente todas las relaciones que se pueden establecer entre los conceptos en los ámbitos especializados. Sin embargo, precisamente por su exhaustividad, resulta difícil de manejar (como ya apuntan Madsen *et al*, 2001). Además, algunos tipos de relación, como por ejemplo las relaciones instrumentales, se repiten en varias ramas de la clasificación y a veces resulta difícil distinguir los matices entre los diversos tipos. Echamos en falta la inclusión de ejemplos para todos los tipos de relación, lo cual ilustraría mejor el significado de cada una de ellas. Valoramos positivamente, no obstante, el esfuerzo realizado para recoger y estructurar todos los tipos de relación con criterios terminológicos, yendo más allá de la clasificación de las relaciones conceptuales propuesta por las normas terminológicas.

Otro aspecto que nos parece muy acertado de esta clasificación es el hecho de que tenga en cuenta las clases conceptuales que intervienen en cada relación (objetos o procesos, sobre todo), así como la consideración de los fenómenos implicados en una acción (agente, objeto/paciente, herramienta y lugar) por medio de las relaciones de actividad.

El trabajo de Nuopponen nos permitirá completar nuestra propuesta de clasificación de las relaciones en función de los resultados obtenidos en el trabajo empírico. En la propuesta de clasificación que presentaremos en el capítulo 6 recuperaremos esta consideración de las relaciones que Nuopponen denomina funcionales.

2.2 Nuevos enfoques de la terminología

A continuación presentamos tres enfoques que revisan y completan los postulados de la Teoría General de la Terminología. Se trata del enfoque funcionalista y dinámico de Sager y Kageura (Sager, 1990; Sager y Kageura, 1994 y Kageura, 2002), la Teoría Comunicativa de la Terminología (Cabré, 1999 y Feliu, 2000 y 2004) y el enfoque sociocognitivo de Termmerman (2000)⁷. En los siguientes apartados se presentan de manera sucinta estos nuevos paradigmas para el estudio de los términos, haciendo especial referencia a sus aportaciones a la clasificación de las relaciones conceptuales.

2.2.1 Enfoque funcionalista y dinámico de la terminología

Sager (1990), en su *Curso práctico sobre el procesamiento de la terminología* y en posteriores publicaciones (Sager y Kageura, 1994 y Sager, 2000) propone una teoría de los conceptos de utilidad para la práctica terminológica.

Sager (1990) destaca la doble función que desempeñan los términos como elementos que permiten la representación del conocimiento especializado y como

⁷ No son las únicas revisiones de la teoría clásica que se han sucedido en los últimos años. Cabe destacar, entre otros, el enfoque socioterminológico propuesto por Gaudin (2003), Gambier (1987 y 1991) y Boulanger (1991 y 1995), y desarrollado en Brasil por Faulstich (1998/1999 y 2002), que considera la importancia de los aspectos sociolingüísticos de los términos. Sin embargo, no lo recogemos en el presente trabajo porque sus principales aportaciones no se refieren a la consideración de las relaciones conceptuales.

elementos que hacen posible la comunicación especializada. En su primera función, los términos permiten categorizar la realidad especializada, y posibilitan el avance del conocimiento mediante la creación de nuevos términos. Sager (2000: 49) pone el acento sobre el carácter dinámico de la terminología, cuya relevancia se deriva de dos motivos fundamentales: por un lado, el conocimiento humano está en constante evolución, y por otro, los recursos léxicos existentes son limitados, y por tanto se hace necesario reutilizarlos y recombinarlos de distintos modos.

La teoría de la terminología propuesta por Sager cumple tres tareas básicas (Sager, 1990/1993: 45). En primer lugar, permite situar al concepto como una entidad discreta en la estructura del conocimiento. En segundo lugar, da cuenta de un conjunto de entidades lingüísticas interrelacionadas asociadas con conceptos agrupados y estructurados según principios cognoscitivos. Por último, establece un vínculo entre los conceptos y los términos por medio de las definiciones.

Sager define los conceptos como “construcciones de los procesos cognoscitivos humanos que ayudan a la clasificación de los objetos mediante la abstracción sistemática o arbitraria” (Sager, 1990/1993: 47). No obstante, la revisión de las definiciones de concepto dadas en las normas, le lleva a proponer que el concepto sea considerado un primitivo axiomático, cuya definición se evita por conveniencia.

En cuanto a las características, las identifica como elementos comunes de los objetos, y diferencia entre características esenciales y no esenciales, coincidiendo con lo visto en la descripción de las características de la TGT.

Las características se pueden expresar como propiedades del concepto, o también como relaciones entre conceptos. Sager argumenta que cada característica se remite a una clase determinada de conceptos que la posee y que se constituye en clase de conceptos en virtud de dicha característica común (Sager, 1990/1993: 50).

2.2.1.1 Clases de conceptos: modelo conceptual estático y dinámico

Sager (1990) recurre a la noción de **clase de conceptos** para la determinación de las relaciones entabladas entre ellos, ya que considera conveniente especificar entre qué tipos de conceptos se establecen. De este modo se puede profundizar en los mecanismos de formación de nuevos términos:

Las clases conceptuales de los conceptos determinan las relaciones posibles entre ellos. Por ejemplo, relaciones de producto o material sólo pueden existir entre entidades materiales. De esta forma surge un modelo que presenta restricciones en la naturaleza de las relaciones entre conceptos en virtud de sus clases. La relación de clases de conceptos con categorías de relaciones conceptuales nos permite una mayor penetración para comprender la formación de conceptos y sus términos complejos.

(Sager, 1990/1993: 63)

Sager afirma que las estructuras conceptuales de los campos especializados se distinguen no sólo por la referencia especial, sino también por la naturaleza de los conceptos que predominan en un campo concreto. Por ello es necesario observar y definir las distintas clases de conceptos para poder aplicar los resultados al estudio de las relaciones conceptuales.

Sager (1990/1993: 50-51) identifica cuatro métodos principales para estructurar los conceptos, mediante los cuales se obtienen cuatro tipos que varían según la materia y la situación. A continuación se describen los cuatro métodos y los cuatro tipos de conceptos a los que dan lugar:

- a) Los conceptos se pueden atribuir a una clase, de manera que se identifican tipos de conceptos. Este método da lugar a **conceptos de clases**. Así, por ejemplo, *Venus* se atribuye a la clase de los *planetas*.
- b) Los conceptos se pueden agrupar en categorías⁸ según lo establecen los rasgos distintivos de las clases o grupo de clases, por ejemplo, según las propiedades comunes. Este método da lugar a **conceptos de propiedades**. Un ejemplo sería el concepto *cuadrúpedo*, que agrupa a todos los animales que tienen cuatro patas.
- c) Los conceptos y los grupos de conceptos pueden diferenciarse mediante un proceso de distinción entre categorías y el establecimiento de relaciones entre ellas. Este método da lugar a **conceptos de relaciones**. Ejemplos de conceptos de relaciones serían *hermana*, *más*, *paralelo*.

⁸ Espanha Gomes, H. y De Almeida Campos, M.L. (1996: 251) definen categoría como “a concept with general application that can be used to group other concepts”. Las categorías permiten sistematizar el conocimiento por medio de alguna propiedad que agrupa a los conceptos.

- d) La interacción de categorías de conceptos se puede efectuar en el nivel de las funciones, para relacionar estados con cambios de estados. Este método da lugar a **conceptos de funciones**. Así, por ejemplo una *reacción química* da lugar a un *compuesto químico*.

En un nivel más general, Sager (1990/1993: 52) distingue entre cuatro clases de conceptos, que son **entidades, actividades, propiedades y relaciones**.

Sager y Kageura (1994) proponen una explicación terminológica para entender la interacción entre la cognición y los procesos de denominación. Así, presentan dos modelos conceptuales diferenciados: por un lado, un **modelo conceptual estático**, que permite dar cuenta de los conceptos existentes en una disciplina, y por otro, un **modelo dinámico**, que permite representar el potencial de creación de nuevos conceptos. En su propuesta destacan la importancia del acto de nombrar en la formación de conceptos y establecen las clases conceptuales en función de los actos de habla básicos.

Modelo conceptual estático

En el modelo conceptual estático intervienen tres clases de conceptos, que son los conceptos de **entidad, relación y actividad**.

Conceptos de entidad

En una primera etapa de la cognición, se identifican las similitudes existentes entre los objetos de la experiencia y se agrupan en clases. Así, se establece la diferencia entre conceptos de entidad individuales y conceptos de clases. Los primeros se denominan por medio de nombres propios, y los segundos por medio de nombres comunes. Cuando establecemos conceptos de clases, nos fijamos en las características comunes de los objetos y los ordenamos y estructuramos mediante procesos de generalización, especialización, abstracción, definición, o bien definiendo las entidades complejas en función de sus componentes.

Conceptos de relación

En una segunda etapa, se consolidan los tipos de entidades combinando o separando las clases en grupos en función de las relaciones existentes entre entidades. Así, se constituye una red compleja de nodos conceptuales unidos por medio de arcos o relaciones. Las relaciones se analizan y se les asigna un nombre. De este modo, identificamos conceptos de relación generales, con validez para todos los ámbitos temáticos (por ejemplo, *similitud, diferencia, comparación*), y

conceptos de relación específicos, que se aplican en ámbitos temáticos especializados (por ejemplo, *ratio*).

Conceptos de actividad

Además de un sistema estático de entidades y clases y sus relaciones, a partir de la experiencia se construyen estructuras conceptuales de actividad, que también se organizan en clases. La estructura de los conceptos de actividad requiere la participación de entidades para que puedan llevarse a cabo. Por ejemplo, identificamos conceptos de actividades como *ir* o *enviar*. Los conceptos de actividades, a su vez, pueden relacionarse mediante conceptos de relación. Por ejemplo, la acción *input* y la acción *output* están vinculadas por una relación funcional.

Este enfoque de la formación de conceptos es útil en un modelo de conocimiento que describe la estructura del conocimiento de un modo estático, teniendo en cuenta los conceptos ya existentes en una disciplina concreta.

Modelo conceptual dinámico

Para poder dar cuenta de la evolución del pensamiento, es necesario considerar un modelo dinámico, en el que la estructura del conocimiento puede modificarse, y los conceptos interactúan dentro de una misma clase y entre distintas clases, con ciertas restricciones lingüísticas o propias del ámbito temático.

En el modelo conceptual dinámico se consideran, en primer lugar, las interacciones entre conceptos, y a continuación se analizan los conceptos para descomponerlos en sus elementos constituyentes. Para llevar a cabo estos dos procesos es necesario:

- a) Expandir el conjunto de los conceptos de relación para poder dar cuenta de la interacción dinámica que se produce entre entidades y procesos.
- b) Permitir la desconstrucción o fragmentación de las entidades, actividades y relaciones existentes. Esto puede hacerse de dos modos: por medio de la subdivisión de conceptos en entidades más pequeñas, o bien considerando que los conceptos poseen rasgos o características, denominados conceptos de propiedades.

Así, los conceptos de propiedades permiten descomponer y recombinar las distintas clases de conceptos (entidades, relaciones y actividades), y a su vez

constituyen una clase conceptual. En un primer nivel, las propiedades se consideran sólo como elementos constituyentes de otros conceptos, mientras que en un nivel superior se consideran como una clase de conceptos.

A continuación presentamos la definición de cada clase conceptual en el modelo conceptual dinámico:

- a) **Entidad:** Clase de conceptos que se obtiene a partir de la abstracción de elementos de nuestra experiencia y reflexión cuya existencia se considera independiente en el espacio y en el tiempo. Las entidades son elementos definibles por separado, necesarios para la identificación y clasificación de las unidades de la experiencia y el conocimiento. Dentro de las entidades podemos distinguir entre *entidades individuales* (designadas por nombres propios) y *entidades de clase* (designadas mediante nombres comunes). La mayoría de los ejemplos observados se consideran elementos de una clase a la que se otorga un nombre común (Sager y Kageura, 1994: 196). Para identificar las clases nos fijamos en las unidades convencionales de observación, que a su vez pueden ser tipos o partes de unidades mayores.
- b) **Relación:** Clase de conceptos obtenida a partir de la abstracción de relaciones físicas, temporales u otro tipo de relaciones ontológicas entre objetos y a partir de las relaciones lógicas establecidas entre entidades, relaciones y actividades (por ejemplo *diferencia, relación*). Las relaciones son la clase de conceptos que identifica las relaciones existentes o establecidas entre dos o más entidades, actividades o propiedades, o bien entre una combinación de ellos. Al clasificar las distintas entidades en unidades mayores o menores construimos un sistema estático que consiste en una red de nodos unidos por arcos que los relacionan entre sí.
- c) **Actividad:** Clase de conceptos obtenida mediante la abstracción de procesos, operaciones o acontecimientos realizados por o con entidades. Su estructura es más compleja que la de las entidades porque requieren la participación directa de entidades para llevarse a cabo.

- d) **Propiedad:** Clase de conceptos derivada del análisis de los componentes y características de entidades, actividades y relaciones. Las propiedades siempre se consideran en un primer nivel de abstracción asociadas a otros conceptos, y sólo en un segundo nivel se constituyen como conceptos independientes. Las propiedades permiten identificar y describir las diferencias existentes entre entidades y actividades, así como identificar rasgos o características de entidades, actividades o relaciones.

Sager y Kageura (1994) dividen cada clase de conceptos en subcategorías. En un primer nivel de análisis, los subtipos identificados parecen tener validez general, pero en niveles más profundos, la subdivisión está condicionada por el ámbito específico que objeto de estudio, que en el caso de Sager y Kageura es la documentación. A continuación ofrecemos el primer nivel de subdivisión de las clases conceptuales propuesto por Sager y Kageura (1994), indicando los ejemplos que identifican en el ámbito de la Documentación:

a) **Tipos de conceptos de entidad**

- a. **Entidad material:** Subclase de los conceptos de entidad que se deriva directamente de los objetos con sustancia física. Las entidades materiales se subdividen en:

i. **entidades materiales animadas:**

- personas: *persona, autor, lector*
- organizaciones: *empresa, departamento, universidad*

ii. **entidades materiales inanimadas:**

- objetos físicos: *diccionario, periódico*
- lugares: *ciudad, estantería*

- b. **Entidad no material:** Entidades que no hacen referencia a un objeto con sustancia física. Se subdividen en:

- i. **entidades clasificadoras:** Subclase de los conceptos de entidades que no se derivan de objetos físicos y que sirven para ordenar y clasificar otros conceptos. Ejemplos de entidades

clasificadoras son *elemento, organización, sistema*.

ii. **entidades abstractas:** Subclase de conceptos de entidades que no se derivan de objetos físicos ni de fenómenos físicamente observables y que representan unidades de conocimiento abstractas, tanto en general como en ámbitos temáticos concretos. Ejemplos de entidades abstractas serían conceptos como *teoría, idea, ecuación, conocimiento*.

iii. **entidades específicas de un ámbito temático:** Cada ámbito temático requiere categorías especiales para sus abstracciones. Algunos ejemplos propios de la Documentación son *artículo, cita, palabra*.

b) **Tipos de conceptos de actividad:** Según Sager y Kageura, esta es la clase más compleja porque presupone relaciones con otros conceptos.

a. Conceptos de acción: Aquellos en los que la actividad se centra en quien la lleva a cabo. Algunos ejemplos de conceptos de acción son *respuesta, petición, búsqueda*.

b. Conceptos de transferencia: Aquellos en los que el foco se centra en la transferencia del objeto con un cambio en su estado externo. Algunos ejemplos de conceptos de este tipo en el ámbito de la Documentación son *recepción o préstamo*.

c. Conceptos de cambio de estado: Conceptos en los que se produce un cambio en el estado interno o externo del objeto. Algunos ejemplos de conceptos de cambio de estado son *prohibición o adaptación*.

d. Conceptos de producción: Conceptos que producen o reproducen entidades. Algunos ejemplos son *diseño o output*.

c) **Tipos de conceptos de relación:** Sager y Kageura afirman que su tratamiento de los conceptos de relación es incompleto porque no considera las relaciones entre conceptos expresadas mediante compuestos.

a. Etiquetas para tipos de relación

- i. General:** *combinación, posición, orden*
- ii. Específica de un ámbito temático:** *paridad, nivel*
- iii. Medida específica de un ámbito temático:** *entropía, sensibilidad*

b. Valores de las relaciones

i. Físicas:

- Espaciales: *proximidad, contigüidad*
- Temporales: *simultaneidad, prioridad*
- Funcionales: *instrumentalidad, influencia*

ii. No físicas:

- Modo de la relación: *jerarquía, disyunción*
- Comparación: *compatibilidad, equivalencia*
- Estatus relativo: *subsidiariedad, incidentalidad*

d) **Tipos de conceptos de propiedades:** Los conceptos de propiedades siempre se consideran asociados a una entidad.

a. Etiquetas de los tipos de propiedades

- i. General:** *estado, estilo, naturaleza*
- ii. Específicas del ámbito temático:** *alcance, condición, forma*
- iii. Medidas específicas del ámbito temático:** *peso, dimensión*

b. Valores de las propiedades

i. Físicas:

- Espaciales: *universalidad*
- Temporales: *periodicidad*
- Funcionales/situacionales: *conveniencia*
- Forma: *angularidad*
- Color: *transparencia*
- Cantidad: *mitad*
- Escala: *macro, micro*

- Otras: *visibilidad*

ii. No físicas:

- General/especial: *universalidad*
- Natural/artificial: *realidad*
- Simple/compleja: *complejidad, polisemia*
- Negativa: *no-, in-*
- Modo: *arbitrariedad, pasividad*
- Representativas del ámbito temático: *análogo, denotación*

Los valores de las propiedades dependen siempre de otras clases de conceptos (entidades y actividades).

2.2.1.2 Relaciones conceptuales en el enfoque funcionalista

Sager (1990) considera que entre los distintos grupos de conceptos pueden establecerse relaciones de diversa naturaleza, si bien se seleccionan únicamente los tipos de relaciones que nos parecen más relevantes en función de las necesidades de un trabajo concreto. Por tanto, los sistemas conceptuales en terminología no tienen validez absoluta, sino que se crean para un fin específico.

En opinión de Sager, las relaciones desempeñan un papel en el proceso de formación de conceptos, que se apoya en la selección de relaciones concretas entre las características de los conceptos y también indirectamente entre conceptos e incluso entre objetos (Sager, 1990/1993: 55).

Sager distingue cuatro tipos de relaciones, que son las **relaciones genéricas** (dentro de las cuales identifica la **relaciones cuasigenéricas**), las **relaciones partitivas**, las **relaciones polivalentes** y las **relaciones complejas**.

Relaciones genéricas

La concepción de las relaciones genéricas de Sager coincide a grandes rasgos con las relaciones lógicas consideradas en la TGT (ver apartado 2.2.4.1). Este tipo de relación identifica los conceptos que pertenecen a la misma categoría, en la que existe un concepto más amplio (denominado **concepto genérico**), del que se dice que está superordinado a un concepto o conceptos subordinados más restringidos (denominados **conceptos específicos**). Todos los conceptos subordinados a un mismo superordinado poseen las mismas características que éste, más alguna característica diferenciadora que permite distinguirlos tanto del

superordinado común como entre sí. Por ello la relación genérica implica tanto una relación vertical como una relación horizontal.

En los niveles inferiores, la extensión del concepto queda limitada, es decir, el concepto se refiere a un menor número de objetos de la realidad, y al mismo tiempo, la intensión se hace más compleja.

Dentro de las relaciones genéricas, Sager distingue las denominadas **relaciones cuasigenéricas** (1990/1993: 58), que son aquellas en las cuales la atribución a una clase no es permanente. Por ejemplo, un *diente de león* es siempre un miembro de la familia de las *Compositae* (relación genérica), pero también puede ser considerado como una hierba, una planta medicinal o una hortaliza (relación cuasigenérica).

Relaciones partitivas

Las relaciones partitivas, que Sager denomina también *relaciones de todo-parte*, indican la conexión entre los conceptos que constan de más de una parte y los conceptos que se refieren a las partes (1990/1993: 59).

Sager recoge los distintos tipos de relación partitiva propuestos por Hann (1978), en función del número y la clase de las partes:

a) Las partes son componentes atómicos del todo; es decir, el todo consta de muchas de las mismas partes. Un ejemplo son las unidades de una escala.

b) Las partes son un conjunto numerado finito. Un ejemplo son las 52 cartas de la baraja de póker.

c) El todo consta de varios grupos de partes numeradas y no numeradas. Un ejemplo son los valores individuales de la baraja.

d) La parte o partes son un componente opcional del todo. Por ejemplo, el equipo de música en un coche.

e) La parte es un componente de un todo, pero puede ser un todo en sí mismo. Por ejemplo, el monitor de un ordenador es un todo en sí mismo, pero al mismo tiempo es un componente de un ordenador.

f) La parte o partes son alternativas, es decir, el concepto incluyente puede tener un concepto específico u otro, pero no los dos simultáneamente. Por ejemplo, un coche puede tener un motor de gasolina o un motor diésel, pero no ambos a la vez.

Relaciones polivalentes

Sager considera preciso evaluar todas las relaciones que un concepto establece con otros conceptos, y no asumir que es suficiente con identificar una única relación. Las relaciones polivalentes se establecen cuando un concepto se coloca en más de una jerarquía dentro de un campo temático específico. Por ejemplo, el concepto *autobús* puede clasificarse como un tipo de *vehículo de pasajeros* o como un tipo de *vehículo terrestre*.

Relaciones complejas

Sager destaca la importancia de las relaciones complejas como reveladoras de la naturaleza de los conceptos en los lenguajes de especialidad. Estas relaciones se corresponden a grandes rasgos con las que en la TGT se denominan no jerárquicas.

A continuación presentamos la lista de relaciones complejas propuesta por Sager, incluyendo los ejemplos de cada relación que se indican en Sager (1990/1993: 62-63). Se trata de una lista amplia, si bien el autor considera que puede ser ampliada en función de las necesidades:

- a) **Relación causa-efecto:** *explosión-lluvia ácida*
- b) **Relación material-producto:** *acero-viga*
- c) **Relación material-propiedad:** *crystal-frágil*
- d) **Relación material-estado:** *hierro-corrosión*
- e) **Relación proceso-producto:** *tejer-tejido*
- f) **Relación proceso-instrumento:** *incisión-escalpelo*
- g) **Relación proceso-método:** *almacenamiento-lio-filación*
- h) **Relación proceso-paciente:** *teñir-textil*
- i) **Relación fenómeno-medida:** *luz-watio*
- j) **Relación objeto-contragente:** *veneno-antídoto*
- k) **Relación objeto-contenedor:** *herramienta-caja de herramientas*
- l) **Relación objeto-material:** *punto-hierro*
- m) **Relación objeto-cualidad:** *gasolina-supercarburante*
- n) **Relación objeto-operación:** *boca de taladro-taladración*
- o) **Relación objeto-característica:** *combustible-sin humo*
- p) **Relación objeto-forma:** *libro-rústica*
- q) **Relación actividad-lugar:** *explotación-mina de carbón*

Propone crear subconjuntos de estas relaciones en función de la clase de conceptos involucrados (entidades, propiedades, actividades o relaciones), y de

nuevo insiste en que las clases conceptuales de los conceptos determinan la naturaleza de las relaciones posibles entre ellos (Sager, 1990/1993: 63).

Según Sager, al establecer las relaciones puede indicarse la naturaleza del criterio que se ha seguido para la subdivisión. Por ejemplo, las impresoras se pueden clasificar en función del modo en que opera el mecanismo de impresión en *impresora láser, impresora matricial e impresora de margarita*. Esto es lo que se denomina **clasificación por facetas**, que permite agrupar los términos en torno a una característica compartida. Las facetas permiten reflejar la multidimensionalidad en los sistemas de conceptos.

Por último, Sager recuerda que, además de relacionar los conceptos entre sí, en ocasiones surge la necesidad de agrupar conceptos dentro de los campos temáticos, tal y como se hace en los sistemas de clasificación habituales en Documentación, como la Clasificación Decimal Universal. Esta manera de relacionar los conceptos se basa en un enfoque de “arriba a abajo”, que permite disponer los conceptos en subconjuntos de conocimiento, si bien no refleja las relaciones que existen entre los términos individuales. A medio camino entre la clasificación y las relaciones conceptuales Sager sitúa los tesauros, en los que se lleva a cabo una división en subcategorías de los campos temáticos mediante encabezamientos por clases de materias.

2.2.1.3 La dinámica de la terminología (Kageura, 2002)

Dentro de la corriente funcionalista, Kageura (2002) nos ofrece una visión dinámica de la terminología. En su estudio *The Dynamics of Terminology*, Kageura se cuestiona si la formación de términos y el crecimiento terminológico responden a algún tipo de sistematicidad, es decir, si se pueden observar patrones motivados en la construcción de términos dentro de un dominio concreto, en su caso el dominio de la recuperación de la información (*information retrieval*). Para demostrar que existe dicha sistematicidad, Kageura analiza la potencialidad dinámica para crear nuevos términos que se observa en una porción sincrónica de la estructura interna de una terminología determinada (Kageura, 2002: 34).

Kageura (2002) lleva a cabo el estudio de los patrones de formación de términos y de crecimiento terminológico en el ámbito temático de la Documentación. En su opinión, si queremos ofrecer una teoría de la terminología, es necesario estudiar la terminología de un dominio concreto en su totalidad, y no

únicamente como fuente de ejemplos que vengan a apoyar la teoría. En su opinión, un buen estudio descriptivo constituye una teoría de los términos.

[...] a theory of terms or terminology should deal with the terminology of a domain in its totality, because it is only with respect to individual domains that the very concept of «term» is consolidated.

(Kageura, 2002: 2)

Kageura (2002: 253) se plantea el análisis desde dos perspectivas. Por un lado, describe los patrones de formación de términos desde un punto de vista conceptual, y por otro, lleva a cabo un estudio cuantitativo.

En cuanto al análisis conceptual, Kageura adopta un enfoque tradicional de la naturaleza de los conceptos, y en este sentido define las categorías conceptuales como elementos que pueden delimitarse claramente, a los que se pueden asignar elementos léxicos sin caer en la ambigüedad. No por ello deja de ser consciente de las nuevas corrientes en terminología, que abogan por una perspectiva más flexible (como Cabré, 1999 y Temmerman, 2000), pero decide centrarse en el aspecto sistemático de la dinámica de la terminología frente a la estructura general de los términos existentes (Kageura, 2002: 253). Su decisión de estudiar la terminología de un dominio en su totalidad hace necesario establecer un sistema conceptual general ajustado a la visión tradicional de los conceptos y las categorías conceptuales, como punto de partida para cualquier estudio de la terminología. Posteriormente podrán tenerse en cuenta aspectos relativos a la flexibilidad o la prototipicidad de los términos.

Kageura emplea las categorías conceptuales definidas dentro del sistema conceptual de dos modos: en primer lugar, al atribuir las categorías a los términos, emplea las clases conceptuales para observar la sistematicidad de la formación de términos. En segundo lugar, al atribuir categorías conceptuales a los morfemas, utiliza las clases conceptuales para observar la regularidad de los patrones combinatorios.

La primera tarea para caracterizar la dinámica de la terminología es descubrir y describir las regularidades conceptuales en los términos existentes. Para ello es necesario observar las relaciones existentes entre los términos y entre los constituyentes de los mismos, así como las relaciones que se dan entre los constituyentes y el tipo de combinaciones conceptuales empleadas en la

construcción de términos (Kageura, 2002: 36). Kageura asume que el rasgo principal que determina la sistematicidad de los términos es el sistema conceptual de un dominio. Las regularidades de la construcción de la terminología se reflejan en las relaciones que se establecen entre los términos y entre sus elementos constituyentes, que Kageura organiza en forma de patrones de especificación conceptual (*conceptual specification patterns*). Estos patrones tienen una utilidad tanto práctica como teórica. Desde un punto de vista empírico, se trata de relaciones establecidas en el interior de los términos. Desde una perspectiva teórica, constituyen un mecanismo para incorporar un aspecto clasificador de la formación de términos (Kageura, 2002: 254).

Kageura agrupa los patrones de especificación, y por tanto las relaciones intraconceptuales, según seis puntos de vista (igualdad, punto de vista funcional, punto de vista parte/todo, punto de vista de los atributos internos o externos y otros), que dan lugar a los siguientes tipos de relación⁹:

- a) **Especificación desde el punto de vista de la igualdad:**
 - i. **Yuxtaposición:** *reader printer*
- b) **Especificación desde el punto de vista funcional**
 - i. **Especificación del rol:** *sample copy*
 - ii. **Especificación de los elementos complementarios (objeto afectado, producto, instrumento y manera):**
data processing, index making, questionnaire survey, multiple access
 - iii. **Especificación de la función:** *editing person*
 - iv. **Especificación del uso:** *expression criterion*
 - v. **Especificación de la finalidad:** *computer program*
- c) **Especificación desde un punto de vista parte/todo**
 - i. **Especificación del todo:** *vocabulary primitive*
 - ii. **Especificación de los constituyentes:** *card catalogue*
 - iii. **Especificación de contenido informativo y representación:** *term list*

⁹ Empleamos aquí la traducción de los patrones de especificación conceptual propuesta en la revisión de Oster (2005: 31-33).

- d) **Especificación desde el punto de vista de los atributos internos**
 - i. **Especificación de los atributos formales:** *block diagram*
 - ii. **Especificación de manera estática:** *equivalence relation*
 - iii. **Especificación de naturaleza:** *natural language*
 - iv. **Especificación del concepto al que se atribuye una propiedad:** *noise ratio.*
 - v. **Especificación de la cantidad:** *two language dictionary*
- e) **Especificación desde un punto de vista relacional/externo**
 - i. **Especificación de estatus:** *basic entry*
 - ii. **Especificación de lugar:** *edge hole*
 - iii. **Especificación de tiempo:** *pre-coordination*
 - iv. **Especificación de origen:** *derived word*
 - v. **Especificación de alcance:** *universal bibliography*
 - vi. **Especificación del agente:** *user needs*
- f) **Otros:**
 - i. **Especificación de diferenciación:** *Venn diagram*

El análisis de los patrones conceptuales que lleva a cabo es general, por lo que se pierde granularidad en la descripción, es decir, se pierde precisión. Para suplir esta falta de granularidad, Kageura afronta la descripción de los patrones cuantitativos de aparición de elementos léxicos o conceptuales en su conjunto, que completarán el análisis conceptual. Adopta el método de interpolación y extrapolación binómica (*binomial interpolation and extrapolation*), que permite calcular el cambio en el número de elementos relevantes para la construcción de terminología más allá del tamaño de los datos actuales. El enfoque cuantitativo revela los patrones típicos de crecimiento, superando la insuficiencia del enfoque puramente conceptual.

Kageura concluye que la terminología del dominio de la Documentación tiende hacia la sistematización de su estructura interna, que se refleja tanto en la

estabilidad de las relaciones establecidas entre elementos léxicos, como en las formas lingüísticas de los términos.

Valoración

El enfoque funcionalista y dinámico de la terminología abre nuevas y productivas vías para el estudio de las relaciones conceptuales. Por una parte, la consideración de las clases conceptuales nos permite describir las relaciones conceptuales de un modo más preciso, pudiendo especificar entre qué clases de conceptos se establece cada relación en particular. Nos apoyaremos en la tipología de clases conceptuales propuesta por Sager y Kageura (1994) para establecer restricciones en las relaciones conceptuales en nuestro trabajo.

La clasificación de las relaciones conceptuales ofrecida por Sager (1990) nos servirá también como punto de partida para la elaboración de nuestro catálogo de relaciones en la parte empírica del estudio.

La visión dinámica de la formación de términos propuesta por Kageura (2002) permite concebir las relaciones entre conceptos como un proceso que posibilita la evolución del conocimiento y la denominación de conceptos nuevos con los recursos limitados de la lengua. No obstante, al observar los patrones de especificación conceptual vemos que en ellos se combinan relaciones generales válidas para cualquier campo con relaciones muy específicas, que sólo tienen validez en un campo concreto. Otro aspecto interesante de la propuesta de Kageura es el análisis cuantitativo de los patrones de especificación conceptual, que completa el estudio de su naturaleza. Sin embargo, creemos que este análisis cuantitativo puede estar demasiado sesgado en función del ámbito especializado objeto de estudio, por lo que las conclusiones a las que Kageura llega para el ámbito de la Documentación podrían no ser válidas para otras áreas de conocimiento.

2.2.2 Teoría Comunicativa de la Terminología

La Teoría Comunicativa de la Terminología, desarrollada por Cabré (1999), supone una revisión profunda de la Teoría General que acerca la Terminología a la Lingüística. Si bien Cabré reconoce la adecuación de los principios de la Teoría General de la Terminología (TGT) para las finalidades que en su origen perseguía su actividad terminológica –principalmente la normalización conceptual y denominativa de los términos–, mantiene que los

avances de la vertiente aplicada de la disciplina y a la aparición de nuevas necesidades terminológicas hacen que la TGT resulte inadecuada:

[...] la teoría da muestras de reduccionismo en diferentes aspectos de sus concepciones y de sus prácticas: la concepción global de la unidad terminológica, la reducción de la unidad terminológica a su condición denominativa, el olvido de los aspectos sintácticos de las unidades terminológicas, la ignorancia de los aspectos comunicativos de los términos, o la insistencia en negar la variación formal y conceptual de los términos.

(Cabré, 1999: 69)

Cabré critica que la homogeneidad postulada por la teoría no se ve refrendada por los resultados del análisis empírico en los diversos ámbitos temáticos, en los que se demuestra la existencia de variación formal y conceptual en los términos, derivada de su naturaleza lingüística.

Desde el convencimiento de que la terminología es una materia interdisciplinar, y como tal está sujeta a distintos enfoques, y teniendo en cuenta la evolución social y la diversidad de situaciones en las que se desenvuelve la actividad terminológica, Cabré propone un nuevo paradigma con el que superar las insuficiencias de la TGT.

La teoría concebida por Wüster y ampliada por sus sucesores de la Escuela de Viena centra su estudio en “el concepto y las relaciones conceptuales como unidades de análisis para llegar a las denominaciones de los conceptos establecidos” (Cabré 1999: 73), es decir, defiende un enfoque onomasiológico que diferencia a la terminología de su disciplina hermana, la lexicografía. La propuesta de Cabré pretende aclarar los aspectos como el objeto y la óptica de estudio, que a su entender no resuelve la TGT, y a continuación propone tres principios y tres condiciones para una teoría comunicativa de la terminología.

En primer lugar, el objeto de estudio de la Terminología para Cabré está constituido por las unidades terminológicas en su dimensión funcional, formal y semántica, teniendo en cuenta su doble sistematicidad. Esta doble sistematicidad se da, por una parte, con respecto al sistema de la lengua, y por otra, en relación con el área de especialidad en que se emplean.

En segundo lugar, Cabré trata de delimitar la óptica del estudio. En este sentido, dado que las unidades terminológicas presentan al menos tres aspectos (lingüístico, cognitivo y comunicativo), será necesario integrar las distintas

“facetas del poliedro en una unidad que sirva para la comunicación especializada” (Cabré, 1999: 84).

Por último, una vez fijados el objeto y la óptica del estudio, Cabré propone tres principios, inherentes a la unidad terminológica, y tres condiciones (posibilidades alternativas que los textos pueden poseer o no). En el siguiente cuadro se resumen estos seis principios y condiciones, que explicamos más adelante (Cabré, 1999: 85-89):

PRINCIPIOS Y CONDICIONES DE LA TEORÍA COMUNICATIVA DE LA TERMINOLOGÍA	
<i>Principio sobre la poliedricidad del término</i>	Las unidades terminológicas son inherentemente poliédricas
<i>Principio sobre el carácter comunicativo de la terminología</i>	Toda unidad terminológica persigue una finalidad comunicativa
<i>Principio sobre la variación</i>	La comunicación especializada comporta variación en diversos grados
<i>Condición de lenguaje natural</i>	Los términos se ven sujetos a las propiedades morfológicas, sintácticas y semánticas generales de las palabras de una lengua
<i>Condición de comunicación especializada</i>	La comunicación especializada se diferencia de la general en la selección y el tratamiento de las unidades léxicas y los recursos textuales
<i>Condición de especialización</i>	La condición de comunicación especializada admite diversidad de niveles

Tabla 4. Principios y condiciones de la Teoría Comunicativa de la Terminología.

Principio sobre la poliedricidad del término

Las unidades terminológicas son inherentemente poliédricas, es decir, integran al mismo tiempo aspectos lingüísticos, cognitivos y sociales.

Principio sobre el carácter comunicativo de la terminología

Toda unidad terminológica persigue una finalidad comunicativa, ya sea directa, ya sea a través de la representación conceptual conducente a un mejor entendimiento entre expertos.

Principio sobre la variación

Todo proceso de comunicación comporta variación, explicitada en formas alternativas de denominación del mismo concepto (sinonimia) o en apertura significativa de una misma forma (polisemia). Este principio admite diferentes grados en función de las condiciones de la situación comunicativa.

Condición de lenguaje natural

La concepción de la terminología debe situarse en el marco de una determinada esfera de conocimiento, en la que se define su objeto. Es la propia situación en un área de conocimiento la que perfila el campo de estudio de la terminología.

La condición de lenguaje natural proyecta una serie de consecuencias propias de la gramática: los términos se ven sujetos a las propiedades morfológicas, sintácticas y semánticas generales de las palabras de una lengua.

Condición de comunicación especializada

La comunicación especializada cumple los requisitos semánticos y pragmáticos, es decir, trata un tema especializado desde una perspectiva especializada, se produce en situaciones profesionales, pertenece al registro formal y adquiere su significado desde una estructura preestablecida. Comparte con la comunicación general las reglas y los procesos lingüísticos y textuales, y se diferencia de ella en la selección de unidades y en la frecuencia con que cada recurso es actualizado en el discurso.

Condición de especialización

La condición de comunicación especializada admite diversidad de niveles o grados de especialidad, y dicho grado de especialización condiciona tanto la densidad terminológica de un texto como la cantidad de variación expresiva para hacer referencia a un mismo concepto.

A medida que disminuye el grado de especialización, el discurso va adquiriendo características que lo acercan al discurso no especializado: en el *plano semántico*, se produce variación conceptual, redundancia, ambigüedad, falta de precisión estricta; en el *plano formal*, se aprecia variación denominativa y uso elevado de fórmulas parafrásticas para explicar el concepto.

En definitiva, Cabré propone (1999: 89) una teoría de base lingüística que da cuenta de la terminología como un conjunto de unidades denominativo-conceptuales del lenguaje natural, que representan el conocimiento especializado dentro de una área temática precisa y que vehiculan la comunicación profesional.

Esta revisión de los postulados de la disciplina terminológica tiene consecuencias metodológicas tanto para la práctica como para los principios terminográficos. En cuanto a estos últimos, al considerar los conceptos como unidades sujetas a variación según ciertos parámetros, este enfoque se aleja de la orientación onomasiológica, que, si bien se sigue considerando predominante,

deja lugar para otros enfoques. En cuanto a la práctica terminográfica, adquiere un papel indispensable el concepto de adecuación, puesto que “cada trabajo en concreto adopta una estrategia en función de su temática, objetivos, contexto, elementos implicados y recursos disponibles” (Cabré, 1999: 137).

Si bien la teoría planteada por Cabré no hace referencia explícita a las relaciones conceptuales, en el marco de su paradigma terminológico, Feliu (2000 y 2004) desarrolla una completa propuesta de clasificación de las relaciones conceptuales que describimos detalladamente en el siguiente apartado.

2.2.2.1 Catálogo de relaciones conceptuales de Feliu (2000 y 2004)

En el contexto de la Teoría Comunicativa de la Terminología, Feliu (2000 y 2004) lleva a cabo un estudio de las relaciones conceptuales establecidas entre los términos cuyo objetivo es servir de base para la extracción de información sobre relaciones a partir de un corpus especializado.

Apoyándose en el estudio de los tipos de relaciones conceptuales propuestos desde la teoría de la terminología y la Semántica Léxica, y basándose en la diversidad de configuraciones posibles para cada tipo de relación, sugiere que una teoría adecuada de las relaciones conceptuales debe cumplir los siguientes requisitos (Feliu, 2000: 36-37):

- a) Debe aportar criterios para identificar los distintos tipos de relaciones a partir de los **elementos lingüísticos y conceptuales**.
- b) Debe contener criterios para **identificar** y **describir** las diferentes configuraciones dentro de cada relación. Estos criterios deberían tener en cuenta la presencia o no de determinadas **propiedades**, y a partir de su presencia, determinar el grado de **relevancia** de cada propiedad.
- c) Debe proporcionar parámetros para **comparar** los diversos tipos de relación y configuraciones, es decir, para determinar cuáles son semejantes entre sí y cómo se pueden distinguir unas de otras.
- d) Debe indicar qué elementos permiten establecer el **grado de pertinencia** de las propiedades de cada configuración a pesar de la diversidad de configuraciones.
- e) Debe describir vías de formación de **nuevas configuraciones**.

Feliu y Cabré (2002) conciben las relaciones como un nexo semántico, es decir, como un predicado particular entre dos o más unidades de conocimiento especializado. Esta definición se formaliza en la fórmula

$$\mathbf{a R b, n}$$

donde **R** representa a la relación, **a** y **b** son las clases conceptuales relacionadas y **n** permite reflejar la recursividad de algunas de las relaciones, es decir, la posibilidad de que **R** se establezca entre un elemento **a** y más de dos elementos situados en el lado derecho de la ecuación. Esta ecuación de lo que se entiende por relación está basada en la propuesta de Otman (1996) que veremos en el apartado 3.6.

El resultado de su estudio sobre los distintos tipos de relaciones conceptuales se resume en el catálogo que propone Feliu (2004: 30-34), en el que para cada relación se incluye información sobre sus propiedades matemáticas, las clases conceptuales implicadas (según lo propuesto por Sager y Kageura, 1994), el número de argumentos oracionales y la expresión de la relación por medios lingüísticos en catalán.

A continuación resumimos su propuesta, incluyendo los tipos y subtipos de relación que se identifican:

1. **Relación de semejanza:** Relación que se establece por equivalencia u oposición entre dos o más conceptos. Se trata de una relación simétrica y transitiva, que presenta dos subtipos:

a) Semejanza positiva:

- Equivalencia total o sinonimia
- Equivalencia parcial o semejanza

b) Semejanza negativa:

- Total u oposición
- Parcial o contraste

Esta relación se establece entre dos o más conceptos y puede presentar recursividad. Las clases conceptuales relacionadas son entidades y actividades, y los conceptos situados a uno y otro lado de la relación deben pertenecer a la misma clase conceptual.

2. **Relación de inclusión:** Relación que se establece por la inclusión de algunas características de un determinado elemento en otro elemento. Es una relación no simétrica y transitiva. Presenta un único subtipo, la **relación de clase**

o **de hiponimia**. Como en el caso anterior, esta relación se produce entre dos o más conceptos, por lo que puede presentar recursividad. Las clases relacionadas son entidades y actividades, y necesariamente los elementos situados a ambos lados de la relación deben pertenecer a la misma clase conceptual.

3. Relación de secuencialidad: Relación que se establece por localización o sucesión en el espacio o en el tiempo de los conceptos que une. Se trata de una relación no simétrica y transitiva (según el subtipo). Presenta dos subtipos:

- a) Secuencialidad espacial:
 - i. De localización
 - ii. De dirección
- b) Secuencialidad temporal:
 - iii. Simultaneidad
 - iv. Anterioridad-posterioridad

Las clases conceptuales relacionadas pueden ser entidades y actividades pero, a diferencia de las dos relaciones anteriores, en este caso los dos elementos relacionados no tienen por qué pertenecer a la misma clase. Es una relación entre dos o más argumentos oracionales (según el subtipo) y puede presentar recursividad.

4. Relación de causalidad: Relación que se establece entre una causa y su efecto. Se trata de una relación no simétrica y no necesariamente transitiva. Presenta dos subtipos:

- a) **Relación causal:** causa-efecto
- b) **Relación procedural:** proceso-resultado
 - v. Proceso-resultado con afectación
 - vi. Proceso-resultado sin afectación

5. Relación instrumental: Relación que se establece entre un instrumento y su función. Se trata de una relación no simétrica y no transitiva que no presenta subtipos. Se produce entre dos conceptos y no presenta recursividad. Las clases relacionadas son entidades y actividades, pero no es necesario que los dos elementos de la relación pertenezcan a la misma clase conceptual.

6. Relación de meronimia: Relación que se establece entre un elemento que constituye un todo y los elementos que conforman sus partes. Este tipo de relación no es simétrica, y tampoco necesariamente transitiva. Feliu identifica un

único subtipo de relación meronímica, la **parte-todo**, en la que se identifican siete subtipos en función de las unidades relacionadas.

Las relaciones meronímicas se producen entre dos o más elementos, y puede presentar recursividad. Las clases conceptuales relacionadas pueden ser entidades, actividades y propiedades. Los conceptos relacionados no deben pertenecer necesariamente a la misma clase, pero Feliu afirma que una propiedad tiene que estar necesariamente asociada a una entidad.

En función del tipo de unidades que se relacionan, y basándose en la clasificación de Winston *et al* (1987), Feliu distingue siete tipos de relación parte-todo:

- a) Relación componente-objeto
- b) Relación miembro-colección
- c) Relación porción-masa
- d) Relación material-objeto
- e) Relación etapa-proceso
- f) Relación característica-actividad
- g) Relación lugar-área

7. Relación de asociación: Relación que se establece por la correlación entre dos o más elementos. Se trata de una relación no necesariamente simétrica ni transitiva, que se establece entre dos o más argumentos oracionales, pudiendo presentar recursividad. Se produce entre entidades y actividades, y los conceptos relacionados no tienen por qué pertenecer a la misma clase conceptual.

Feliu observa la expresión léxica de las relaciones, y ofrece un catálogo de marcadores léxicos en lengua catalana mediante los cuales se expresa cada relación.

A continuación se reproduce un cuadro que resume el catálogo de relaciones de Feliu, basado en Feliu y Cabré (2002):

Tipo	Subtipo	Propiedades ¹⁰
Semejanza	Positiva:	
	- Equivalencia total o sinonimia	+ ⇔
	- Equivalencia parcial o semejanza	+ ⇒
	Negativa:	
- Total u oposición	+ ⇔	
- Parcial o contraste	+ ⇒	
Inclusión	Inclusión de clase o hiponimia	- ⇔ + ⇒
Secuencialidad	Lugar:	
	- Localización	- ⇔
	- Dirección	~ ⇒
	Tiempo:	
- Simultaneidad	- ⇔	
- Anterioridad-posterioridad	+ ⇒	
Causalidad	Causalidad:	
	- causa-efecto	- ⇔
	- proceso-resultado: con afectación	~ ⇒
- proceso-resultado: sin afectación		
Instrumento	Instrumento-función	- ⇔ ~ ⇒
Meronomia	Parte todo (<i>a</i> y <i>b</i> pueden ser):	- ⇔ ~ ⇒
	<i>Componente-objeto</i>	- ⇔ ~ ⇒
	<i>Miembro-colección</i>	- ⇔ ~ ⇒
	<i>Porción-masa</i>	- ⇔ ~ ⇒
	<i>Material-objeto</i>	- ⇔ ~ ⇒
	<i>Etapas-proceso</i>	- ⇔ ~ ⇒
	<i>Lugar-área</i>	- ⇔ ~ ⇒
		~ ⇔ ~ ⇒
Asociativa		~ ⇔ ~ ⇒

Tabla 5. Tipología de relaciones conceptuales (Feliu y Cabré, 2002).

Valoración

¹⁰ El símbolo ⇔ indica simetría, el símbolo ⇒ representa la transitividad de la relación. Los símbolos +/-/~ indican la presencia ausencia o no obligatoriedad de una propiedad en particular.

Esta clasificación presenta varios aspectos que consideramos positivos. En primer lugar, en su confección se han tenido en cuenta las tipologías propuestas tanto por la terminología como por la Semántica Léxica. En segundo lugar, considera el modelo conceptual de clases propuesto por Sager y Kageura (1994) y describe las clases conceptuales que pueden estar implicadas en cada relación. Por último, hace explícitas las propiedades matemáticas de simetría y transitividad de las relaciones, que desempeñan un papel importante a la hora de formalizar las relaciones conceptuales.

En cuanto a los aspectos mejorables de la tipología, echamos en falta la presencia de ejemplos concretos de los conceptos entre los que se establece cada relación, ya que se centra en la expresión de las relaciones mediante marcadores léxicos.

Esta propuesta de clasificación de las relaciones nos permitirá determinar la propia estructura de nuestro catálogo, al que añadiremos relaciones en función de las necesidades de nuestro estudio.

2.2.3 Enfoque sociocognitivo de la terminología (Temmerman, 2000)

Temmerman defiende un enfoque sociocognitivo de la terminología, basado en la observación inicial de que las categorías conceptuales poseen una naturaleza prototípica, es decir, no siempre pueden definirse a partir de sus características necesarias, como propone la Teoría General de la Terminología, y suficientes. Temmerman defiende que los conceptos no siempre tienen una estructura bien delimitada, sino que existen ejemplos más representativos (es decir, más prototípicos) que otros. Las unidades de comprensión (*units of understanding*) pueden describirse en función de unidades de información más o menos esenciales, dependiendo del tipo de unidad de comprensión de que se trate y de otros factores como la perspectiva desde la que se aborda la unidad. Así, propone su representación por medio de plantillas de comprensión (*templates of understanding*).

En su obra *Towards New Ways in Terminology Description*, Temmerman (2000) establece cinco principios para una Teoría Sociocognitiva de la Terminología, que se resumen en la siguiente tabla, presentada a continuación (Temmerman, 2000: 223):

PRINCIPIOS DE LA TEORÍA SOCIOCOGNITIVA DE LA TERMINOLOGÍA	
<i>Primer principio</i>	La Terminología sociocognitiva parte de las unidades de comprensión, que suelen poseer estructura prototípica
<i>Segundo principio</i>	La comprensión es un acto estructurado. Una unidad de comprensión tiene estructura intra- e intercategorial y funciona mediante modelos cognitivos
<i>Tercer principio</i>	Dependiendo del tipo de unidad de comprensión y del nivel y el tipo de especialización del emisor y el receptor de la comunicación, variará el tipo de información que se considere esencial para la definición
<i>Cuarto principio</i>	Las sinonimia y la polisemia son funcionales para el avance del conocimiento, y por tanto deben ser descritas
<i>Quinto principio</i>	Las unidades de comprensión están en constante evolución. Los modelos cognitivos desempeñan un papel en el desarrollo de las nuevas ideas, lo que implica que las denominaciones son motivadas

Tabla 6. Principios de la Teoría Sociocognitiva de la Terminología (Temmerman, 2000: 223).

El primer principio hace referencia la **naturaleza de las unidades terminológicas**, que Temmerman denomina *unidades de comprensión (units of understanding)*. La mayoría de estas unidades presentan **estructura prototípica**¹¹, es decir, no se corresponden con categorías claramente delimitables, sino que los contornos de los conceptos son difusos. En su propuesta distingue tres tipos de unidades de comprensión, a saber:

- **Categorías colectivas** (*Umbrella categories*): unidades de comprensión concebidas en la mente sobre la base de la experiencia. Ejemplo: *biotechnology (biotecnología)*.
- **Entidades**: Unidades de comprensión percibidas de manera objetiva. Ejemplo: *intron (intrón)*.
- **Actividades**: Unidades de comprensión que en un primer momento se conciben en la mente y posteriormente se materializan y desarrollan gracias a la experiencia. Ejemplo: *blotting (transferencia)*.

¹¹ Zawada y Swanopoele (1994) también defienden la naturaleza prototípica de los conceptos especializados. Su estudio se centra en la naturaleza, el papel y el estatus cognitivo de las características que definen los conceptos de la mineralogía.

Mientras que las entidades pueden definirse por medio de sus características necesarias y suficientes, las categorías colectivas y las actividades poseen una estructura prototípica y forman parte de marcos o modelos cognitivos idealizados (*Idealized Cognitive Models*, ICM)¹².

Esta afirmación nos lleva al segundo principio de la teoría sociocognitiva de la terminología, según el cual, la **comprensión es un hecho estructurado**, y por tanto una unidad de comprensión se estructura de forma intracategorial e intercategorial y funciona en el seno de modelos cognitivos.

Desde el punto de vista intracategorial se distinguen distintos módulos de información, como los rasgos (*features*), la definición nuclear (*core definition*), la información histórica o la información procedural, cuya relevancia depende del tipo de categoría de que se trate (entidad, actividad o categoría colectiva). Desde el punto de vista intercategorial se hace necesario desvelar el dominio y la intención del modelo cognitivo.

El tercer principio mantiene que la **definición de un concepto variará** según el tipo de unidad de comprensión y el nivel de especialización del emisor y del receptor de la definición. Temmerman distingue entre dos tipos de unidades de comprensión, *nociones* (entidades) y *categorías* (categorías colectivas y actividades). Las primeras se pueden describir en el marco de la TGT por medio de sus características necesarias y suficientes, pero para las segundas propone la sustitución de las definiciones tradicionales por plantillas (*templates*) que permitan describir los aspectos más borrosos y flexibles de la unidad de comprensión. Estas plantillas se componen de distintos módulos de información que recogen información intra- e intercategorial, que será más o menos esencial dependiendo del tipo de categoría y de la perspectiva desde la que se aborda la unidad.

A continuación reproducimos la plantilla propuesta por Temmerman para todos los tipos de categorías (entidades, actividades y categorías colectivas):

¹² El término *marco* se entiende aquí en el sentido propuesto desde la Lingüística por Fillmore (1985), mientras que el término *ICM* proviene de las teorías de Lakoff (1987).

Category/term: Type of category: entity activity collective category ...etc Core definition: Intracategorial information: is a part of consists of parts is a type of has the following types aims use application ...etc. Intercategorial information: perspectives domains intentions Historical information:

Fig. 1. Plantilla para la definición de las unidades de comprensión (Temmerman, 2000: 122).

En esta plantilla se puede apreciar que la información relativa a las relaciones conceptuales aparece en el apartado sobre información intercategorial. Temmerman se centra en las relaciones lógicas y ontológicas (partitivas), así como en el uso, los objetivos y la aplicación. Esta lista se plantea como una lista abierta, en la que pueden introducirse nuevos tipos de relación intercategorial en función de las necesidades de cada trabajo terminológico.

El cuarto principio postula que **el avance del conocimiento está ligado a la sinonimia y la polisemia**, y por consiguiente es necesario describir estos fenómenos. Para superar las limitaciones de la TGT, Temmerman defiende que los términos deben estudiarse en el seno del discurso y desde un punto de vista diacrónico, ya que, si bien las unidades claramente delimitadas tienden al isomorfismo, a la univocidad y a la monorreferencialidad, a medida que evoluciona su significado, aumenta la polisemia debido a cambios en la concepción, en la percepción y en los medios de expresión.

Frente a un enfoque tradicional de carácter puramente onomasiológico, el enfoque sociocognitivo propone que la evolución de las unidades de comprensión se estudie tanto desde una perspectiva onomasiológica como semasiológica.

Temmerman defiende que la polisemia es funcional en los lenguajes especializados: existe una constante discusión entre los especialistas que puede provocar que lo que es unívoco en un momento concreto pase a ser polisémico. Desde el punto de vista onomasiológico (que tiene en cuenta el proceso de lexicalización), la polisemia resulta de la reflexión humana sobre el mundo, pero desde el punto de vista semasiológico, la polisemia es el resultado de la evolución del significado.

Las categorías prototípicas explotan todo su potencial polisémico, por lo que cada nueva variante puede incorporarse a una categoría prototípica, elevando el número de variantes semánticas agrupadas alrededor de una estructura prototípica.

La (cuasi-)sinonimia, por su parte, aparece por la búsqueda de precisión al expresarse, relacionada con la perspectiva adoptada por el emisor del mensaje. La existencia de la sinonimia se justifica porque los mecanismos de denominación pueden desencadenar diferentes posibilidades de lexicalización. La sinonimia también es funcional, ya que contribuye a establecer distintos matices que ya estaban presentes cuando se desarrolló el concepto. Según Temmerman, el emisor, en tanto que usuario de la lengua, es quien decide sobre qué elemento o característica del concepto quiere poner el acento.

Esta funcionalidad que tienen tanto la polisemia como la sinonimia acarrea las siguientes consecuencias para el trabajo terminológico práctico. Por un lado, se hace imperativa una distinción entre las unidades de conocimiento bien delimitadas que tienden a la univocidad y las categorías de estructura prototípica, en las cuales existen polisemia y sinonimia, por lo que es necesario desarrollar métodos y principios para describirlas.

Por otro lado, la Terminología como disciplina científica debe abrirse a los métodos de investigación lingüísticos para la descripción del cambio diacrónico basados en el análisis de corpus. De este modo será posible estudiar el papel de la lengua en el proceso creativo del conocimiento.

El quinto y último principio del enfoque sociocognitivo de la terminología hace referencia a la **evolución permanente** a la que se encuentran sometidas las unidades de conocimiento. Los factores que determinan esta evolución son la necesidad de profundizar en la comprensión de las unidades, la interacción entre los distintos participantes en la comunicación y la estructura prototípica de las

unidades, que es a la vez causa y efecto de la evolución del significado. Por ello, Temmerman defiende el estudio diacrónico de las unidades terminológicas y mantiene que los modelos cognitivos, entre ellos los modelos metafóricos, desempeñan un papel fundamental en el desarrollo de nuevas ideas.

Estos cinco principios teóricos llevan a Temmerman a considerar que el trabajo de descripción terminológica que debería consistir en tres tipos de análisis: análisis de la estructura prototípica, análisis de los modelos cognitivos y análisis diacrónico.

Valoración

La propuesta de Temmerman, como la de Weissenhofer que vimos en el apartado 2.1.5.1, considera que algunas unidades de comprensión no pueden definirse mediante una serie de condiciones necesarias y suficientes. Su concepción tiene en cuenta los procesos cognitivos que se desarrollan en la evolución de la ciencia, aspecto que aporta una visión más realista de la terminología en la comunicación especializada.

En nuestro trabajo no vamos a tener en cuenta (al menos en principio) la posible prototipicidad de los conceptos; sin embargo, consideramos interesante la aportación de las plantillas para la definición de los conceptos.

Esta propuesta se sitúa un paso más allá de las definiciones realizadas en lenguaje natural, en el camino hacia la formalización de la información conceptual mediante el empleo de las plantillas de definición. Asimismo, nos parece muy acertada la consideración de que los conceptos forman parte de marcos o esquemas idealizados.

Otro aspecto interesante de este trabajo es la insistencia en el estudio de los términos empleados en el discurso como medio para estudiar la evolución de los conceptos. Nos sumamos a esta consideración, puesto que estimamos que los conceptos especializados deben estudiarse, en la medida de lo posible, en su hábitat natural, que es el discurso especializado.

2.3 Propiedades matemáticas de las relaciones conceptuales

Uno de los elementos que nos va a permitir definir formalmente las relaciones conceptuales son sus propiedades lógicas o matemáticas. Los autores que se han ocupado del estudio de las relaciones desde la Semántica Léxica (Lyons 1977; Cruse, 1986 y 2004; Evens, 1988) y algunos de los que las han

tratado desde la terminología (Otman, 1996; Feliu, 2000 y 2004; Oster, 2005) emplean para su descripción propiedades como la transitividad, la simetría o la reflexividad, si bien no todos ellos las definen con precisión.

Las propiedades de las relaciones tal y como las describimos en este apartado proceden de la lógica matemática, que considera las relaciones como objetos matemáticos, formados por pares (o series) de conjuntos que representan conceptos o palabras (dependiendo del enfoque). Si consideramos las relaciones en este sentido, podemos explotar sus propiedades matemáticas, descritas desde la lógica para su formalización (Evens, 1988: 24).

En este apartado nos proponemos recoger las aportaciones de estos autores en lo que a las propiedades de las relaciones se refiere. Para explicar las propiedades de las relaciones utilizaremos de momento una fórmula sencilla. Así, consideramos la siguiente fórmula:

$$\mathbf{a R b}$$

donde **R** es el nombre de la relación y **a** y **b** son los conceptos que pertenecen a una determinada clase entre los que se establece dicha relación.

2.3.1.1 *Transitividad*

Una relación se considera transitiva si para todos los valores de **a**, **b** y **c**, se cumple que si **a R b** y **b R c**, entonces **a R c**. Por poner un ejemplo de la lengua general, citado por Lyons (1977) y Cruse (2004), si **a** es más alto que **b**, y **b** es más alto que **c**, entonces se cumple que **a** es más alto que **c**.

Conviene definir asimismo qué se entiende por relación intransitiva y por relación no transitiva. Una relación es intransitiva cuando se cumple que si **a R b** y **b R c**, entonces la relación **a R c** es imposible desde el punto de vista lógico. Así, por ejemplo, la relación “ser padre de” que se establece entre padre e hijo es una relación intransitiva, puesto que si **a** es el padre de **b**, y **b** es el padre de **c**, es imposible que **a** sea el padre de **c**.

Una relación no transitiva es aquella en la que si **a R b** y **b R c**, no podemos concluir que la relación entre **a** y **c** se cumpla. Por ejemplo, una relación no transitiva sería que si Pedro es amigo de Luis, y Luis es amigo de Juan, no podemos concluir que Pedro y Juan sean amigos, ni que no lo sean.

La transitividad es la propiedad de las relaciones más empleada a la hora de hacer inferencias, y en ella se basan muchos de los sistemas de inteligencia

artificial para organizar conceptos relacionados mediante una relación hiponímica. La existencia de relaciones transitivas permite la aplicación en los sistemas de inteligencia artificial del denominado principio de la herencia, sobre el que volveremos más adelante.

Se puede deducir que la relación también es válida entre los dos extremos de la relación compuesta si una relación es transitiva y se produce la composición de relaciones (ver apartado 2.3.1.8), siempre y cuando los rasgos semánticos considerados no comporten pertenencia (Otman, 1996: 77).

2.3.1.2 Simetría

Una relación es simétrica si y solo si se cumple que si $a R b$, entonces $b R a$, es decir, si a está relacionado con b mediante la relación R , se puede asegurar que b también está relacionado con a mediante dicha relación. Se produce una relación de implicación en los dos sentidos. Un ejemplo claro sería la relación “estar casado con”. Si José está casado con Carmen, se cumple que Carmen está casada con José. Un ejemplo claro de relación simétrica es la sinonimia, que trataremos con más detalle en el apartado 2.4.5. Las relaciones simétricas son relaciones inversas de sí mismas (Evens, 1988: 24, como se verá en el apartado 2.4.2.4 sobre conversividad, reciprocidad y relaciones inversas).

Al igual que en el caso de la transitividad, podemos hablar de relaciones asimétricas y de relaciones no simétricas. Una relación asimétrica es aquella en la que, si $a R b$, entonces no se cumple que $b R a$. Por ejemplo, si José es el padre de Luis, entonces Luis no puede ser el padre de José. Una relación no simétrica sería aquella en la que no podemos asegurar que si se cumple que $a R b$, entonces se cumple $b R a$. Por ejemplo, si José quiere a Luis, no se cumple necesariamente que Luis quiera a José.

2.3.1.3 Reflexividad

Una relación es reflexiva si y solo si se cumple que $a R a$, es decir, cada uno de los elementos que abarca la clase de objetos a está relacionado consigo mismo mediante la relación R . Un ejemplo sería la relación *tener la misma edad que*, o *tener el mismo nombre que*.

Cruse (2004) considera que esta propiedad de las relaciones no es muy útil para el análisis semántico, puesto que se aplica a un solo elemento. No obstante,

otros autores, como Evens (1988) y Lyons (1977), la mencionan entre las propiedades relevantes de las relaciones. Evens defiende que la relación que se establece entre un todo y sus partes (ver meronimia) es reflexiva, ya que cualquier objeto forma parte de sí mismo.

De manera análoga a lo que ocurre con las relaciones transitivas y simétricas, podemos hablar de relaciones irreflexivas y no reflexivas. Un ejemplo de relación irreflexiva sería la de que una persona no puede ser su propia madre, y un ejemplo de relación no reflexiva sería que si Juan conoce el peso de Luis, esto no implica necesariamente que Juan conozca su propio peso.

Según Lyons, si una relación es simétrica, transitiva y reflexiva, se dice que se trata de una **relación de equivalencia**.

Las tres propiedades presentadas hasta ahora son aquellas que merecen la consideración de todos los autores. A continuación presentamos otras propiedades que merecen nuestra consideración de cara al análisis formal de las relaciones.

2.3.1.4 *Conversividad, reciprocidad y relaciones inversas*

La **conversividad** es una propiedad de las relaciones existentes entre relaciones que tiene que ver con la direccionalidad. Otman (1996) afirma que para cada relación concreta es posible encontrar una relación que podríamos considerar su **inversa**, aunque no siempre se utilice o se cumpla en los dos sentidos. Siguiendo nuestra formulación inicial de las relaciones, si **a R b**, entonces existe una relación **R'** que es la que se establece entre **b** y **a**. Dicho de otro modo, si se cumple la relación **R** entre **a** y **b**, entonces sabemos inmediatamente que entre **b** y **a** se cumple la relación **R'**. Un ejemplo sería que si una bicicleta tiene por parte un manillar, entonces el manillar es parte de la bicicleta. Igualmente, sabemos que si José es hijo de Luis, entonces Luis es el padre de Jose, es decir, “ser hijo de” y “ser padre de” son relaciones inversas.

Cruse (2004) define también las relaciones **semi-conversas**, que son aquellas en las que la reciprocidad se da en uno de los sentidos pero no en el otro. Por ejemplo, en la relación que se establece entre ser médico y ser paciente, ya que si José es el médico de Luis, entonces se cumple que Luis es su paciente; por el contrario, si Luis es el paciente de José, José no es necesariamente su médico, puesto que puede ser su dentista o su psicólogo. La implicación se cumple solo en

uno de los sentidos, es decir, “ser médico de” es el semi-converso de “ser paciente de”.

2.3.1.5 Relación uno-a-uno

Evens (1988) identifica esta propiedad de las relaciones, que se cumple si y solo si, cuando $a R b$ y $c R b$, entonces a y c tienen que ser el mismo concepto. Asimismo debe ser cierto que si $a R b$ y $a R c$, entonces b y c son necesariamente el mismo concepto. Así, por ejemplo, si a está situado justo a la izquierda de b y c no es igual a a , entonces es imposible que c esté justo a la izquierda de b .

2.3.1.6 Recursividad de las relaciones

Como hemos apuntado en el apartado 2.2.2.1, Feliu (2000) tiene en cuenta en su descripción de las relaciones conceptuales la posible recursividad de las mismas. Esta propiedad hace referencia al hecho de que las relaciones pueden tener más de dos argumentos situados a la derecha de la relación. Algunas relaciones pueden darse entre un elemento a y n elementos, siendo n un número ilimitado de conceptos que pueden estar relacionados con el concepto a por medio de la relación R . La recursividad es uno de los aspectos que sirven a Feliu para caracterizar las relaciones conceptuales.

Un ejemplo de relación recursiva sería la relación entre un todo y sus partes, la denominada relación meronímica. Un todo puede tener –y de hecho es bastante habitual que así suceda– más de un elemento constituyente. Así, una bicicleta tiene ruedas, manillar, sillín, etc.

2.3.1.7 Principio de la herencia

A continuación dedicaremos un apartado al principio de la herencia, que se ha mencionado anteriormente como una de las consecuencias de la transitividad de las relaciones hiponímicas, entre otras. No se trata de una propiedad de las relaciones conceptuales en sentido estricto, sino de un principio relevante a la hora de estructurar formalmente el conocimiento y realizar inferencias.

Según Otman, la herencia en las redes semánticas se deriva del principio de economía. Al menos en teoría, todas las propiedades de un nudo superior en una jerarquía pueden ser heredadas por un nudo inferior. La noción de herencia

puede expresarse mediante la fórmula *Si a subsume b, entonces a es más general que b*.

L'héritage dans les réseaux sémantiques relève d'un principe d'économie. Il découle de l'hypothèse selon laquelle toutes les propriétés d'un noeud supérieur d'une hiérarchie sont disponibles (au moins virtuellement) aux niveaux inférieurs. En ce sens, la notion d'héritage est assimilable à la notion de subsomption qui peut se résumer par la formule suivante : *Si A subsume B, A est plus général que B*.

(Otman, 1996: 62)

En la práctica, no todas las características se heredan, y un mismo subordinado puede heredar características de varios superordinados. En algunos casos, las características de dichos superordinados pueden ser incompatibles entre sí.

Según Otman, podría considerarse que los conceptos más genéricos realizan copias de sus “bienes” y las transmiten a sus descendientes sin perder por ello su propia naturaleza. El ascendiente no tiene por qué transmitir todos sus atributos a su descendencia. Algunos de esos atributos, heredados a su vez de sus ascendientes, pueden no poseer un valor determinado, sino ofrecer una serie o un rango de posibles valores entre los que el descendiente ha de elegir.

Otman (1996) identifica cuatro posibles sistemas de herencia basándose en las definiciones de herencia en la programación orientada a objetos¹³, que son herencia simple, herencia múltiple, herencia estática y herencia dinámica:

- a) **Herencia simple**, que se produce cuando un objeto puede tener un único ascendente.

¹³ La Programación Orientada a Objetos (POO) es un paradigma de programación que trata de amoldarse al modo de pensar del hombre, potenciando el mantenimiento, la extensión y la reutilización del software generado bajo este paradigma. El elemento básico de este paradigma no es la función (elemento básico de la programación estructurada), sino un ente denominado objeto. Un objeto es la representación de un concepto para un programa, y contiene toda la información necesaria para abstraer dicho concepto: los datos que describen su estado y las operaciones que pueden modificar dicho estado, y determinan las capacidades del objeto. Para más información sobre la POO puede consultarse http://pisuerga.inf.ubu.es/lsi/Invest/Java/Tuto/I_1.htm.

- b) **Herencia múltiple**, que se da si un objeto puede tener varios ascendientes, es decir, hereda características de varios superordinados a la vez, lo que puede originar conflictos entre los atributos heredados.
- c) **Herencia estática**, que implica una reproducción tal cual de la estructura de un objeto.
- d) **Herencia dinámica**, que es aquella que puede variar para objetos individuales en función del momento de ejecución de un program.

Meyer *et al* (1997: 104) hablan del fenómeno de la herencia únicamente en referencia a las relaciones genéricas. A la vista del ejemplo que proponen (*optical disk, compact disk y CD-Rom*), Meyer *et al* distinguen entre dos tipos de características:

- a) las que en el concepto genérico tienen un valor único que se mantiene en todos los conceptos más específicos
- b) las que en el concepto genérico ofrecen un abanico de valores entre los que el concepto específico “escoge”

Parece, por tanto, que los conceptos subordinados heredan algunas características y otras no. Meyer *et al* mantienen que en la relación genérica el subordinado hereda todas las características del superordinado, y también añade o especializa al menos una de ellas (Meyer *et al*, 1997: 104). Algunos de los valores de las características permanecen, pero otros se modifican. En algunos casos, en el concepto específico aparecen características que no aparecían en el original.

In knowledge-based systems such as CODE, inheritance is typically achieved automatically: any characteristic name and value introduced at one level of a generic-specific hierarchy will automatically inherit to all specific concepts, unless specified by the user at a lower level.

(Meyer *et al*, 1997: 116)

En Meyer *et al* (1992: 958) se hace referencia al papel que pueden desempeñar los mecanismos de herencia dentro de una base de conocimiento terminológico (*Terminological Knowledge Base, TKB*). Por un lado, facilitan el trabajo del terminólogo, que no necesita repetir información de un nivel de la jerarquía al siguiente. Por otro lado, le permiten identificar conflictos cuando se producen cambios en un nivel jerárquico que “se filtran” a niveles inferiores.

Feliu (2000: 60-61), por su parte, distingue cuatro tipos de herencia:

- a) **Herencia monotónica simple:** Herencia de las propiedades de un nodo sólo de un antecesor, y ese valor no puede ser eliminado ni sustituido en ningún punto de la ontología. Sería equivalente a la herencia simple descrita por Otman (1996).
- b) **Herencia múltiple:** Cada nodo puede heredar propiedades de más de un padre. Este tipo de herencia también ha sido descrito por Otman (1996).
- c) **Herencia por defecto:** Este tipo de herencia se emplea en los sistemas informáticos de representación del conocimiento en los que se pueden asignar valores por defecto a un superordinado, que automáticamente se heredan en los niveles inferiores de la taxonomía. Sin embargo, en el nodo inferior se pueden sobrescribir localmente los valores de una propiedad heredada.
- d) **Herencia ortogonal:** Este tipo de herencia consiste en agrupar la información y permitir la herencia múltiple sólo en algunos grupos específicos, pero no en todos.

En definitiva, como afirma Otman (1996: 63), la herencia es un mecanismo que permite compartir y reutilizar las características existentes dentro de las redes semánticas o sistemas de conceptos.

Desde la conceptología, Weissenhofer (1995: 9) llama la atención sobre el hecho de que la herencia no es un principio que necesariamente se cumpla en los sistemas conceptuales establecidos en terminología, ya que en ocasiones se recurre a las denominadas *características independientes* (ver apartado 2.1.2) para establecer el sistema conceptual. Dichas características independientes pueden combinarse y aparecer en distintos niveles del sistema de forma aleatoria y sin tener en cuenta el principio de la herencia.

It should be noted that inheritance is not a strict principle in concept systems. There are also independent characteristics which follow each other on different levels in a concept hierarchy and can be combined with each other at random.

(Weissenhofer, 1995: 9)

2.3.1.8 Composición de relaciones (Otman, 1996: 60-61)

La formalización de los conceptos por medios informáticos permite la composición automática de las relaciones, pero conviene cuestionarse cuándo es pertinente y funcional llevar a cabo la composición de relaciones, y cómo representarlas sin multiplicar su número en exceso.

Otman ejemplifica una composición de relaciones irrelevante el ejemplo del elefante y el peso de la trompa. El peso de la trompa es una característica de una de las partes del elefante, pero eso no implica que sea una característica esencial del elefante, al menos en este caso (tal vez el peso de los colmillos pudiera ser relevante, pero no el de la trompa). Un ejemplo de composición de relaciones que resulta funcional en su modelo de representación semántica es el de las almendras garrapiñadas. Estas almendras están formadas por azúcar y almendras. Las dos partes se pueden relacionar porque el azúcar recubre a las almendras. En este caso, y dado que aparece en la definición, resulta interesante establecer el lazo entre las dos partes.

2.4 Las relaciones conceptuales en la Semántica Léxica

Una vez presentados estos nuevos enfoques teóricos de la terminología y sus aportaciones en lo que se refiere a la naturaleza de los conceptos y sus relaciones, consideramos necesario revisar algunas aportaciones que, desde la Semántica Léxica y la Psicología –e incluso la propia terminología– completan la explicación sobre algunas relaciones conceptuales que presentan mayor complejidad. Nos referimos a la relación genérica (denominada *hiponimia* en la Semántica Léxica), las relaciones partitivas (que en Lingüística se denominan *meronímicas*), la relación de sinonimia y las relaciones causales. Estas aportaciones nos permitirán profundizar en la naturaleza de estas relaciones y afinar nuestra clasificación en el trabajo empírico.

Antes de estudiar las relaciones, conviene aclarar algunos conceptos básicos que se manejan en Semántica Léxica. En la bibliografía sobre relaciones se emplean distintas denominaciones para referirse a ellas en función de si se consideran como relaciones entre palabras o entre sus representaciones mentales.

La primera distinción suele establecerse entre **relaciones léxicas** y **relaciones semánticas**. Según Evens (1988), las relaciones semánticas son aquellas que conectan conceptos, mientras que las relaciones léxicas conectan

palabras. Para Murphy (2003), las relaciones semánticas son aquellas que se definen mediante paradigmas semánticos, frente a las relaciones léxicas, que definen cualquier tipo de relación paradigmática entre palabras, ya sean de naturaleza fonética, morfológica, morfosintáctica o incluso semántica. En definitiva, las relaciones léxicas son aquellas en las que se ven implicadas las palabras, frente a las relaciones contenidas en el lexicón mental, que serían propiamente las relaciones semánticas.

Dentro de las relaciones semánticas, la Semántica Léxica suele distinguir entre **relaciones paradigmáticas** y **relaciones sintagmáticas**. Evens (1988) define las relaciones sintagmáticas como relaciones que conectan palabras que concurren frecuentemente en el mismo texto, y por ello se denominan en ocasiones relaciones colocacionales. Son relaciones entre palabras que aparecen juntas en una determinada estructura sintáctica.

Las relaciones paradigmáticas, por su parte, relacionan palabras que expresan de alguna forma el mismo significado (o la misma parte de un significado). El conjunto de palabras relacionadas conforma un paradigma. Las relaciones que se consideran paradigmáticas son la hiponimia, la meronimia y la sinonimia. Entre las relaciones sintagmáticas se incluyen las relaciones de restricción o modificación, las relaciones argumentales y las relaciones derivacionales. Generalmente las relaciones paradigmáticas se establecen entre palabras de la misma categoría gramatical, mientras que en las relaciones sintagmáticas entran en juego palabras pertenecientes a diversas categorías gramaticales.

Cruse (2004: 147), por su parte, defiende que las relaciones sintagmáticas y paradigmáticas funcionan en un tándem, en el que las primeras delimitan el espacio en el que operan las segundas.

La información paradigmática aparece normalmente en oraciones que son siempre ciertas. La información sintagmática, por su parte, aparece con frecuencia en oraciones ocasionales que describen situaciones particulares.

A diferencia de Evens (1988), Cruse (2004: 147) separa las **relaciones derivacionales** de las relaciones sintagmáticas. Estas relaciones permiten crear familias de palabras a partir de la misma raíz léxica. Su importancia radica en que participan en la estructuración del vocabulario de una lengua.

Por último, puede establecerse una tercera oposición dicotómica entre **relaciones semánticas** y **relaciones conceptuales**. Generalmente se habla de relaciones semánticas cuando nos referimos a la representación formal de las relaciones lingüísticas (Jackendoff, 1983: 95), mientras que cuando se habla de relaciones conceptuales se hace referencia al plano de la representación mental. En el ámbito de la terminología, Kageura (2002: 56) considera para el estudio de las relaciones entre términos que no es ni útil ni necesario distinguir estos dos tipos de relación, si bien se decanta por el término *conceptual*, puesto que *semántico* podría dar la impresión de referirse a algo que depende de la estructura Lingüística y excluye otros aspectos relacionados con el concepto.

Otros autores, como Béjoint y Thoiron (2000: 8), consideran que la diferencia entre lo conceptual y lo semántico no es de naturaleza sino de grado. Hablan de rasgos y relaciones conceptuales cuando se refieren a los términos, y de rasgos y relaciones semánticas cuando se refieren a las palabras.

Valoración

El presente trabajo tiene como objeto de estudio las relaciones entre conceptos en los ámbitos especializados; por tanto, preferimos la denominación *relación conceptual*, por ser la de uso más corriente en terminología. No obstante, si consideramos la naturaleza de las relaciones desde la perspectiva de la Semántica Léxica, el tipo de relaciones que estamos analizando entraría dentro de las denominadas relaciones paradigmáticas, puesto que se basan en la coincidencia de algún rasgo semántico entre conceptos.

Analizamos a continuación las relaciones de hiponimia, meronimia, sinonimia y las relaciones causales.

2.4.1 Hiponimia

La relación de hiponimia se corresponde con la relación que en terminología hemos denominado relación lógica o genérica, o relación que se establece entre un superordinado y sus subordinados. Es una de las relaciones conceptuales más estudiadas desde la Lingüística y la más relevante para la estructuración del vocabulario, ya que está íntimamente ligada al modo en el que las lenguas articulan la experiencia (Cruse, 2002: 20).

En este apartado recogemos las aportaciones de dos de los principales lingüistas que se han dedicado al estudio de las relaciones desde la Semántica

Léxica: Lyons (1977) y Cruse (1986, 2002 y 2004). Recogemos las distintas definiciones y la expresión de las propiedades más relevantes de este tipo de relación, lo que nos permitirá afinar nuestra comprensión de su naturaleza y complejidad.

Lyons (1977) define la hiponimia como una relación que se mantiene entre un lexema subordinado específico y un lexema superordinado más general. Cruse (2002), por su parte, la define como una relación paradigmática que se establece entre elementos pertenecientes a la misma categoría gramatical. Este autor aborda la caracterización de la hiponimia teniendo en cuenta diversos puntos de vista. La hiponimia puede caracterizarse mediante definiciones lógicas (extensionales o intensionales), colocacionales (atendiendo a la normalidad colocacional), componenciales (teniendo en cuenta los rasgos semánticos necesarios y suficientes) o desde la teoría de los prototipos. Veamos brevemente cada una de las definiciones de hiponimia.

Desde el punto de vista **lógico**, la hiponimia puede definirse en términos de inclusión unilateral, es decir, si **a** es un hipónimo de **b**, **a** implica **b**, pero no se cumple que **b** implica **a**.

Desde el punto de vista **colocacional**, la hiponimia se define del siguiente modo:

a es un hipónimo de **b** si y sólo si los contextos normales de **a** son un subconjunto adecuado de los contextos normales de **b**.

Desde el punto de vista **componencial**, **a** es un hipónimo de **b** si y solo si los rasgos que definen a **b** son un subconjunto de los rasgos que definen a **a**. Esta definición sólo es válida cuando los significados de **a** y **b** pueden caracterizarse en términos de rasgos necesarios y suficientes.

En opinión de Cruse (2002), las tres definiciones vistas anteriormente resultan demasiado estrictas, ya que no tienen en cuenta las intuiciones de los hablantes sobre la relación de hiponimia.

Para superar esta restricción, Cruse (2002) propone caracterizar la relación hiponímica desde la teoría de los prototipos. En dicha teoría, una categoría conceptual puede definirse en función de sus rasgos semánticos, pero estos no tienen por qué ser necesarios y suficientes. Cuantos más rasgos semánticos coincidentes con la categoría conceptual superordinada posea una categoría conceptual subordinada, mejor ejemplo será de dicha categoría superordinada.

Así, Cruse propone establecer una escala de importancia de los rasgos semánticos para determinar la prototipicidad de las categorías conceptuales con respecto a su hiperónimo.

Un aspecto importante de la relación de hiponimia es determinar entre qué elementos se establece. Para Cruse, la hiponimia es una relación paradigmática que se da en las principales categorías gramaticales (sustantivos, adjetivos, verbos), pero que es especialmente significativa en sustantivos. Las entidades relacionadas mediante esta relación pueden ser morfemas, lexemas o frases. Cruse (2002) alude también a la posibilidad de relacionar lo que él denomina microsentidos (*microsenses*) de las unidades léxicas, que son los que permiten definir, por ejemplo, un *cuchillo* como un tipo de *cubierto* o como un tipo de *arma*, dependiendo del ámbito en el que nos encontremos. En cualquier caso, al ser la hiponimia una relación paradigmática, los elementos relacionados deben pertenecer a la misma categoría gramatical. En cuanto a la naturaleza léxica o conceptual de la relación hiponímica, Cruse, desde una perspectiva lingüístico-cognitiva, considera que toda relación entre significados es de naturaleza conceptual, resolviendo así el dilema semántico/conceptual al que aludíamos al principio de esta sección. Sin embargo, reconoce que algunos elementos léxicos pueden tener significados que no se recogen en los conceptos asociados a ellos. Por ejemplo, la diferencia de significado existente entre las palabras inglesas *cat* y *pussy* no está incluida en el concepto que representa al objeto gato.

Cruse (2002) hace hincapié en que cuando se establece una relación de hiponimia se establece simultáneamente una relación de incompatibilidad entre los hipónimos de un mismo hiperónimo. Es decir, ambas relaciones funcionan siempre en combinación. Siempre que hay un hipónimo tiene que haber al menos un cohipónimo incompatible con él.

Como hemos visto, la hiponimia suele estudiarse en términos lógicos como una relación de inclusión de clase. Sin embargo, en dicha relación no está claro si es el hipónimo el que está incluido en el hiperónimo o viceversa, ya que depende de si consideramos la extensión o la intensión del concepto. Así, si consideramos el punto de vista de la extensión, el concepto subordinado está incluido en el superordinado, mientras que, atendiendo a la intensión, el concepto superordinado estaría incluido en el subordinado.

En cuanto a sus propiedades, la relación de hiponimia es una relación transitiva, que se puede definir en términos de implicación unilateral. Al mismo tiempo, la hiponimia es una relación asimétrica, ya que la implicación se cumple únicamente en un sentido.

Lyons (1977) añade que en muchos casos, un hipónimo nominal encapsula el significado de un modificador adjetival y lo combina con el significado del lexema superordinado. Los cohipónimos de un mismo superordinado contrastan en significado, y la naturaleza de este contraste se explica por la diferencia que se aprecia en la modificación sintagmática del superordinado encapsulada en cada subordinado.

Cruse (2002 y 2004) identifica una variedad especial de hiponimia, la **taxonimia**, de gran relevancia para la estructuración del conocimiento especializado. Cruse define la taxonimia como una relación vertical que permite describir los lexemas subordinados como un “tipo de” su superordinado. Un taxónimo siempre especifica el significado de su superordinado, restringiendo los rasgos distintivos que lo diferencian del superordinado. La taxonimia permite articular un dominio de manera eficiente. Para ello, las categorías implicadas en una taxonomía deben cumplir los requisitos de ser interiormente coherentes, externamente distintas e informativas al máximo (Cruse, 2002: 13). Cruse sugiere que un taxónimo debe especificar la esencia del hiperónimo.

Las jerarquías taxonómicas se centran en las categorías de nivel básico. Estas categorías son muy informativas y están altamente diferenciadas, y representan el nivel de la taxonomía más alto en el que los elementos se asocian con imágenes visuales claras. El resto de los niveles taxonómicos se definen en relación con el nivel básico. Así, podemos identificar categorías superordinadas, que agrupan más de una categoría básica y las categorías subordinadas, incluidas en las categorías básicas.

Lyons (1977), por su parte, identifica un tipo especial al que denomina **cuasi-hiponimia**. Un ejemplo de este tipo sería la relación establecida entre *forma* y *redondo*, o *sabor* y *salado*. Esta relación no es una relación paradigmática, puesto que no se establece entre lexemas pertenecientes a la misma categoría gramatical.

Valoración

La relación de hiponimia es fundamental para la estructuración del vocabulario en general, y del vocabulario especializado en particular. En los ámbitos especializados tienen especial relevancia las taxonomías, en las que los conceptos se estructuran en función de sus rasgos distintivos con respecto a su superordinado y a sus cohipónimos. En cuanto a las propiedades de la relación, la transitividad de la relación de hiponimia es incuestionable. Otra propiedad que distingue esta relación es su asimetría, es decir, el hecho de que la relación sólo sea válida en un sentido. En la explicación aquí presentada se intuye la importancia de las relaciones entre los cohipónimos de un mismo hiperónimo. Esta relación, que Cruse denomina *incompatibilidad* es la que más adelante trataremos como cohiponimia.

2.4.2 Meronimia

La relación que se establece entre un todo y sus partes (que en terminología hemos denominado relación parte-todo o partitiva) se suele denominar meronimia en la Semántica Léxica, y es la otra relación fundamental para la estructuración del conocimiento, junto con la hiponimia. Al igual que ésta, ha sido estudiada en profundidad, no sólo desde la Lingüística, sino también desde otros ámbitos como la Psicología. La meronimia es una relación compleja, en la que pueden distinguirse varios subtipos. En este apartado recurrimos de nuevo a Cruse (2004) y Lyons (1977) para completar nuestra visión de la relación, y recogemos otras importantes contribuciones realizadas desde la Psicología de Winston *et al* (1987), junto con la de Pribbenow (2002) e Iris *et al* (1988). Completaremos nuestra definición de la meronimia con las aportaciones que, desde la propia terminología, han hecho autores como Van Campenhoudt (1996) y Barrière (2002).

Las relaciones meronímicas poseen similitudes con las relaciones hiponímicas, como apuntan Lyons (1977), Pribbenow (2002) y Cruse (2004). Ambas relaciones son asimétricas y transitivas¹⁴ en el nivel conceptual, y ambas permiten formar jerarquías. Sin embargo, existen diferencias evidentes entre ellas.

¹⁴ Más adelante volveremos a tratar el asunto de la transitividad de las relaciones meronímicas, que es más complejo que en el caso de la hiponimia.

Mientras que la relación hiponímica se basa en la similitud entre conceptos y permite la herencia de rasgos semánticos entre el hiperónimo y los hipónimos, en las relaciones meronímicas la relación se basa en el conocimiento sobre el mundo y se establece entre conceptos distintos. Pribbenow (2002) apunta que en las relaciones meronímicas se da una especie de herencia “hacia arriba”, que permite llegar al todo a partir de las partes (sobre todo en tipos de partes como el material).

Desde la terminología, Van Campenhoudt (1996: 55) se plantea también la problemática de la relación parte-todo. Para él, la frontera que separa la relación meronímica de la hiponímica es, en ocasiones, difusa. Propone el siguiente ejemplo: la *marea alta* es: ¿un momento de la marea? ¿una parte de la marea? ¿un tipo de marea? Si recurrimos a la teoría de conjuntos, esta tampoco nos sirve de gran ayuda, porque la inclusión, como hemos visto, depende del punto de vista.

Al igual que la hiponimia, la meronimia puede definirse en términos de inclusión con distintas direcciones, según se observe desde el punto de vista de la extensión o de la intensión (Cruse, 2004). Sin embargo, no resulta sencillo definirla como una implicación lógica. En opinión de Van Campenhoudt, es difícil distinguir entre la inclusión espacio-temporal, la inclusión de un conjunto dentro de otro y la pertenencia a un conjunto.

A su vez Lyons (1977) considera la dificultad de distinguir entre hiponimia y meronimia. Hace referencia a los lexemas que no se refieren a nombres contables que denotan objetos discretos (como los nombres abstractos o los nombres concretos incontables), en los que las partes pueden concebirse como tales o como hipónimos de un superordinado. Lo mismo ocurre con los nombres colectivos (como por ejemplo *ganado* o *clero*), en los que la relación con los individuos que pertenecen al grupo puede considerarse una relación meronímica o una relación de cuasi-hiponimia (como vimos en el apartado anterior).

En cuanto a la transitividad de la relación partitiva, Cruse (1979 y 2004) distingue entre dominios de función únicos (cuando la parte únicamente puede serlo de un todo) y dominios de función variables (cuando existen varios holónimos posibles). En el caso de que el dominio de función sea único, la transitividad no se mantiene cuando intentamos decir que la parte pertenece a un todo superior. Por ejemplo, en el ejemplo *la mano tiene dedos y el cuerpo humano tiene una mano*, resulta extraño decir que *el cuerpo tiene un dedo*, es decir, no se

mantiene la transitividad. Sin embargo, cuando las partes pueden tener dominios de función variables, es posible decir que un holónimo superior posee dicha parte (por ejemplo, en el ejemplo en inglés *a door has a handle* y *a handle has a door*, *handle* puede formar parte de otros todos, como por ejemplo *a door* y *a cup*). Para Cruse, cuando el dominio de función del todo es variable, se mantiene la transitividad.

Para Winston *et al* (1988), la relación meronímica es siempre una relación transitiva. Los posibles fallos en la transitividad de la relación meronímica se producen cuando se combinan distintos tipos de meronimia en el mismo silogismo.

Iris *et al* (1988) dedican parte de su estudio al análisis de la transitividad en las relaciones parte-todo y concluyen que el hecho de que la relación partitiva sea transitiva depende del tipo meronimia de que se trate. Así, la transitividad se mantiene en las relaciones entre conjunto y subconjunto y en los todos segmentados, pero no es válida cuando hablamos de componentes funcionales ni tampoco en el modelo miembro-colección.

Por otro lado, Cruse (2004) distingue entre **accesorios** (partes de las que se puede decir que están adosadas al todo) y **partes “verdaderas”** (las que forman parte integral del todo). Mientras que en partes verdaderas se mantiene la transitividad, en las segundas resulta extraño asociarlas con un todo superior. Por ejemplo, decir que la rodilla está adosada a la pierna, o que el tronco está adosado al árbol no son expresiones habituales, puesto que rodilla y tronco son partes integrales de su todo correspondiente.

Como vemos, la relación de meronimia resulta más difícil de delimitar que la relación de hiponimia. Para Cruse (2004: 151) existen muchos casos fronterizos en los que es difícil identificar la meronimia. Por este motivo, propone cinco propiedades prototípicas de la meronimia, que contribuyen a determinar la prototipicidad de un ejemplar de esta relación. Estas propiedades son **necesidad, integralidad, discreción, motivación y congruencia**.

- a) **Necesidad:** Algunas partes son necesarias para la constitución de un todo, mientras que otras son opcionales. No se trata de necesidad en el sentido lógico, sino en un sentido canónico, es decir, para que se considere que un todo está bien formado, debe

poseer una determinada parte. Un ejemplo de partes necesarias desde el punto de vista canónico serían los *dedos en una mano*.

- b) **Integración:** Esta propiedad hace referencia al grado de integración entre la parte y el todo. Algunas partes se pueden describir más fácilmente que otras sin hacer alusión al todo al que pertenecen. Compárense, por ejemplo, el mango de una cuchara y el tirador de una puerta. Mientras que resulta difícil definir el mango de la cuchara sin mencionar el todo, no es extraño concebir el tirador de una puerta sin hacer referencia explícita al todo en el que se integra.
- c) **Discreción:** Esta propiedad se refiere a la facilidad para distinguir entre las partes de un mismo todo. Así, por ejemplo, resulta más fácil distinguir entre el brazo y el cuerpo que la punta de la lengua con respecto a la lengua. Cuanto más sencilla resulte esta distinción, más prototípica es la relación entre el todo y la parte.
- d) **Motivación:** La motivación tiene que ver con la función que la parte desempeña con respecto al todo. Cuanto más clara esté dicha función, mejor ejemplo de la relación meronímica será la que se establece entre dicha parte y el todo. Por ejemplo, el tirador de una puerta sirve para abrirla y cerrarla.
- e) **Congruencia:** Esta propiedad tiene tres características: el rango, que hace referencia al nivel de generalidad del merónimo con respecto al del holónimo; la fase, que determina si el todo y las partes coinciden en el tiempo, y la clase conceptual a la que pertenecen el todo y las partes.

Veamos a continuación las clasificaciones de distintos tipos de meronimia que, desde la semántica, la Psicología y la terminología han propuesto Iris *et al* (1988), Winston *et al* (1988) y Barrière (2002), respectivamente.

Iris *et al* (1988) hablan de una “familia” de relaciones, más que de una relación aislada. Las clasificaciones propuestas dependen de la naturaleza de la relación que se establece entre el todo y la parte. Estas autoras desarrollan cuatro modelos (o *schemata*), cada uno de los cuales expresa un tipo de relación parte-

todo. Estos cuatro esquemas de las relaciones meronímicas son el componente funcional, el todo segmentado, la relación miembro-colección y la relación conjunto-subconjunto:

1. **Componente funcional:** El primer esquema contempla la parte como una unidad que funciona dentro de un todo. La parte contribuye al todo, no solo como unidad estructural, sino como un elemento esencial de la actividad del mismo. Un ejemplo es la relación establecida entre *rueda* y *bicicleta*.

2. **Todo segmentado:** Este esquema se refiere a un todo que se divide en partes que comparten con el todo su naturaleza. Se tiene la sensación de que el todo precede a la parte. En este caso el todo está especificado y tiene unos límites discretos. Esta relación es la que se establece entre una *tarta* y las *porciones* en que se divide.

3. **Miembro-colección:** La relación de la parte con respecto al todo es la que se establece entre un conjunto y los elementos discretos que lo componen. Se puede definir un conjunto o una colección en función de las propiedades de sus miembros. Un ejemplo propuesto por estas autoras es la relación que se establece entre *oveja* y *rebaño*.

4. **Conjunto-subconjunto:** Este modelo se basa en el anterior. Los subconjuntos suelen poseer etiquetas propias dentro del conjunto. Cuando tenemos conjuntos y subconjuntos podemos tener una jerarquía. Este último modelo de relación meronímica es la que, en su opinión, puede confundirse con la relación de inclusión y la noción de jerarquía. Un ejemplo propuesto por ellas es la relación entre *fruta* y *comida*.

Esta clasificación supone un avance para el estudio de de las relaciones meronímicas, que podemos completar con la clasificación sugerida desde la Psicología por Winston *et al* (1988).

Winston *et al* (1988) proponen una teoría de las relaciones basada en la descomposición de las mismas en elementos comparables. Esta división les permite establecer una jerarquía de las relaciones de inclusión, que pueden ser de tres tipos: inclusión espacial, meronimia e inclusión de clase. La primera es la más simple y posee como único elemento la inclusión; la segunda incorpora, además de la inclusión, un segundo elemento, la conexión. La tercera relación de inclusión, y la más compleja, es la hiponimia o inclusión de clase, que añade el elemento de similitud a los dos anteriores. Su enfoque defiende que las relaciones

no deben verse como un todo unitario, sino que se pueden descomponer en elementos que permiten compararlas (Chaffin y Herrmann, 1988).

Winston *et al* (1987) basan su clasificación de las relaciones meronímicas en la aparición de tres propiedades, que son la **funcionalidad**, la cualidad de ser **homeomérico**¹⁵ y la condición de ser **separable** o no. Posteriormente añaden una cuarta propiedad, la **simultaneidad** (Chaffin *et al*, 1988). Estas propiedades se pueden definir como sigue:

- **Funcional:** La relación entre la parte y el todo es funcional cuando las partes están en una posición espacial o temporal específica que determina su función dentro del todo. La parte posee una función especial con respecto al todo, como por ejemplo una *rueda* con respecto a un *coche*. Esta propiedad coincide con la que Cruse denomina motivación.

- **Homeomérica:** La relación es homeomérica cuando la naturaleza de una parte es la misma que la del resto de las partes y la del todo al que pertenece. Por ejemplo, una gota de agua sigue siendo agua, mientras que un árbol no es lo mismo que un bosque.

- **Separable:** Las partes y el todo pueden estar físicamente conectadas o no. Por ejemplo, el *asa* con respecto a la *taza*. Esta propiedad coincide con lo que Cruse (2004) denomina discreción.

- **Simultánea:** este rasgo de las relaciones meronímicas se añade en Chaffin *et al* (1988). Las partes simultáneas pertenecen a objetos que poseen la mayoría de sus partes al mismo tiempo (*rueda-coche*), a diferencia de las partes no simultáneas (*adolescencia-crecimiento*). Este rasgo de las relaciones meronímicas se ha añadido para distinguir las relaciones característica-actividad de las componente-objeto. Para Van Campenhoudt (1996: 75), este rasgo es demasiado estricto, puesto que dejaría fuera los componentes facultativos de un objeto, como por ejemplo los accesorios. Este es uno de los rasgos de la propiedad prototípica de congruencia identificada por Cruse (2004).

¹⁵ Algunos autores, como Díez Orzas (1999), optan por mantener la forma inglesa a falta de una traducción satisfactoria. Si bien el término *homeomérico* no es muy frecuente en español, se emplea en el lenguaje filosófico para referirse a la posibilidad de descomponer un todo en partes semejantes (*totalización homeomérica*).

En función de estos rasgos de las propiedades meronímicas, Winston *et al* (1987) distinguen seis subtipos de relaciones meronímicas: componente funcional-objeto integrado, miembro-colección, porción-masa, material-objeto, fase-actividad y lugar-área. A continuación las definimos, aportando ejemplos:

a) **Componente funcional-objeto integrado:** Relación que se establece entre los componentes que cumplen una función y el todo al que pertenecen dichos componentes. Siguiendo a Cruse (1986), las partes que componen un objeto integrado pueden ser canónicas (una mano a la que le faltan dedos es una mano defectuosa) o facultativas (una cocina que no tienen frigorífico no es necesariamente una cocina defectuosa). Un ejemplo de este tipo de relación sería la que se establece entre *taza* y *asa*.

b) **Miembro-colección:** Los miembros de una colección no tienen que cumplir una función específica o tener una disposición estructural determinada con respecto al todo o al resto de las partes. No debe confundirse con la relación entre un miembro y una clase: lo que determina que un concepto pertenezca a una clase es su similitud con el resto de los conceptos de dicha clase, mientras que el hecho de ser miembro de una colección lo determina la proximidad espacial o la conexión social. Un ejemplo de este tipo de relación sería la que se establece entre *bosque* y *árbol*.

c) **Porción-masa:** Lo que diferencia las porciones de otras partes es que son homeoméricas con respecto al todo, es decir, la naturaleza de las porciones es similar a la del resto de las partes y a la del todo. Un ejemplo de este tipo de relación es la que se establece entre *pizza* y *porción*.

d) **Material-objeto:** El material no se puede separar del objeto, es una parte intrínseca del mismo. Un ejemplo de este tipo de relación es la que se establece entre *agua* y *oxígeno*.

e) **Característica-actividad:** Se utiliza para designar las características y fases de procesos y actividades complejos. Un ejemplo sería la relación que se establece entre *comprar* y *pagar*.

f) **Lugar-área:** Relación entre áreas y localizaciones concretas. Los lugares no son partes funcionales del todo, pero se mantiene el rasgo homeomérico, es decir, todos los lugares de un área son similares al resto de los lugares y a toda el área. Un ejemplo sería la relación que se establece entre *campo de fútbol* y *portería*.

En Tabla 7 se resume la tipología de las relaciones meronímicas de Winston *et al* (1987), indicando cuáles son sus propiedades y un ejemplo de cada una.

Relaciones	Ejemplos	Funcional	Homeomérica	Separable	Simultánea
Componente-objeto	<i>handle-cup</i>	+	-	+	+
Miembro-colección	<i>tree-forest</i>	-	-	+	+
Porción-masa	<i>slice-pie</i>	-	+	+	+
Material-objeto	<i>gin-martini</i>	-	-	-	+
Característica-actividad	<i>paying-shopping</i>	+	-	-	+
Lugar-área	<i>oasis-desert</i>	-	+	-	-

Tabla 7. Tipos de relaciones meronímicas y elementos de la relación (Winston *et al*, 1987)

En un artículo posterior Chaffin *et al* (1988) añaden un séptimo tipo de meronimia, la meronimia parte-proceso. Dado que en la teoría de la terminología se suele incluir este tipo de relación entre las relaciones no jerárquicas secuenciales, optamos por no incluirla dentro de las meronimias.

Desde la terminología, Van Campenhoudt (1996: 73-74) rechaza la división de las meronimias que él denomina “temporales” (es decir, la meronimia característica-actividad y la meronimia fase-proceso, caracterizadas por la sucesión en el tiempo de las partes) por considerar que la división entre ambas es demasiado sutil, basada en pruebas dependientes de la lengua en la que se formulan (en virtud del rasgo “separable”). En su opinión, la distinción entre las meronimias temporales se puede establecer simplemente teniendo en cuenta si las partes son canónicas o facultativas (es decir, dependiendo de si aparecen siempre en el proceso o no).

Barrière (2002: 95) propone una jerarquía de las relaciones partitivas que integra las propuestas de Iris *et al* (1988) y Chaffin y Hermann (1988) con las propuestas por Barrière (1997) tal y como se muestra en la Fig. 2:

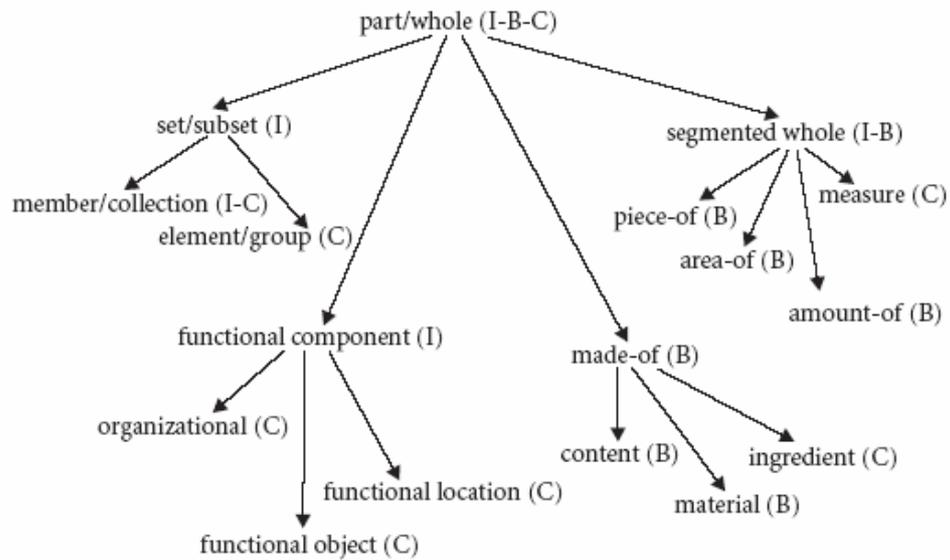


Fig. 2. Jerarquía integrada de las relaciones parte-todo (Barrière, 2002: 95).

En esta figura, los subtipos de relación meronímica van acompañados de una letra que indica qué autores proponen este subtipo de relación. Así, la I indica que es una relación propuesta por Iris *et al* (1988), la C representa los tipos propuestos por Chaffin y Hermann (1988) y la B se refiere a las aportaciones de Barrière (1987).

Pese a que añade subtipos a las relaciones meronímicas propuestas por Winston *et al* (1987), consideramos que los nuevos tipos de meronimia identificados por Barrière (1997) –*piece-of*, *area-of*, *amount-of*, *made-of*, *content* y *material*– se hallan ya incluidos en dicha propuesta de Winston *et al*, y por ello consideramos que se complica innecesariamente la clasificación.

Valoración

Como hemos visto, la relación meronímica es una relación muy compleja. Un aspecto que complica este tipo de relación es su naturaleza prototípica, descrita por Cruse (2004) de forma explícita, pero que se encuentra implícita en el resto de los estudios.

Dicha naturaleza prototípica, junto con la consideración de distintas propiedades específicas de las meronimias, lleva a Winston *et al* (1987) a proponer una clasificación de las relaciones meronímicas que, al menos en principio, parece abarcar todos los casos y solucionar las dudas sobre la transitividad de la relación. Como hemos visto, para Winston *et al* (1987) los

posibles fallos en la transitividad de la relación meronímica se producen cuando se combinan distintos tipos de meronimia.

En nuestro trabajo empírico adoptaremos esta propuesta de clasificación, dado que en nuestra opinión es la que más luz aporta con vistas a la formalización de las relaciones conceptuales, y además ya ha sido tomada en consideración en trabajos previos de formalización de las relaciones conceptuales como los de Otman (1996) y Feliu (2004).

2.4.3 Sinonimia

La relación de sinonimia en sentido estricto no es una relación conceptual, sino una relación establecida entre denominaciones que se refieren al mismo concepto. De hecho, en la terminología clásica orientada a la normalización, la sinonimia se considera denostable, y por tanto se propone su eliminación en aras de una comunicación especializada unívoca, en la que para cada concepto debería existir una única denominación.

Sin embargo, como hemos visto en las revisiones de la teoría general de la terminología, la sinonimia es un hecho presente en la lengua en general y, aunque en menor medida, también está presente en la comunicación especializada. Cabré (1999) y Temmerman (2000) consideran que la variación denominativa es un hecho que debe tenerse en cuenta en los trabajos terminológicos descriptivos. De hecho, para Temmerman incluso desempeña un papel fundamental en el desarrollo del conocimiento.

En la Lingüística y en la Semántica Léxica, la sinonimia se concibe como una relación paradigmática que se establece entre palabras cuyas similitudes semánticas son más destacadas que sus diferencias (Cruse, 2004).

Cruse (2004) distingue tres grados de sinonimia, a saber, la **sinonimia absoluta**, la **sinonimia proposicional** y la **cuasi-sinonimia**:

- a) La **sinonimia absoluta** se da cuando existe una identidad absoluta entre los significados de dos palabras. No son muchos los casos en los que se puede identificar este tipo de sinonimia.
- b) La **sinonimia proposicional** se define en términos de implicación. Dos elementos léxicos se consideran sinónimos proposicionales si la sustitución de uno por otro en una expresión con unas determinadas propiedades o valores de

verdad no afecta a dichas propiedades. Sin embargo, se aprecian algunas diferencias en el significado expresivo, en el nivel estilístico, o en el nivel discursivo. Este tipo de sinónimos son muy abundantes, sobre todo en los casos en los que existe un significado emotivo especial. Cruse (2004: 155) propone como ejemplo la relación entre *violin* y *fiddle* en inglés que, si bien se refieren al mismo objeto *violín*, difieren en cuanto al grado de coloquialidad.

- c) La **cuasi-sinonimia** se produce en los casos en que los usuarios de la lengua tienen intuiciones sobre una cierta relación de semejanza entre las palabras, aunque esta no sea total. Entre los cuasi-sinónimos existen ciertas diferencias permisibles. Un ejemplo propuesto por Cruse (2004: 157) es la relación que se establece entre *fog* y *mist*, que se pueden considerar sinónimos, aunque se diferencian en un matiz de grado.

Con respecto a las propiedades de este tipo de relación, Lyons (1977) la considera un tipo especial de relación de hiponimia transitiva, simétrica y reflexiva.

En su catálogo de relaciones conceptuales, Feliu (2000) considera la sinonimia como un tipo de semejanza positiva. Dentro de la semejanza positiva se distingue entre equivalencia total o sinonimia y equivalencia parcial o semejanza. Frente a estas define las relaciones de semejanza negativa, entre las que se encuentran las relaciones de oposición y contraste. Para esta autora, todas las relaciones de semejanza son simétricas y transitivas.

Valoración

Como hemos visto, la sinonimia no es en sentido estricto una relación conceptual, sino una relación entre denominaciones.

En nuestro trabajo consideraremos la sinonimia como una relación entre denominaciones cuya representación resulta útil en los trabajos descriptivos, puesto que permite dar cuenta de la variación denominativa presente en los lenguajes de especialidad.

2.4.4 Relaciones causales

La conexión existente entre una causa y el efecto que produce es una de las relaciones conceptuales más presentes los ámbitos de especialidad. Como afirman Khoo *et al* (2002: 51), la relación causa-efecto participa en los procesos de toma de decisiones y resolución de problemas, y por ello es relevante en todas las áreas de la ciencia y la tecnología. Para Barrière (2001), la consideración de las relaciones causales en las bases de conocimiento terminológico añade una dimensión dinámica al modelo de un dominio, dado que las relaciones causa-efecto se producen en acciones (*events*).

Este tipo de relación suele aparecer en las tipologías propuestas desde la terminología, si bien no siempre se le asigna el mismo estatus, y en ocasiones su concepción varía dentro de las distintas tipologías.

Al igual que la hiponimia y la meronimia, la causalidad no es una relación simple, sino que su naturaleza es compleja, y en ella se pueden discernir distintos tipos. En este apartado presentamos tres clasificaciones de las relaciones causales propuestas en el marco de la terminología. Se trata de la clasificación propuesta por Nuopponen (1994), que ya esbozábamos en el apartado 2.1.5.2, la propuesta de Feliu (2004) apuntada en el apartado 2.2.2.1 y la jerarquía de las relaciones causales propuesta por Barrière (2002) con vistas a la extracción semiautomática de información sobre relaciones.

Nuopponen (1994) ofrece una clasificación de las relaciones causales situándolas dentro de las relaciones conceptuales de efecto definidas por Wüster y que se han considerado en el apartado 2.1.5.2. Su clasificación distingue entre **relaciones causales de consecuencia** y **coordinación causal de conceptos**. Las relaciones causales de consecuencia se establecen entre conceptos que se refieren a la causa y al efecto, respectivamente, mientras que la coordinación causal entre conceptos se establece entre conceptos que se refieren a causas o efectos. Las primeras dan lugar a secuencias de conceptos simples (por ejemplo, *humedad-corrosión*), que se pueden conectar para formar cadenas causales (por ejemplo, *humedad-corrosión- agujeros*). En las cadenas causales el efecto de la secuencia causal simple anterior se convierte en la causa de la segunda secuencia, y así sucesivamente.

La coordinación causal de conceptos puede subdividirse en **policausalidad** –en la que se combinan varios conceptos que se refieren a causas– y **poliefectualidad**, en la que se coordinan varios conceptos que se refieren a efectos. La policausalidad puede dividirse en **disyunción causal** (conceptos que se refieren a causas alternativas) y **conjunción causal** (conceptos que se refieren a causas que coocurren para dar lugar a un mismo efecto). La poliefectualidad se puede dividir también en **disyunción y conjunción “efectual”**, dependiendo de si los efectos se excluyen unos a otros o se dan simultáneamente.

Otro aspecto interesante sobre las relaciones causales recogido por Nuopponen (1994) es la discusión sobre el tipo de conceptos que se relacionan mediante relaciones causales. Nuopponen defiende que en las relaciones causales intervienen tanto acciones como hechos, que dan lugar, respectivamente, a causas de producción y causas explicativas.

En la tipología de las relaciones conceptuales de Feliu (2004: 37), la relación de causalidad se identifica como una relación de secuencialidad que puede darse entre entidades, actividades y propiedades, sin que los dos elementos de la relación tengan que pertenecer necesariamente a la misma clase conceptual. Feliu distingue entre las **relaciones causales** (causa-efecto) y **relaciones procedurales** (proceso-resultado).

Barrière (2002), por último, propone establecer jerarquías de relaciones, del mismo modo en el que se crean jerarquías de conceptos. Este proceso es lo que esta autora llama *semantic relation refinement*, que permite subdividir las relaciones en relaciones más específicas, con el fin de abordar las relaciones entre conceptos con distintos niveles de detalle dependiendo de la granularidad del estudio. Barrière (2002: 98) propone la jerarquía de relaciones causales que se muestra en la Fig. 3:

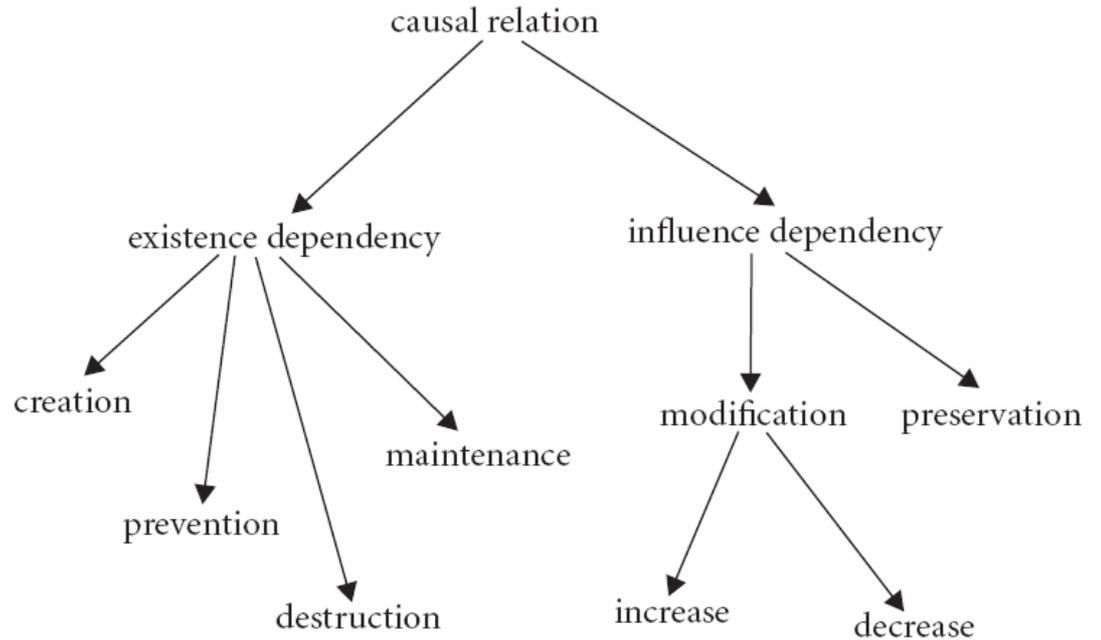


Fig. 3. Tipología de las relaciones causales propuesta por Barrière (2002: 98).

Las relaciones en las que la existencia del concepto depende de la causa (*existence dependency*) se basan en la tendencia que tiene una entidad o relación a existir o no, mientras que las relaciones de influencia afectan a las características que poseen las entidades o actividades (es decir, no son las entidades o características las que tienden a existir o no, sino que aumenta o disminuye el valor de las características). La dependencia para la existencia presenta cuatro casos distintos, que son creación, prevención, destrucción o mantenimiento. Las relaciones en las que la causa influye sobre la entidad o la relación pueden subdividirse en dos casos: modificación y conservación. La modificación puede suponer un aumento o una disminución de las características de las actividades o entidades.

Otro aspecto relevante de las relaciones causales discutido por Barrière es el concepto de certeza (*certainty*) de la relación. Mientras que otros tipos de relación, como las relaciones de hiponimia o meronimia, poseen valores de verdad simples, las relaciones causales poseen un componente de probabilidad e incertidumbre que, en opinión de Barrière, es necesario tener en cuenta en la representación de las relaciones causales. Su propuesta para codificar el grado de certeza de una relación causal en una representación del conocimiento consiste es

emplear una escala de cinco valores para representar grados de certidumbre (*critical, expected, possible, unexpected y excluded*).

Valoración

La complejidad de las relaciones causales queda de manifiesto en las clasificaciones de las mismas propuestas por las autoras revisadas en este apartado. Sin embargo, no nos parece seguro que sea necesario distinguir entre distintos tipos de causalidad en la formalización de las relaciones. Creemos que, al menos *a priori*, la consideración de tantos tipos de causalidad podría multiplicar innecesariamente el número de relaciones consideradas.

Durante el estudio empírico determinaremos si la subdivisión de las relaciones causales es útil desde el punto de vista metodológico para la representación de las relaciones conceptuales, empleando para ello un editor de ontologías.

2.5 Recapitulación y conclusiones

En este capítulo hemos presentado las relaciones conceptuales tal y como se conciben en el enfoque clásico de la terminología y en algunas de las revisiones que se han sucedido en los últimos años desde la propia terminología. Hemos completado la visión de las relaciones propuesta desde la terminología con las aportaciones de la Semántica Léxica y la Psicología.

En la primera sección del capítulo se ha resumido la perspectiva de la Teoría General de la Terminología. Así, hemos visto qué se entiende por concepto, cuáles son sus rasgos distintivos o características, y cómo se forman nuevos conceptos a partir de conceptos existentes. Por último, se han descrito las principales relaciones conceptuales contempladas desde la TGT.

La Teoría General de la Terminología destaca la importancia de las relaciones conceptuales como elemento estructurador del conocimiento especializado en un área de especialidad concreta. Constatamos, no obstante, que la noción de relación conceptual se contempla como una especie de primitiva axiomática que no se define de forma explícita, sino que se clasifica directamente sin aportar una definición clara.

La TGT se ha centrado especialmente en las relaciones genéricas y partitivas entre conceptos, limitándose a ofrecer una lista más o menos exhaustiva

de otras relaciones no jerárquicas, cuya utilidad para la organización del conocimiento se reconoce. Sin embargo, las aportaciones que se ofrecen desde la TGT resultan limitadas, ya que no profundizan en la naturaleza de las relaciones, y, a excepción de las contribuciones más recientes de Weissenhofer (1995) y Nuopponen (2005), resultan insuficientes para reflejar la complejidad de algunas relaciones.

El enfoque funcionalista y dinámico de la terminología defendido por Sager y Kageura arroja luz sobre las relaciones conceptuales al proponer que se consideren las clases conceptuales implicadas en cada relación. El catálogo de relaciones propuesto por Sager (1990) es uno de los más completos, sin caer en la excesiva prolijidad de otras propuestas como la de Nuopponen (2005). Por ello, lo tomaremos como punto de partida en nuestro trabajo empírico para elaborar nuestro catálogo de relaciones conceptuales.

La Teoría Comunicativa de la Terminología acerca posiciones entre la Terminología y la Lingüística, y destaca la naturaleza lingüística del término. Creemos que la propuesta de clasificación de las relaciones aportada por Feliu (2004) desde la TCT es un punto de partida para la formalización de las relaciones, puesto que toma en consideración propiedades como la transitividad y la simetría y además propone un formalismo para las relaciones que nos permite tener en cuenta tanto las clases conceptuales implicadas en cada relación como las propiedades lógicas de las relaciones. Por ello la retomaremos en nuestro trabajo empírico al proponer nuestro propio catálogo de relaciones conceptuales.

El enfoque sociocognitivo de la terminología (Temmerman, 2000) considera la prototipicidad de las categorías conceptuales y presenta una plantilla para recoger las relaciones intra- e intercategoriales que puede utilizarse como instrumento en la formalización de las relaciones conceptuales. Destacamos su defensa de la funcionalidad de la sinonimia y la polisemia en el avance del conocimiento especializado.

En el apartado 2.3 hemos definido las propiedades lógicas de las relaciones. A pesar de que tradicionalmente no se han tratado en terminología, creemos que su consideración es fundamental de cara a la formalización de las relaciones empleando un editor de ontologías, ya que la transitividad y la herencia son el principal principio estructurador de las jerarquías de conceptos. El resto de propiedades nos permitirán hacer razonamientos e inferencias en la búsqueda de

conceptos, teniendo en cuenta las relaciones conceptuales que se establecen entre ellos y la naturaleza de los mismos.

Por último, el estudio de las relaciones desde la Semántica Léxica y la Psicología, junto con algunas aportaciones recientes de la Terminología, nos permite comprender mejor la naturaleza de algunas relaciones complejas, como la hiponimia, la meronimia, las relaciones causales y la sinonimia.

**CAPÍTULO 3: FORMALIZACIÓN DE LAS RELACIONES
CONCEPTUALES EN TERMINOLOGÍA Y LINGÜÍSTICA**

3 Formalización de las relaciones conceptuales en Terminología y Lingüística

Uno de los objetivos del presente trabajo es la propuesta de un modelo de formalización de las relaciones conceptuales. Por ello, en este capítulo revisamos las propuestas de formalización realizadas desde la Terminología y desde la Lingüística. En primer lugar, repasaremos brevemente las clasificaciones que tradicionalmente se realizan con fines documentales para ordenar y recuperar la información contenida en todo tipo de documentos.

En la sección 3.1 se estudia cómo se organizan los conceptos en la Teoría General de la Terminología por medio de sistemas conceptuales. Veremos cómo se estructuran y representan estos sistemas, y analizaremos las limitaciones que plantean los sistemas de conceptos en terminología.

En la sección 3.2 presentamos los nuevos modelos que propone la terminología para la organización del conocimiento especializado. En primer lugar, describimos las bases de conocimiento terminológico sugerida por Meyer y sus colaboradores a principios de los noventa (3.2.1). A continuación presentamos el modelo de Redes Semántico-Terminológicas de Otman (3.2.2), el empleo de grafos e hipertexto propuestos por Antia (3.2.3), los esquemas relacionales de Oster (3.2.4) y la aplicación de las funciones léxicas y el diccionario explicativo-combinatorio de Mel'cuk (1997) a la terminología basada en corpus de Dancette y L'Homme (3.2.5).

La sección 3.3 está dedicada al empleo más reciente de las ontologías para la estructuración de conceptos. En ella presentaremos varios proyectos que han aplicado las ontologías en el proceso terminográfico, como los proyectos ONCOTERM (3.3.1), CAOS (3.3.2), la propuesta termontográfica de Kerremans y Temmerman (3.3.3) y el proyecto GENOMA-KB (3.3.4). Completaremos esta sección con una revisión de las ventajas e inconvenientes del empleo de ontologías para el trabajo terminográfico (3.3.5).

Una vez vistas las formalizaciones en terminología, dedicaremos la sección 3.4 a presentar las propuestas de creación de ontologías y bases de datos léxicas hechas desde la Lingüística. Describiremos algunas de las propuestas más relevantes como WordNet (3.4.1), la ontología Mikrokosmos (3.4.2) y la base de datos léxica FrameNet (3.4.3).

Concluiremos con la recapitulación y conclusiones de lo visto en este capítulo (3.5), que nos ha permitido observar la evolución en las formas de representación del conocimiento especializado en los últimos años hasta llegar a la consideración de las ontologías como un instrumento útil para la terminología.

Antes de pasar a la descripción de la formalización de las relaciones conceptuales en Terminología, consideramos que es necesario hacer un repaso de los sistemas de clasificación empleados desde hace siglos en la Documentación con el fin de organizar y recuperar documentos y la información contenida en ellos.

La organización del conocimiento es una capacidad exclusivamente humana que nos distingue del resto de los animales. Esta organización se lleva a cabo mediante la clasificación, que consiste en la ordenación de clases aplicando un método. Las primeras clasificaciones conocidas se remontan al siglo IX a. C. en la Biblioteca de Asurbanipal en Nínive. A lo largo de los siglos, se han desarrollado otras clasificaciones, como la de Calímaco en la Biblioteca de Alejandría (siglo III a. C.) o la Ars Magna de Ramón Llull (siglos XIII-XIV), hasta llegar a los sistemas de clasificación más modernos, como el Sistema Dewey (1873) y la más conocida Clasificación Decimal Universal (1893).

En épocas más recientes se han desarrollado formas de estructurar el conocimiento como las taxonomías y tesauros, similares a la organización de conceptos que se lleva a cabo en Terminología. Dadas estas similitudes con las terminologías, describiremos brevemente qué es un tesoro y cómo se organiza.

Un tesoro se puede definir de dos modos distintos, dependiendo de si atendemos a su estructura o a su función (Cordón *et al*, 2001). Por su estructura, un tesoro es un vocabulario controlado y dinámico de términos que mantienen entre sí relaciones semánticas y genéricas y que se aplica a un campo particular del conocimiento. Por su función, un tesoro es un instrumento de control de la terminología, utilizado para trasladar a un lenguaje más estricto el lenguaje natural utilizado en los documentos. Nos interesa sobre todo fijarnos en la estructura del tesoro, en la que se hace referencia a las relaciones semánticas y genéricas que se establecen entre los términos de un vocabulario controlado. Veamos, pues, cómo es la estructura interna de un tesoro y cómo se hacen explícitas en él las relaciones entre términos.

Los componentes fundamentales de los tesauros son los términos, denominados también **palabras clave**. Los términos pueden ser simples o compuestos según su composición, dividiéndose en identificadores e indicadores de función. Atendiendo a su relevancia, los términos pueden considerarse **descriptores** (términos preferentes) y **términos secundarios** o equivalentes (términos no preferentes).

Una de las características fundamentales de los términos de un tesoro es que están interconectados mediante relaciones jerárquicas, relaciones asociativas y relaciones de equivalencia. Las relaciones jerárquicas se dan entre términos genéricos y términos subordinados o específicos. Las relaciones asociativas se establecen cuando la organización se realiza de forma horizontal por asociación de ideas afines desde una perspectiva semántica. Las relaciones de equivalencia se establecen entre los descriptores y los términos secundarios. Este tipo de relación sirve para marcar la preferencia por el empleo de los primeros frente a los segundos.

Además de las relaciones, el uso o significado de los términos se aclara mediante las denominadas notas de aplicación, que permiten definir el término o aclarar su uso.

La formalización de las relaciones en los tesauros se realiza por medio de notaciones alfabéticas que se colocan precediendo al término al que hacen referencia. En la Tabla 8 se recogen de forma resumida las notaciones utilizadas para designar las relaciones entre términos (Currás, 2005: 106-107):

<i>Notación (español/ inglés)</i>	<i>Significado</i>
NA/SN	Nota de aplicación
USE/USE	Indica el descriptor (término preferente) que se elige entre varios términos sinónimos o cuasi-sinónimos
UP/UF	Indica el término equivalente (no preferente)
TC/TT	Término cabecera. Se aplica a los términos que identifican el nombre de la clase más amplia a la que pertenece el concepto específico
TG/BT	Término genérico
TGG/BTG	Término genérico (genérico)
TGP/BTP	Término genérico (partitivo)
TE/NT	Término específico
TEG/NTG	Término específico (genérico)
TEP/NTP	Término específico (partitivo)
TR/RT	Término relacionado

Tabla 8. Notaciones utilizadas para representar las relaciones semánticas entre términos en los tesauros (Currás, 2005).

A modo de ejemplo, presentamos en la Fig. 4 una captura de pantalla del tesauro Eurovoc, que abarca todos los ámbitos de actividad de las Comunidades Europeas y permite indizar los documentos en los sistemas de documentación de las instituciones europeas y de sus usuarios¹⁶. En dicha captura podemos ver el descriptor *ayuda humanitaria*, ubicado dentro del microtesauro *política de cooperación*. Este concepto está relacionado con el descriptor *ECHO* (utilizado para referirse a la Oficina Humanitaria de la Comunidad Europea) y que posee tres conceptos subordinados (*ayuda a los refugiados*, *ayuda a los siniestrados* y *voluntario internacional*). Estos conceptos más específicos, a su vez, establecen distintas relaciones con otros conceptos contenidos en el tesauro, a los que se puede acceder mediante hipervínculos.

En los últimos tiempos, los expertos en Biblioteconomía y Documentación también han vuelto sus ojos hacia las herramientas procedentes de la ingeniería

¹⁶ Este tesauro multilingüe se encuentra disponible de forma gratuita en todas las lenguas oficiales de la Unión Europea a través de la página <http://europa.eu/celex/eurovoc> (Fecha de consulta: julio de 2006).

del conocimiento, en concreto en las ontologías, como un avance más en la representación de los conceptos y las relaciones que se establecen entre ellos¹⁷.

Como vemos, la organización de términos y conceptos no es una actividad exclusiva de la Terminología. La Documentación lleva siglos dedicada a proponer formas de estructuración de los conceptos de un dominio de especialidad, si bien los fines que persigue no son los mismos. Conviene, pues, tener en cuenta las aportaciones de esta disciplina para la Terminología.

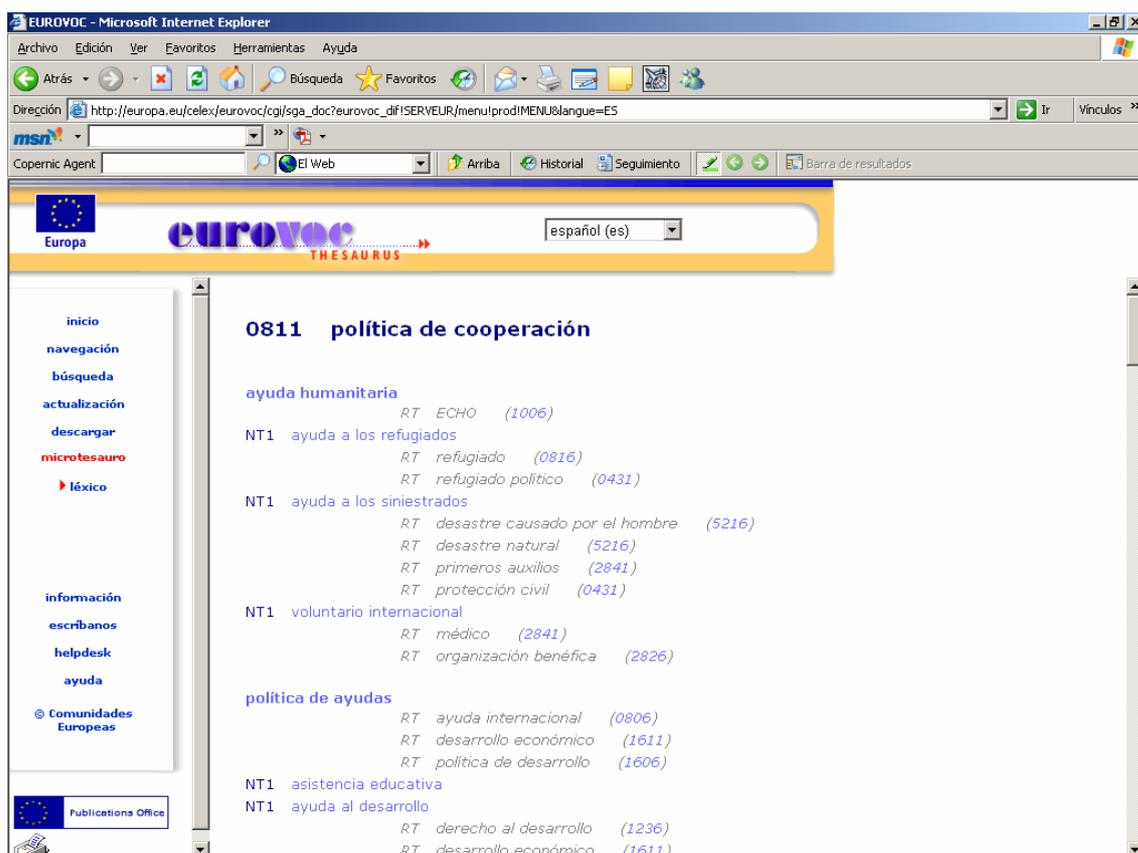


Fig. 4. Ejemplo del descriptor *ayuda humanitaria* procedente del tesauro Eurovoc.

3.1 Representación de las relaciones conceptuales en la TGT: sistemas de conceptos

La terminología de un dominio, entendida como el conjunto de conceptos y sus denominaciones en un ámbito del saber concreto, no se presenta como una

¹⁷ Para obtener una visión amplia y reciente sobre la aplicación de ontologías en Documentación, recomendamos la lectura de García Jiménez (2004), Arano (2005) y Currás (2005).

colección arbitraria de términos, sino que se estructura en los denominados **sistemas de conceptos**. La gestión de la terminología busca la generación de sistemas que nos permiten ordenar los conceptos en función de las relaciones que se establecen entre ellos dentro de un dominio específico (Wright y Budin, 1997: 337).

Este sistema ha de ser coherente y reflejar las relaciones existentes entre los conceptos, puesto que la posición que ocupa un concepto dentro del sistema está determinada por las relaciones con el resto de conceptos del dominio. Esta organización dependerá del punto de vista desde que se analicen los conceptos, que a su vez está condicionado por el ámbito temático y la perspectiva desde la que se considere.

Wüster (1998: 57) establece una diferencia clara entre sistema de conceptos y **clasificación temática**. Para Wüster, un conjunto ordenado de conceptos cualesquiera no constituye un sistema de conceptos *per se*. Para asignarle esta etiqueta es necesario que exista una relación conceptual de tipo lógico u ontológico. En el caso de que las relaciones no pertenezcan a ninguno de esos dos tipos, Wüster habla de clasificaciones temáticas. En su opinión, las relaciones temáticas son relaciones de ocurrencia, que sólo aparecen ocasionalmente en un documento. Wüster reserva este tipo de relaciones para los sistemas temáticos empleados en el campo de la documentación y la información.

Según la norma ISO 704 (2000: 14), los sistemas de conceptos deben cumplir las siguientes funciones:

- a) Modelar las estructuras de los conceptos a partir de los conocimientos propios de un dominio.
- b) Clarificar las relaciones entre conceptos.
- c) Servir de base para una terminología unificada y normalizada.
- d) Facilitar la comparación de los conceptos y denominaciones en varias lenguas.
- e) Facilitar la redacción de definiciones.

La norma identifica distintos sistemas de conceptos dependiendo del tipo de relaciones conceptuales que se elijan para su estructuración. Así, se habla de **sistemas genéricos**, **sistemas partitivos** y **sistemas asociativos** de conceptos, y en las normas se especifican distintas recomendaciones para su representación.

Arntz y Picht (1989) distinguen entre **sistemas monojerárquicos** y **sistemas polijerárquicos**. Los sistemas monojerárquicos son aquellos en los que en cada nivel se emplea un único criterio de ordenación. Los sistemas polijerárquicos son aquellos en los que se emplean varios criterios de ordenación en el mismo nivel de abstracción. Estos dan lugar a sistemas combinados, que surgen cuando un concepto se divide atendiendo a varios tipos de características. Los sistemas monojerárquicos suelen ser más claros que los polijerárquicos, pero estos reflejan en mayor medida la multidimensionalidad de los conceptos. Mediante la combinación de dos o más tipos de relaciones (genéricas y partitivas, por ejemplo) se pueden obtener **sistemas mixtos**, aumentando de este modo la expresividad del sistema, y por tanto su valor como método de ordenación.

Estos autores emplean asimismo la noción **campo de conceptos** para referirse a un grupo de conceptos estructurados únicamente de forma temática, que sirve como punto de partida para el establecimiento de un modelo conceptual. El establecimiento de dicho modelo se efectúa a partir del análisis de las características de cada concepto, así como de las relaciones entre conceptos.

Nuopponen (1994) propone la consideración de criterios cuantitativos, cualitativos y estructurales (formales) para la clasificación de las relaciones conceptuales y los sistemas de conceptos. Teniendo en cuenta el criterio estructural, Nuopponen propone la siguiente clasificación de los sistemas de conceptos:

a) **Sistemas de conceptos jerárquicos**: Dentro de ellos distingue entre sistemas de conceptos lógicos, sistemas de conceptos con relaciones de contacto (excepto los sistemas temporales) y sistemas de conceptos jerárquicos mixtos.

b) **Sistemas de conceptos secuenciales**: En este tipo de sistemas los conceptos aparecen en una secuencia o en filas de conceptos alternativos o coordinados. Dentro de ellos distingue entre los sistemas de conceptos temporales, sistemas de conceptos causales (cadenas causales) y sistemas conceptuales en desarrollo.

c) **Sistemas de conceptos heterárquicos**: Los sistemas de conceptos heterárquicos son aquellos en los que conceptos distintos pueden constituir el punto de partida para un sistema conceptual según un punto de vista determinado. Dentro de ellos se situarían los sistemas conceptuales causales, los sistemas funcionales, los sistemas de interacción y los sistemas satélite. Estos últimos

ofrecen una gran flexibilidad, ya que se basan en el mapa conceptual contenido en la mente de los hablantes.

Los manuales de terminología y las normas ofrecen una serie de recomendaciones para representar gráficamente las distintas relaciones conceptuales. Así, la norma ISO 704 (2000) propone que para las relaciones genéricas se utilice una representación gráfica en forma de árbol, que refleje las series genéricas horizontales y verticales, o bien una lista de conceptos con una notación alfanumérica separada por puntos para indicar los diferentes niveles.

Las relaciones partitivas, por su parte, se suelen representar gráficamente mediante una estructura de corchetes, y su notación numérica, según la norma ISO 704 (2000), es similar a la de las relaciones genéricas. Otros autores, como por ejemplo Arntz y Picht (1989), recomiendan separar los distintos niveles mediante guiones en la presentación de un sistema conceptual partitivo en forma de lista, para diferenciarlos de los sistemas basados en relaciones genéricas.

En cuanto a las relaciones no jerárquicas, la norma ISO 704 (2000) no da ninguna indicación sobre cómo han de representarse gráficamente ni en forma de lista. Sin embargo, en algunos de los ejemplos que ofrece las representa gráficamente mediante flechas con dos punteros.

En la Fig. 5 ofrecemos un ejemplo de representación gráfica del sistema de conceptos para los instrumentos de escritura recogido en la norma ISO/704 (2000: 15)¹⁸.

Como puede verse, en un primer nivel el concepto superordinado *instrument d'écriture* se subdivide mediante la relación genérica en los conceptos subordinados *marqueur*, *crayon* y *stylo*. *Crayon* a su vez se subdivide en cuatro conceptos subordinados en función de su uso. En un nivel inferior, se indican las partes del concepto *porte-mine*, con lo que se introduce una relación partitiva en el sistema. Las partes *corps* y *mécanisme d'avancement de la mine* tienen a su vez subpartes. Por último, el concepto *mine* se relaciona con el concepto *graphite* mediante la relación asociativa producto-composición.

¹⁸ Para la representación gráfica hemos empleado el programa GeneSis para la generación de sistemas de conceptos, descrito en Monterde Rey (2003).

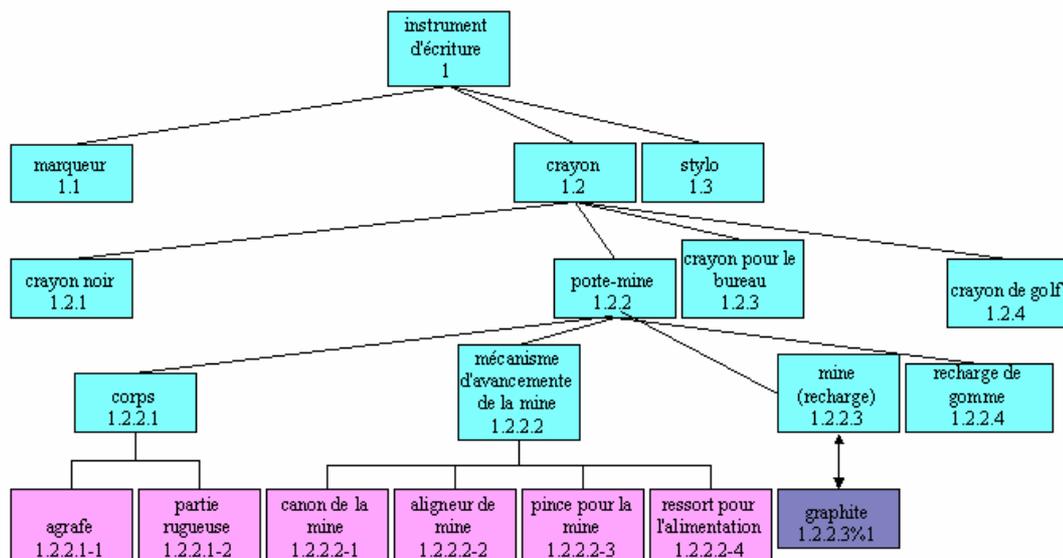


Fig. 5. Representación gráfica del sistema conceptual de *instrument d'écriture* (ISO 704, 2000: 15)

Además de la representación gráfica, las normas proponen que los sistemas de conceptos se presenten en formato de lista numerada. El sistema presentado en forma de notación para el concepto *instrument d'écriture* es el siguiente:

- 1 instrument d'écriture
 - 1.1 marqueur
 - 1.2 crayon
 - 1.2.1 crayon noir
 - 1.2.2 porte-mine
 - 1.2.2.1 corps
 - 1.2.2.1-1 agrafe
 - 1.2.2.1-2 partie rugueuse
 - 1.2.2.2 mécanisme d'avancement de la mine
 - 1.2.2.2-1 canon de la mine
 - 1.2.2.2-2 aligneur de mine
 - 1.2.2.2-3 pince pour la mine
 - 1.2.2.2-4 ressort pour l'alimentation
 - 1.2.2.3 mine (recharge)
 - 1.2.2.3%1 graphite
 - 1.2.2.4 recharge de gomme
 - 1.2.3 crayon pour le bureau
 - 1.2.4 crayon de golf
 - 1.3 stylo

Valoración

La TGT propone formalizar las relaciones entre conceptos mediante la representación gráfica de los conceptos y sus relaciones y su presentación en forma de listas de conceptos a los que se asigna una notación numérica. Si bien estas representaciones resultan expresivas en los sistemas de conceptos basados en relaciones jerárquicas (genéricas y partitivas), es difícil dar cuenta de otros tipos de relaciones. En definitiva, la propuesta de representación de la TGT resulta plana y no da cuenta de todos los posibles tipos de relaciones que pueden darse en una red de conceptos, y en ellos la multidimensionalidad de los sistemas de conceptos se puede representar sólo de una manera parcial. Como ya apuntaban Felber y Budin en 1989, las áreas de especialidad no siempre pueden representarse íntegramente por medio de sistemas de conceptos (Felber y Budin, 1989: 167).

Como ya afirmaba Sager (1990/1993: 55), el desarrollo de los instrumentos informáticos que permiten representar las relaciones entre conceptos reclama una profundización en la representación de dichas relaciones.

El hecho de que ahora podamos modelar sistemas relativamente complejos y sus relaciones internas con un ordenador ha influido en la teoría hasta el punto de haber provocado un gran aumento en el número de relaciones entre conceptos consideradas útiles para estructurar los campos conceptuales. Coincidimos con Sager en que el punto de vista simplista del pasado que proponía las relaciones entre los conceptos mediante tres tipos (genérico, partitivo, otros) ya no tiene validez ni apoyo (Sager, 1990/1993: 55).

3.2 Nuevos modelos para la formalización de relaciones en Terminología

En esta sección presentamos algunas de las nuevas formas de representación de conceptos y de las relaciones que entre ellos se establecen. Todas ellas sugieren modos de superar las limitaciones de los sistemas conceptuales tradicionales que hemos visto en el apartado anterior. En el proyecto COGNITERM (apartado 3.2.1) se sugiere la creación de bases de conocimiento terminológicas que proporcionen distintos puntos de acceso a la información conceptual. Otman (apartado 3.2.2) propone la creación de redes semántico-terminológicas que estructuren los conceptos en función de las relaciones que los unen, considerando las propiedades lógicas de cada relación y los sistemas de

herencia. Antia (apartado 3.2.3) se fija en las posibilidades de representación de los conceptos mediante grafos conceptuales que reflejen todas las relaciones posibles entre conceptos. Oster (apartado 3.2.4), por su parte, propone la formalización de las relaciones semántico-conceptuales detectadas en el seno de los términos compuestos mediante esquemas relacionales. Por último, Dancette y L'Homme (apartado 3.2.5) recurren a las funciones léxicas para dar cuenta de las relaciones que se establecen entre los conceptos de un diccionario sobre términos del comercio minorista.

3.2.1 Bases de conocimiento terminológico: proyecto COGNITERM

Se atribuye a Ingrid Meyer y a su equipo de la Universidad de Ottawa la paternidad del término *Terminological Knowledge Base* (TKB, base de conocimiento terminológico), que se define como una base de conocimiento en el sentido que se le da en inteligencia artificial, es decir, una base de datos que contiene información altamente estructurada sobre los atributos y relaciones de los conceptos (Meyer *et al*, 1997: 109).

Meyer *et al* (1997: 102) dividen las características en función del tipo de vínculo que existe entre la característica y el concepto que describe. Así, distinguen entre **atributos** y **relaciones**. Los atributos son características inherentes al concepto, mientras que las relaciones son características que proporcionan información sobre otros conceptos vinculados con el descrito. Estos autores hablan de un tercer tipo de características, las **funciones** o **finalidades**, cuyo lugar en la tipología de las características no queda definido.

Meyer *et al* (1997) insisten en la importancia de las relaciones para la gestión de la terminología, ya que facilitan la adquisición del conocimiento especializado.

Among the types of characteristics that concepts may have, relations play a very important role in terminology management because they are fundamental to understanding how concepts within a subject field are interlinked, which in turn is crucial to the process of acquiring subject-field knowledge.

(Meyer *et al*, 1997: 103)

La formalización de los atributos y relaciones conceptuales entre términos se realiza mediante el sistema de gestión del conocimiento CODE (Conceptually Oriented Description Environment). Esta herramienta no sólo permite almacenar el conocimiento y recuperarlo a partir de la base, sino que además facilita el análisis conceptual en el momento de la elaboración de la base de conocimiento (Meyer *et al*, 1997:108). Las unidades básicas contenidas en el sistema CODE son los conceptos, las características y la facetas, entendidas como información relacionada con el tipo de características seleccionadas. Los conceptos en CODE se representan por medio de rectángulos que constituyen nodos de un gráfico unidos mediante flechas que representan las relaciones.

Este sistema aprovecha el principio de la herencia entre conceptos superordinados y subordinados. Esta herencia se consigue de forma automática, es decir, los nombres y los valores de las características que se introducen para un concepto en un determinado nivel de la jerarquía genérico-específico son heredados automáticamente por los conceptos situados por debajo de él en la jerarquía, a no ser que quien edita la base de conocimiento especifique lo contrario en el nivel inferior.

Una de las herramientas que permiten la detección de inconsistencias es la denominada *Characteristic Comparison Matrix* (CCM, matriz de comparación entre las características), que es un explorador que permite la comparación de la estructura intensional de los conceptos. La CCM sirve como punto de partida para seleccionar las características esenciales que merecen aparecer en la definición de los conceptos, y al mismo tiempo permite identificar sinónimos. A su vez, la CCM sirve como herramienta para la edición de las características y facilita su introducción de manera consistente en la base de conocimiento terminológico.

En Meyer *et al* (1992) se identifican las principales ventajas o avances que supone CODE/COGNITERM como base de conocimiento terminológico, en comparación con las bases de datos terminológicas tradicionales. La primera ventaja es la capacidad de codificar la información conceptual de forma explícita, haciendo visibles las relaciones conceptuales, favoreciendo la consistencia y permitiendo la representación gráfica.

La segunda ventaja repercute en la adquisición y sistematización de la información, ya que al organizar la información por medio de mecanismos de herencia se facilita el trabajo del terminólogo y la detección de inconsistencias.

La tercera ventaja tiene que ver con la posibilidad de recuperar la información a partir de la información conceptual, ya que COGNITERM permite la búsqueda de información de manera onomasiológica, es decir, a partir de la información sobre características.

El sistema de gestión original CODE evolucionó a uno nuevo denominado IKARUS y posteriormente a DOCKMAN, que es el que se está desarrollando en la actualidad. Lamentablemente, este sistema no se encuentra disponible para su consulta a través de internet.

3.2.1.1 *Gestión de la multidimensionalidad en COGNITERM*

En teoría, la estructuración de los conceptos según las relaciones existentes entre ellos debería resultar sencilla. Sin embargo, como apuntan Meyer *et al* (1997), esta pretensión de simplicidad rara vez puede llevarse a la práctica en sistemas conceptuales reales. Esto se debe a que las diversas áreas de conocimiento no se estructuran en forma de árbol, tal y como nos propone la teoría clásica de la terminología, sino que los conceptos se interrelacionan formando redes complejas, debido a que algunos conceptos pueden tener más de un concepto genérico, o se pueden clasificar en función de más de una característica. Este fenómeno es lo que se conoce como multidimensionalidad:

Real concept systems, of course, are normally more complex than the “neat-and-tidy” tree structures found in terminology textbooks. Some areas of knowledge structure may be more like a very tangled web than like a tree, since some concepts may have more than one generic concept, leading to the complex phenomenon of multiple inheritance [...] A very common source of complexity within concept systems is multidimensionality, a phenomenon that occurs when an object can be classified according to more than one characteristic.

(Meyer *et al*, 1997: 106)

Bowker (1997) define la multidimensionalidad como un fenómeno de la organización conceptual que se produce cuando los conceptos se clasifican atendiendo a varios criterios dentro del mismo sistema conceptual. La multidimensionalidad puede observarse tanto desde una perspectiva de arriba-abajo (*top-down perspective*) como desde una perspectiva de abajo-arriba (*bottom-up perspective*) (Meyer y Mackintosh, 1996: 13). En el primer caso consideramos que un concepto superordinado puede ser clasificado según diversos

puntos de vista (por ejemplo, los *vehículos* pueden clasificarse en *motorizados* o no, o en *vehículos terrestres, marítimos* o *aéreos*), mientras que en el segundo caso un concepto puede considerarse subordinado de dos o más conceptos distintos (por ejemplo, un *avión* puede considerarse un *vehículo aéreo* o un tipo de *vehículo motorizado*).

Según Bowker (1997: 137) en el origen de la multidimensionalidad puede haber diferentes necesidades sociales, diferentes niveles de percepción y cognición, y, en ocasiones, puede darse una mala interpretación de las relaciones existentes entre los fenómenos naturales. Asimismo, la multidimensionalidad puede deberse a la evolución de los conceptos y los términos.

Bowker propone distintas estrategias para gestionar la multidimensionalidad en el trabajo terminológico. Según Bowker (1997: 137), los terminólogos deberían tener en cuenta la multidimensionalidad de los ámbitos temáticos, para lo cual propone un conjunto de directrices que permiten reconocer y representar la multidimensionalidad.

Para identificarla, propone las siguientes estrategias:

- a) Trabajar con una representación gráfica, a ser posible en forma de árbol, del sistema de conceptos.
- b) Considerar todos los modos posibles de clasificar el área temática en todos los niveles de clasificación.
- c) Reconocer los indicadores de multidimensionalidad, como la repetición de conceptos o el uso de pares léxicos distintos para indicar contraste.

Para representarla, Bowker propone lo siguiente:

- a) Trabajar con las dimensiones de una en una.
- b) Ordenar las dimensiones según su importancia relativa.
- c) Distinguir las dimensiones claramente en la representación gráfica.
- d) Indicar claramente la característica que subyace en cada dimensión.

Revisar la representación gráfica del sistema de conceptos a medida que se profundiza en el conocimiento del ámbito temático.

Valoración

Sin duda, la principal aportación del sistema de gestión de bases de conocimiento terminológico CODE es que, por primera vez en el desarrollo de la gestión terminológica, se afronta la información conceptual sobre atributos y

relaciones de forma estructurada y explícita. Esto supone un avance con respecto a los sistemas de gestión de la terminología tradicionales, en los que la información sobre atributos y relaciones se encuentra dispersa en lenguaje natural en las definiciones.

Esta ventaja facilita la recuperación y consulta de la información de manera onomasiológica, es decir, a partir de la información conceptual.

Sin embargo, como señala Moreno Ortiz (2002), este sistema parece estar creado sólo para trabajar con dominios conceptuales pequeños, en los que las relaciones conceptuales pueden controlarse fácilmente. Además, tampoco se describe con claridad cuáles son las categorías de datos que se incluyen (parece que se resumen a un solo atributo), ni cuáles son sus posibilidades de intercambio de información.

3.2.2 Las redes semántico-terminológicas (Otman, 1996)

Otman (1996) propone un modelo teórico para la representación formal de las unidades terminológicas a través de lo que él denomina redes semántico-terminológicas (*réseaux sémantico-terminologiques*, RST). Dado que su objetivo es la representación formal de las unidades y las relaciones, su definición de las relaciones se aproxima a la lógica y a la ingeniería del conocimiento. Otman propone que todas las relaciones pueden representarse mediante una fórmula del tipo **R (a, b)**, donde **R** es el nombre de la relación y **a** y **b** son los conceptos entre los que se establece dicha relación.

Otman nos presenta un modelo de representación de la terminología por medio de redes organizadas de manera jerárquica, en el que cada concepto constituye un nodo que se relaciona con otros mediante relaciones verticales –que tienen una vocación clasificadora y en las que se aplica la herencia– y relaciones horizontales, que tienen una vocación descriptiva, y en las cuales no se aplica la herencia. Las relaciones tienen la doble misión de **reagrupar** las nociones con cierto grado de parentesco y **distinguir** las nociones en el seno de clases y subclases por sus rasgos distintivos.

Otman se limita a incluir sustantivos en las RST, alegando que verbos y adjetivos pueden referirse siempre a un sustantivo. Una noción importante que se introduce en las RST es la de **architérmino**, que permite agrupar bajo una sola etiqueta, es decir, en un solo nodo de la red, varias denominaciones

terminológicas que remiten a la misma noción. El architérmino es, en este contexto, una representación convencional de una serie de términos sinónimos (diversas variantes). Los conceptos designados por términos pueden ser concretos, abstractos o ficticios, y pueden designar entidades genéricas o individuales.

Según Otman (1996: 55), una relación permite establecer vínculos nocionales entre varios conceptos. Cada relación se define por medio de un nombre con el que se designa la relación (por ejemplo *tipo-de*), por el tipo de objetos que admiten dicha relación, por la existencia de ciertas relaciones asociadas a esos objetos y, en ocasiones, por la duración de su validez. Las relaciones se conciben como vectores que vinculan los conceptos, entre los que funcionan los mecanismos de herencia.

Únicamente puede haber un arco entre dos nodos cualesquiera, es decir, dos términos cualesquiera sólo están unidos por una relación semántica, aunque Otman es consciente de la restricción que esto conlleva, ya que algunos términos de un dominio no mantienen ninguna relación directa entre sí.

Para Otman, el interés de la terminología por las relaciones conceptuales está vinculado a un ámbito temático concreto y a unas necesidades concretas, y por este motivo no considera los sistemas conceptuales en sentido absoluto, sino en función de unas necesidades específicas de la comunicación especializada. Otman (1996: 101) defiende que ciertas relaciones son más pertinentes en determinados dominios que en otros, y que ciertas relaciones muy particulares sólo se conciben en dominios específicos. Su estudio se centra en las relaciones que llama primitivas porque se dan en la mayoría de los dominios.

Con el fin de representar adecuadamente toda la red de relaciones de un ámbito temático, Otman presenta un catálogo de relaciones conceptuales en el que todas ellas estén bien definidas. En toda red semántica existen relaciones, pero dichas relaciones no son únicamente "hilos" que vinculan unos conceptos con otros, sino que poseen su propia semántica, sus restricciones de uso y sus límites. Para Otman, sólo si somos capaces de describir claramente las propiedades que identifican cada relación, podremos explotarlas en un modelo de representación formal:

Il n'y a pas de réseaux sans relations. Mais les relations des réseaux sont plus que de simples liens, de simples fils qui relient les noeuds les uns aux autres. Ces relations ont leur propre sémantique, leurs restrictions

d'emploi, leurs limites. Elles possèdent aussi un certain nombre de propriétés qui permettent de les exploiter dans un modèle de représentation formel.

(Otman, 1996: 53)

En efecto, se considera que en un dominio del saber con un elevado nivel de coherencia, cada concepto establece relaciones con todos y cada uno de los conceptos del sistema (Otman, 1996: 30). Esto es lo que Otman denomina **sistema auto-referente**. Es el conjunto de enlaces y las propiedades asignadas a esos enlaces lo que determina la definición del término en la RST.

Otman (1996: 151-153) agrupa las relaciones bajo los epígrafes de **relación asociativa** y **relación distintiva**. Entre las primeras se sitúan las relaciones tipo-de, parte-de, función-de y la relación de proximidad. Entre las distintivas encontramos las de contraste-con y equivalente-a. A continuación resumimos cómo entiende Otman cada uno de estos tipos de relación.

a) Relación tipo-de

Es una relación jerárquica, siempre transitiva que relaciona hiperónimos e hipónimos. Todos los hiperónimos e hipónimos designan clases. Considera únicamente la relación de inclusión de una clase hiponímica en una hiperonímica de orden superior (se excluye la pertenencia individual), y cada término se define en relación con el hiperónimo inmediatamente superior.

En cada cadena hiperonímica, un elemento se designa como **objeto de base** y se marca como tal en la representación. Tal y como vimos que proponía Cruse (2002), Otman considera que existe un nivel taxonómico de base que se considera el más informativo y a partir de el cual se pueden identificar hiperónimos e hipónimos. Los elementos hiperonímicos situados en el extremo superior de las cadenas en un dominio se consideran elementos primitivos. Otman (1996) admite la posibilidad de hiperónimos múltiples, es decir, de que un mismo subordinado posea varios superordinados. Es decir, contempla la herencia múltiple.

b) Relación parte-de

Otman (1996) distingue cinco tipos de relaciones meronímicas, correspondientes en parte con las identificadas por Winston *et al* (1987). Son las siguientes:

- objeto/elemento

- conjunto/miembro
- masa/porción
- objeto/constituyente
- actividad/fase

Como lo que le interesa a Otman es explotar la transitividad de ese tipo de relación, se aplicará la transitividad sólo entre partes de la misma categoría, puesto que, como vimos en el apartado 2.3.1.1, la transitividad se mantiene siempre que consideremos el mismo tipo de subrelación meronímica.

c) Relación *función-de*

Este tipo de relación engloba los conceptos de efecto y fin, es decir, las relaciones que tradicionalmente se han llamado relaciones causa-efecto y aquellas que indican la finalidad de un proceso o acción. Otman las considera porque constata la utilidad que en terminología se asigna a las definiciones funcionales. La relación función-de explica la relación que existe entre una entidad (objeto físico, persona, concepto) y la acción que implica. En el modelo de Otman se adopta el principio de atribución de una denominación.

d) Relación *de proximidad*

La relación de proximidad equivale a las relaciones pragmáticas o relaciones asociativas que veíamos en las tipologías de relaciones de la TGT. Este tipo de relación permite llamar la atención sobre una relación semántica fuerte entre dos conceptos sin que sea absolutamente necesario o posible formalizarla. Propone, del mismo modo que en las relaciones de función-de, que se especifique cuál es la relación en cada ejemplo concreto.

Dentro de las denominadas relaciones distintivas, Otman identifica las relaciones de contraste y las relaciones de equivalencia, que se definen a continuación.

e) Relación *contraste-con*

Los términos se presentan en ocasiones en series que pueden mantener relaciones de oposición. En esa relación puede indicarse, como en las anteriores, cuál es el rasgo de oposición mediante una palabra.

f) Relación equivalente-a

Este tipo de relación se refiere a la sinonimia y cuasi-sinonimia. Para Otman, dos términos son sinónimos si ocupan el mismo lugar en el la red semántico-terminológica, responden a la misma definición, mantienen una relación de equivalencia en el seno de un mismo sistema semiótico y son sustituibles uno por otro en un contexto determinado. La sinonimia como relación posee las propiedades de simetría y transitividad. Asimismo, la consideración de las relaciones de equivalencia permite a Otman utilizar las redes semántico-terminológicas en entornos multilingües.

Este tipo de organización permite que el usuario especializado de la red semántico-terminológica (documentalista, indizador, traductor, redactor técnico, entre otros) interroge al sistema de manera onomasiológica, teniendo en cuenta aspectos semánticos como la sinonimia o qué términos comparten una determinada característica. La propuesta de Otman permite la extracción de una parte de la red (en forma de nudos y arcos) y su observación desde un punto de vista determinado (por ejemplo, desde el punto de vista de una relación determinada). De este modo, la propuesta de Otman permite considerar la multidimensionalidad de las relaciones entre conceptos.

Otman es consciente de que algunos parámetros no aparecen representados en su modelo de redes semántico-terminológicas, tales como el conocimiento difuso o aproximado (*fuzzy*), el tiempo y la cronología, la cuantificación, la circularidad y el plurilingüismo.

En la Fig. 6 se reproduce un ejemplo de representación gráfica de la red semántico-terminológica para el concepto de la ingeniería artificial *règle* (Otman, 1997: 248).

En esta representación gráfica los conceptos se representan como cuadros que forman nodos unidos por relaciones. En cada relación aparece una etiqueta que indica el tipo de relación que se establece (por ejemplo *sorte-de* o *partie-de*). Las puntas de flecha que aparecen en los arcos que representan las relaciones indican la direccionalidad de la relación.

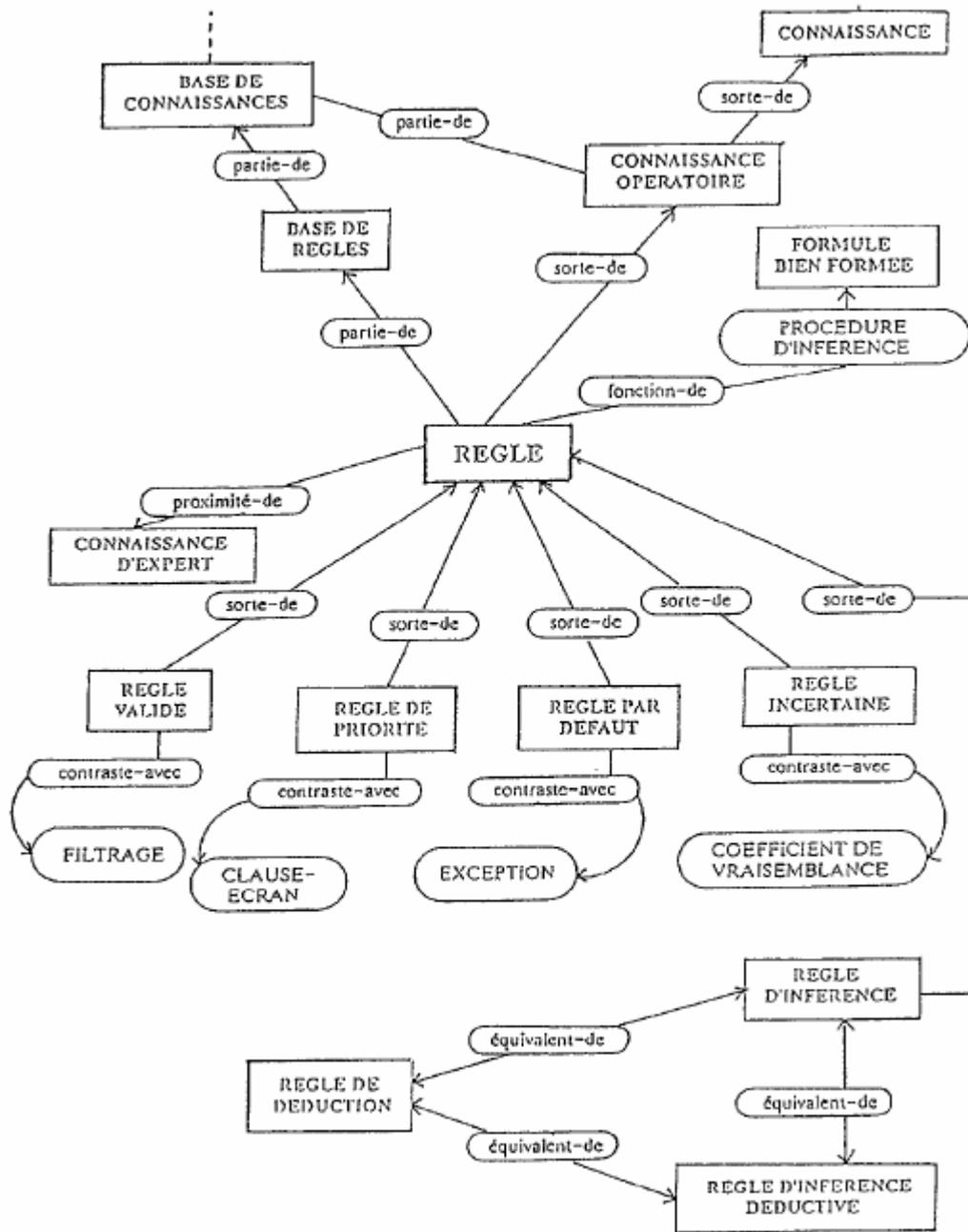


Fig. 6. Fragmento de la red semántico-terminológica del concepto *règle* (Otman, 1997: 248).

Valoración

En la propuesta de Otman destacamos su concepción de las relaciones desde un punto de vista lógico, teniendo en cuenta sus propiedades y aprovechando las posibilidades de los sistemas de herencia entre las relaciones, que garantizan la economía del sistema.

Al igual que vimos en CODE, su representación tiene como fin el permitir recuperar la información de manera onomasiológica, teniendo en cuenta las propiedades de los términos y las relaciones que se establecen entre ellos. Este tipo de consultas onomasiológicas posibilitan la consideración de la multidimensionalidad, ya que gracias a ellas podemos recuperar la información conceptual favoreciendo un punto de vista frente a otros.

Su modelo de representación conceptual resulta adecuado; sin embargo, carece de una herramienta que le permita organizar y recuperar la información conceptual de manera sencilla. Creemos que el empleo de un editor de ontologías puede contribuir a esta tarea.

3.2.3 Formalización mediante grafos conceptuales e hipertexto (Antia, 2000)

Si bien el objetivo principal de Antia (2000) es proponer un marco alternativo para la terminología orientada a la planificación lingüística, su obra aporta algunos elementos sugerentes en lo que a la formalización de las relaciones conceptuales se refiere.

Para Antia (2000: 100), las relaciones que se establecen entre los conceptos especializados tienen implicaciones para la adquisición, representación y extracción del conocimiento. Reconoce que los formatos tradicionales de representación resultan limitados a la hora de reflejar aspectos como la multidimensionalidad, la diversidad de las relaciones posibles entre conceptos (que no siempre son jerárquicas) y la naturaleza no estática del conocimiento, entre otros (Antia, 2000: 166).

Antia (2000) plantea los retos de la representación de conceptos en función de dos dicotomías que hacen referencia a dos dimensiones diferentes. Por un lado, distingue entre la representación de relaciones basadas en aspectos lingüísticos y relaciones basadas en los conceptos. Entre las primeras, Antia (2000) incluye la sinonimia absoluta y los distintos tipos de variantes denominativas (variantes ortográficas y morfológicas, entre otras). Las relaciones basadas en los conceptos, por su parte, no tienen que ver con la forma lingüística de los términos, y entre ellas se incluyen las relaciones jerárquicas y asociativas consideradas por la terminología clásica.

Por otro lado, plantea una segunda dicotomía derivada de la organización del conocimiento a nivel micro- y macroestructural. Así, en el nivel microestructural organizamos categorías de información dentro de un mismo concepto, mientras que en el nivel macroestructural consideramos las relaciones entre distintas categorías de información.

Antia (2000: 167) recurre a tres conceptos clave para plantear su hipótesis sobre la representación del conocimiento: **asociatividad**, **no linealidad** y **(re)generación**. El concepto de asociatividad se refiere a la naturaleza del conocimiento almacenado. Los elementos de conocimiento no son entidades discretas que se encuentran aisladas, sino que aparecen en agrupaciones (*chunks*), de manera que la activación de un elemento del grupo es capaz de activar otros dentro de él.

En segundo lugar, Antia mantiene que el acceso al conocimiento no se produce de modo lineal, sino que, dado que el conocimiento se encuentra interconectado de una manera no estática, existen distintas configuraciones posibles. El acceso al conocimiento comporta un cierto grado de incertidumbre o azar.

El tercer aspecto clave sobre el conocimiento especializado es la creatividad del proceso de (re)generación. El conocimiento nuevo se genera a partir de procesos de inferencia, es decir, del razonamiento deductivo ejercido sobre el conocimiento almacenado. En opinión de Antia, un formalismo de representación válido debe permitir hacer inferencias para dar cuenta de la creatividad del conocimiento.

Teniendo en cuenta estas tres consideraciones sobre el conocimiento especializado (la asociatividad, la no linealidad y la creatividad mediante inferencias), Antia propone la representación de su estructura por medio de **grafos conceptuales** e **hipertexto**, que son modelos de representación tomados de la ingeniería del conocimiento. A continuación describimos brevemente cómo se puede emplear cada uno de estos mecanismos para formalizar las relaciones conceptuales.

Los grafos conceptuales son un formalismo de representación del conocimiento desarrollado por Sowa (1984) y empleado por primera vez en

terminología por Hook y Ahmad (1992)¹⁹. Los grafos permiten representar la estructura interconectada del conocimiento de una manera flexible, con diferentes configuraciones y puntos de acceso para llegar a los mismos datos. Por otro lado, los grafos permiten hacer inferencias y, por tanto, (re)generar conocimiento nuevo.

Los grafos conceptuales se presentan en forma de cuadrados –que representan tipos de conceptos (o individuos)– y círculos, que simbolizan las relaciones conceptuales. Los conceptos y las relaciones se unen por medio de puntas de flecha que expresan la direccionalidad de la relación establecida. En la Fig. 7 se reproduce el ejemplo proporcionado por Antia (2000: 169) de un grafo conceptual complejo para el concepto jurídico *joint resolution*:

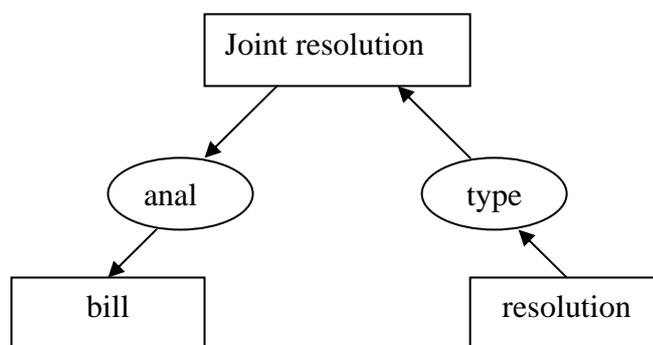


Fig. 7. Grafo conceptual complejo para el concepto *joint resolution* (Antia: 2000, 169).

La parte izquierda de este grafo conceptual se puede leer del siguiente modo: “analogy of joint resolution is bill”. La parte derecha del grafo se puede leer como “type of resolution is joint resolution”. Si leemos en la dirección opuesta de las flechas, las relaciones se pueden leer como “bill is analogous to joint resolution” y “joint resolution is type of resolution”, respectivamente.

La flexibilidad de este formalismo y su adecuación para la representación de la terminología radica en el hecho de que las características de los conceptos pueden reinterpretarse y considerarse a su vez como nodos relacionales a partir de los cuales es posible recuperar la información sobre los conceptos a los que unen.

La segunda propuesta de Antia para la representación de la estructura del conocimiento consiste en la explotación de los sistemas de **hipertexto**. El

¹⁹ Otros autores como Barrière (1997 y 2002) también plantean el empleo de grafos conceptuales para la formalización de las relaciones en terminología.

hipertexto como representación del conocimiento emplea la misma estructura de nodos y enlaces que los grafos conceptuales. La diferencia entre ambas es que, mientras que la implementación en el caso de los grafos es gráfica, la implementación del hipertexto se lleva a cabo en un entorno textual. El empleo de hipertexto en los recursos terminológicos facilita el acceso a la información de manera no lineal.

Valoración

Los grafos conceptuales resultan claramente más expresivos que los sistemas conceptuales en forma de árbol creados en la TGT. De este modo se pueden superar, en parte, las limitaciones de los formatos tradicionales.

Sin embargo, la presentación de este tipo de formalismo por parte de Antia no tiene su reflejo en una herramienta concreta. Antia hace referencia al prototipo diseñado por Hook y Ahmad (1992) pero ese prototipo no está disponible y sólo puede accederse a su descripción a través de la bibliografía.

En cuanto al hipertexto, creemos que no puede considerarse una formalización de las relaciones conceptuales en sentido estricto, sino un instrumento que permite el acceso no lineal a la información, si bien facilita en gran medida el empleo de recursos terminológicos.

3.2.4 Representación de las relaciones semánticas mediante esquemas relacionales (Oster, 2005)

Oster (2005) propone el estudio de las relaciones semánticas que se establecen entre los constituyentes de los términos complejos (formados por más de un lexema) que combinan conceptos en el ámbito de la cerámica. Para ello propone utilizar como instrumento de análisis un conjunto de relaciones intraconceptuales descritos como **esquemas relacionales**.

Las relaciones pueden describirse de un modo general válido para cualquier área de conocimiento, pero también pueden formularse de manera específica para un ámbito temático concreto. Oster define la relación semántica **R** entre dos conceptos **a** y **b** como “la combinación de las funciones que **a** y **b** desempeñan el uno con respecto al otro” (Oster, 2005: 101). Dicha definición puede expresarse mediante el formalismo:

$$R(a, b) = A - B$$

donde **A** es la función que **a** desempeña con respecto a **b** y **B** es la función que **b** desempeña con respecto a **a**.

Las **funciones semánticas** se definen en el sentido propuesto por la gramática funcional de Dik (1989), es decir, como el rol semántico que un concepto desempeña con respecto a otro. Ejemplos de funciones semánticas serían AGENTE, META o RECEPTOR, entre otros.

A pesar de que Oster se centra en las relaciones semánticas existentes en el interior de los términos, considera que también son útiles para la descripción de las relaciones que se dan entre las unidades terminológicas presentes en un texto y las relaciones entre conceptos en un sistema terminológico (Oster, 2006: 4).

Su descripción de las relaciones se basa en la noción de **esquema relacional**, definido como una estructura de conocimiento abstracta que sirve como mecanismo para reconocer e identificar los tipos de relación que unen uno o más conceptos (Oster, 2006: 6). Los esquemas son mecanismos cognitivos que facilitan el procesamiento de la información.

Oster propone una clasificación que consta de ocho esquemas básicos, divididos en esquemas más específicos. Siguiendo a Lakoff (1987), establece una primera distinción entre modelos proposicionales, modelos esquemáticos de imagen y modelos metafóricos. A continuación presentamos de manera sucinta su propuesta de clasificación de los esquemas relacionales. Los ejemplos están tomados de Oster (2005: 106-112):

A. MODELOS PROPOSICIONALES

1. **Esquema de acción:** un médico (AGENTE) opera (ACCIÓN) a un paciente (META)

2. Extensiones del esquema de acción

2.1 *Esquema instrumental:* Un médico (AGENTE) opera (ACCIÓN) al paciente (META) empleando un bisturí (INSTRUMENTO).

2.2 *Esquema final:* en el esquema anterior, la finalidad podría ser extirpar un cuerpo extraño.

3. **Esquema de proceso:** en el proceso de combustión, la entidad PROCESADA es el material que se quema.

4. Extensiones de los esquemas de acción y de proceso

4.1 *Esquema causal*: El calentamiento global (CAUSA) puede tener como efecto la desertización (EFECTO)

5. Esquemas de estado

5.1 *Esquema atributivo*: un gato (CERO) negro (ATRIBUTO)

5.2 *Esquema de identificación*: este plan (CERO) es una locura (ENTIDAD ESPECIFICADORA)

5.3 *Esquema de oposición*: largo (CERO)-corto (ENTIDAD OPUESTA)

5.4 *Esquema de localización espacial*: el barco (CERO) cruza el río (LUGAR)

5.5 *Esquema de localización temporal*: El próximo concierto (CERO) será el viernes (TIEMPO)

6. **Esquema de posición**: relación que une a un perro (META) con su dueño (POSITOR)

B. MODELOS ESQUEMÁTICOS DE IMAGEN

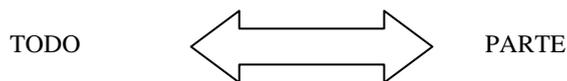
7. **Esquema contenido-contenedor**: vino (CONTENIDO)-botella (CONTENEDOR)

8. **Esquema partitivo**: bicicleta (TODO)-rueda (PARTE)

C. MODELOS METAFÓRICOS

9. **Esquema de analogía**: la expresión *corriente eléctrica* remite a la relación de analogía entre el DOMINIO ORIGEN de los líquidos y el DOMINIO META de la electricidad.

Su propuesta se amplía considerablemente al aplicarla a los términos compuestos específicos del ámbito de la cerámica. No reproducimos aquí la clasificación completa puesto que excede los límites de nuestra presentación teórica, pero reproducimos un ejemplo prototípico de los esquemas relacionales que Oster identifica en los términos del ámbito de la industria cerámica, que nos servirán para ilustrar su propuesta. Tomamos el esquema partitivo, en el que se identifican las relaciones más específicas COMPONENTE FUNCIONAL – TODO, MIEMBRO – TODO, MATERIAL – TODO y COMPONENTE ESENCIAL – TODO. Ejemplos de términos polilexemáticos del ámbito de la industria cerámica en los que se identifican estas relaciones son *horno a rodillo*, *rueda dentada*, *malla metálica* y *fundente de zinc*. El esquema relacional se puede representar gráficamente del siguiente modo:



Valoración

Los esquemas relacionales constituyen un instrumento flexible para la representación y clasificación de las relaciones semánticas que se establecen entre los constituyentes de los términos complejos. Consideramos que también pueden serlo para la representación y clasificación de las relaciones entre conceptos, al menos en el nivel más general, que no depende del ámbito temático en el que nos movemos.

Además, dado que nuestro estudio empírico también se realiza sobre términos de la industria cerámica, las relaciones específicas propuestas por Oster nos pueden ser de ayuda para la formalización de las relaciones conceptuales del mismo ámbito utilizando un editor de ontologías.

No obstante, la perspectiva desde la que Oster aborda su trabajo difiere de la nuestra, ya que estudia las relaciones semánticas en el interior de los términos compuestos en alemán y en español con fines contrastivos.

3.2.5 Aplicación de las funciones léxicas a la terminografía basada en corpus

Marie-Claude L'Homme (2004) y otros colaboradores del OLST (Observatoire de Linguistique Sens-Texte) de la Universidad de Montreal proponen la aplicación de un enfoque léxico-semántico a la estructuración de la terminología que permita interpretar los datos procedentes de corpus electrónicos para la construcción de diccionarios especializados.

L'Homme recurre a la **Teoría del sentido-texto** (*Meaning-Text Theory, MTT*) propuesta por Mel'cuk (1997), y en concreto al componente lexicológico de dicha teoría, la *Explanatory and Combinatorial Lexicology* (ECL).

La teoría del Sentido-Texto es una teoría lingüística que pretende describir la correspondencia Sentido \Leftrightarrow Texto por medio de la construcción de modelos formales. Un modelo Sentido \Leftrightarrow Texto es una máquina virtual cuyo *input* son las representaciones de sentidos enunciados y cuyo *output* es un conjunto de textos que contienen todas las paráfrasis que permiten explicar el sentido dado en la entrada. Como afirma Polguère (1998), la teoría sentido-texto puede

considerarse como una *máquina virtual*²⁰ que permite traducir los sentidos a enunciados, denominados textos.

A diferencia de lo que ocurre en otros modelos lingüísticos formales, en la Teoría del Sentido-Texto el diccionario de una lengua se considera el corazón del modelo de dicha lengua. El *Explanatory and Combinatorial Dictionary* (Diccionario Explicativo y Combinatorio, ECD) es un lexicón formalizado de orientación semántica. Se dice que el ECD es un diccionario **explicativo** porque describe el sentido de una unidad léxica por medio de una definición analítica que descompone dicho sentido en sus elementos constitutivos. Se dice que es un diccionario **combinatorio** en la medida en que describe la combinatoria sintáctica (régimen) y léxica (colocaciones) asociada a cada unidad léxica. La combinatoria léxica se describe por medio de las **funciones léxicas**, que definiremos más adelante.

En cuanto al tipo de relaciones entre palabras representadas en el ECD, en este diccionario se representan tres tipos básicos de relaciones, que son las **relaciones semánticas** (paradigmáticas), las **relaciones sintácticas** (sintagmáticas) y las **relaciones léxicas** (tanto paradigmáticas como sintagmáticas). Las primeras se expresan por medio de la definición (zona semántica de la entrada), y abarcan la sinonimia y la proximidad semántica, entre otras. La definición formula un significado discreto por medio de conceptos elementales, de modo que cada sentido de una palabra puede descomponerse semánticamente.

Las relaciones sintácticas se representan en la zona sintáctica por medio de patrones de régimen sintáctico (*government patterns*) que se establecen entre una palabra y otras que pueden aparecer en la misma oración.

Las relaciones léxicas se recogen en la zona léxica de la entrada y se establecen entre la palabra que aparece en la entrada y aquellas que pueden sustituirla en un texto, o bien unirse a ella en combinaciones de palabras más o menos fijas. Este tipo de relaciones implican lo que se conoce como **funciones**

²⁰ El concepto de *máquina virtual* proviene de la informática, y en concreto del lenguaje Java. Una máquina virtual es un tipo de software que funciona como interfaz entre el código binario del compilador y el microprocesador que ejecuta las instrucciones del programa. Permite construir programas con independencia de la plataforma que se utilice.

léxicas. Dado que son éstas las relaciones que nos pueden ser de mayor utilidad en nuestro enfoque, vamos a describirlas y definir las con cierto grado de detalle.

Una función léxica (*lexical function, LF*) se define como una función en sentido matemático que representa una idea muy general, tal como *muy*, *empezar* o *implementar*, o bien un cierto rol semántico-sintáctico. Mediante la función léxica se relaciona una palabra clave con el conjunto de palabras y frases que expresan el significado o rol que corresponde a dicha función. Dentro de las funciones léxicas se distingue entre **funciones léxicas básicas** (como la sinonimia, la antonimia, el concepto genérico, la función, etc.), **funciones no estándar** (que tienen un significado demasiado específico u se dan sólo en casos muy particulares) y **funciones léxicas compuestas** (son combinaciones sintácticas de funciones léxicas básicas relacionadas entre sí).

Mel'cuk y Zholkovsky (1988) identifican cerca de sesenta funciones léxicas. La presentación exhaustiva de todas ellas excede los límites de este trabajo, pero aun así creemos conveniente presentar algunas funciones léxicas (hemos limitado su número a siete) que podrían ser relevantes para la representación de relaciones conceptuales entre términos.

Símbolo de la FL	Definición	Ejemplo
Syn ²¹	Sinónimo	Syn (shoot) = fire
Gener	Concepto genérico	Gener (gas) = substance
A_i	Propiedad estándar de los participantes en una situación dada.	A_i (speed) = with a speed of
Loc	Preposición que denota la localización estándar en el espacio	Loc (height) = at a height
Caus (normalmente ocurre en composición con otra LF)	Causar, ocasionar que ocurra alguna situación	CausFunc (dinner) = prepare [something for dinner]
Sympt	Síntoma	Sympt (fear, hair) = his hair stood on end
Result	Resultado	Result (buy) = have

Tabla 9. Funciones léxicas estándar que pueden aplicarse en terminología (Mel'cuk y Zholkovsky, 1988)

Veamos a continuación cómo pueden aplicarse las funciones léxicas para la formalización de relaciones conceptuales entre términos.

²¹ Mel'cuk y Zholkovsky (1988) distinguen entre sinónimos con un significado más amplio, más limitado o con significados que se interseccionan.

L'Homme (2004) analiza el mismo concepto (*program*) a partir de su aparición en un corpus textual desde una perspectiva conceptual clásica de la terminología y desde la perspectiva léxico-semántica. A partir de este análisis, la autora extrae las siguientes conclusiones sobre los términos:

- a) Los términos pueden expresar múltiples significados.
- b) Los términos pueden contraer una gran variedad de relaciones con otros términos, y no únicamente relaciones jerárquicas (hiperonimia y meronimia). La comprensión de estas relaciones permite identificar los distintos significados de los términos.
- c) Cada relación es válida para un significado concreto del concepto.

L'Homme extrae asimismo conclusiones sobre la naturaleza y frecuencia de aparición de las relaciones:

- a) La mayor parte de las relaciones conceptuales identificadas son relaciones no jerárquicas, como relaciones causa-efecto, relaciones de instrumento o relaciones entre un concepto y sus propiedades.
- b) Algunas relaciones se producen entre unidades léxicas que no son sustantivos. Por ejemplo, las acciones pueden ser expresadas mediante verbos y las propiedades por medio de adjetivos.
- c) La mayoría de las relaciones implican términos considerados como unidades lingüísticas y no como etiquetas para los conceptos.

En definitiva, los términos se comportan como otras unidades léxicas, y por tanto deben ser tratados como tales.

L'Homme establece cuatro implicaciones del empleo de las funciones léxicas para la terminología computacional. En primer lugar, en un enfoque léxico-semántico las unidades consideradas pueden pertenecer a distintas clases de palabras, no únicamente sustantivos. Tradicionalmente la terminología se ha ocupado de las unidades no predicativas (sobre todo sustantivos), dejando de lado las unidades predicativas, como los verbos, adjetivos, adverbios y ciertos sustantivos que se refieren a acciones y propiedades. Al tener en cuenta estas unidades no predicativas como actores semánticos (*semantic actants*, es decir, expresiones que se refieren a los participantes en la unidad léxica descrita), dichas unidades pueden describirse de un modo más sistemático.

En segundo lugar, el hecho de que el trabajo terminológico propuesto por L'Homme esté basado en corpus conlleva inevitablemente la aparición de polisemia, que desde el enfoque léxico-semántico se contempla como un fenómeno natural de las unidades léxicas. El enfoque léxico-semántico es mucho más compatible con el estudio de datos procedentes de corpus, puesto que en los textos el conocimiento se representa por medio de unidades lingüísticas.

La tercera implicación del empleo del enfoque léxico-semántico en el trabajo terminológico es que las relaciones semánticas se asignan a los sentidos de los términos distinguidos previamente.

Por último, el enfoque léxico-semántico permite dar cuenta de una gran variedad de relaciones que pueden aplicarse a términos que designan tanto entidades como actividades o propiedades. La principal ventaja que conlleva el empleo de funciones léxicas para la representación de relaciones semánticas entre términos es que permiten capturar una amplia gama de relaciones sin tener que recurrir a distintos modelos (L'Homme, 2007: 34).

Una vez vistas las implicaciones que tiene el empleo del enfoque léxico-semántico para el trabajo terminológico basado en corpus, veremos un ejemplo de aplicación de las funciones léxicas para la representación de las relaciones entre términos.

Dancette y L'Homme (2002) aplican las funciones léxicas de Mel'cuk a la representación de relaciones taxonómicas y partitivas como puerta de acceso al conocimiento en un diccionario multilingüe especializado sobre la distribución (*Dictionnaire analytique de la distribution / Analytical Dictionary of Retailing*, Dancette y Rhetoré, 2000). Estas dos autoras se centran en las relaciones jerárquicas porque son las que se consideran más relevantes en terminología, ya que permiten acceder a la organización jerárquica del conocimiento. Su objetivo es hacer explícitas las relaciones semánticas entre términos que aparecen en el diccionario y permitir la distinción de matices de significado dentro de un término.

Las funciones léxicas pueden representarse mediante la fórmula

$$f(x)=y$$

donde **f** es la propia función, **x** es la palabra clave e **y** el valor de la función léxica. Por ejemplo, la fórmula **Anti**(durable good) = non durable good representa la relación de antonimia entre estos *durable good* y *non durable good*,

Gener(private brand) = brand representa la relación hiponímica establecida entre *brand* y *private brand*.

La clasificación de las relaciones semánticas por medio de funciones léxicas permite distinguir distintos tipos de relación entre la entrada del diccionario y otras unidades presentes en un mismo artículo, proporcionando al mismo tiempo distintas vías de acceso a las entradas del diccionario desde los artículos correspondientes a términos relacionados.

Sin embargo, las funciones léxicas estándar propuestas por Mel'cuk no incluyen las relaciones taxonómicas y partitivas, ya que Mel'cuk considera que se refieren al conocimiento enciclopédico y por tanto no son funciones propiamente dichas. Para superar este inconveniente, Dancette y L'Homme proponen adaptar el modelo original con el fin de formalizar este tipo de relaciones, fundamentales en terminología.

Dancette y L'Homme (2004) establecen una distinción entre las relaciones léxico-semánticas consideradas **relaciones terminológicas típicas** o clásicas, las **relaciones argumentales y circunstanciales** y **otras relaciones** (propiedades, unidades de medida, utilidad). Además de estas relaciones paradigmáticas, Dancette y L'Homme proponen el estudio en terminología de las **relaciones sintagmáticas y derivacionales**. Las relaciones sintagmáticas, que se establecen entre sustantivos y verbos, o bien entre sustantivos y adjetivos o adverbios, se consideran mucho menos productivas en las terminologías.

Entre las relaciones terminológicas típicas o clásicas estas autoras incluyen las relaciones de sinonimia y antonimia, las relaciones taxonómicas y las relaciones meronímicas.

Las funciones léxicas que permiten formalizar las relaciones de hiperonimia/hiponimia son la función **Gener** y su inversa, la relación **Spec**. Mediante estas funciones pueden identificarse taxonomías y establecerse jerarquías en el diccionario. No obstante, no permiten distinguir entre las diferentes dimensiones conceptuales (facetas). Ejemplos de estas funciones son los siguientes:

- **Gener**(debit card) = magnetic card
- **Spec**(store layout) = free-flow layout

Los cohipónimos, por su parte, pueden representarse por medio de las funciones léxicas **Syn**∩ y **Anti**. La primera representa la relación que existe entre

unidades léxicas que tienen componentes semánticos en común y componentes semánticos que los diferencian, como por ejemplo **Syn** \cap (convenience good) = unsought good. La función léxica **Anti** representa la relación entre los hipónimos que mantienen una relación de oposición (ej. **Anti**(horizontal integration) = vertical integration). Estas funciones relacionan términos que se encuentran en el mismo nivel de la jerarquía con respecto al mismo superordinado y aclaran la diferencia entre ellos.

En lo que se refiere a las relaciones partitivas, las dos funciones léxicas estándar permiten capturar uno de los subtipos de relaciones parte-todo. Se trata de las funciones **Mult** y su inversa **Sing**. **Mult** se refiere a una colección y **Sing** se refiere a un miembro dentro de una colección. Ejemplos extraídos por las autoras del diccionario de la distribución son **Mult**(item) = // assortment y **Sing**(retailer cooperative) = // independent retailer.

Dancette y L'Homme son conscientes de que existen otros tipos de relación partitiva que las funciones léxicas estándar no tienen en cuenta. Para superar esta limitación, proponen tres funciones léxicas adicionales que permiten capturar la relación entre un todo y un componente funcional, la relación entre un proceso y sus fases y la relación entre un todo y las partes ocasionales (es decir, que pueden aparecer o no). A continuación ofrecemos un ejemplo de cada una:

- Relación entre un componente funcional y el todo:
 - **Part**(label) = bar code
- Relación entre un todo y una parte opcional u ocasional:
 - **Part_{occ}**(department store) = post office
- Relación entre un proceso y sus fases:
 - **Phase**(adoption process) = awareness

Las relaciones argumentales y circunstanciales se definen como relaciones paradigmáticas que se establecen entre las clases nominales. Este tipo de relaciones se dan entre los predicados y sus argumentos (relaciones argumentales), o bien indican las circunstancias en las que se produce un predicado (relaciones circunstanciales). Para dar cuenta de las relaciones argumentales y circunstanciales, Dancette y L'Homme destacan, entre otras, las siguientes funciones léxicas:

- **Obj**, que relaciona una acción y el objetivo que persigue dicha acción
- **Recip**, que relaciona una acción con el receptor de la misma
- **Ag**, que indica quién realiza una acción o es responsable de ella
- **Instr**, que relaciona una acción y el instrumento con que se lleva a cabo
- **Med**, que relaciona una acción con el medio para llevarla a cabo
- **Loc**, que relaciona una acción con el lugar en el que suele ocurrir

Bajo la denominación “otras relaciones”, Dancette y L’Homme recogen las propiedades, las unidades de medida, la relación entre una acción y su resultado, las relaciones causales y la utilidad de un instrumento. Las propiedades se refieren a las características técnicas asociadas a la definición de un concepto.

Otro aspecto propuesto por L’Homme (2007) tiene que ver con la asignación de etiquetas semánticas (*semantic labels*) a las unidades léxicas especializadas para representar las distintas clases conceptuales. Las etiquetas semánticas proporcionan una primera identificación de la categoría conceptual a la que pertenece una unidad léxica. Estas etiquetas se definen en una lista finita construida para un vocabulario específico. Así, por ejemplo, en un ámbito temático se pueden definir etiquetas semánticas para algunas categorías conceptuales, como ACTIVITÉ HUMAINE o LIEU.

Dancette y L’Homme concluyen que las funciones léxicas estándar son un instrumento útil para la representación de algunas de las relaciones que se establecen entre los términos tal como las define Mel’cuk, mientras que otras requieren algún tipo de adaptación a las necesidades del trabajo terminológico. Sin embargo, si bien las funciones léxicas tienen una demostrada utilidad a la hora de hacer explícitas las relaciones semánticas, no permiten representar la información enciclopédica y los aspectos culturales que reflejan los matices entre significados, inevitablemente presentes en los diccionarios especializados basados en corpus. El acceso a las oraciones expresadas en lenguaje natural a partir de las cuales se extraen las relaciones podría ser un modo de superar esta dificultad.

Valoración

Las funciones léxicas proporcionan un modelo para formalizar un gran número de relaciones semánticas entre términos. En la aplicación de Dancette y

L'Homme sirven para hacer explícitas las relaciones en un diccionario, proporcionando así distintos puntos de acceso a la información conceptual contenida en él.

Además, las funciones léxicas tienen la ventaja de que pueden aplicarse a distintas clases de palabras, y no se limitan únicamente a los sustantivos. Asimismo, se aplican de forma independiente a cada uno de los sentidos del término, lo que permite distinguir matices de significado.

La división de las relaciones paradigmáticas no jerárquicas en relaciones argumentales, relaciones circunstanciales y otras relaciones nos parece muy acertada, ya que permite poner orden entre este tipo de relaciones, que tradicionalmente se ofrecen en los manuales de terminología como una especie de *cajón de sastre* en el que las relaciones se recogen sin ninguna estructuración. En la elaboración de nuestro catálogo de relaciones recogeremos esta división de las relaciones no jerárquicas en argumentales y circunstanciales y otras relaciones.

Un posible aspecto negativo de las funciones léxicas es que la estructura del formalismo se basa en la introducción de abreviaturas que no siempre resultan transparentes para el usuario de la obra terminográfica. Otro posible inconveniente de esta propuesta es que la lista de funciones léxicas es muy amplia, por lo que el formalismo podría dejar de ser manejable por contener un excesivo número de funciones léxicas demasiado específicas.

3.3 Formalización de las relaciones conceptuales en terminología por medio de ontologías

En la sección anterior hemos visto algunos modelos para la formalización de las relaciones conceptuales desarrollados en los últimos quince años con el fin de superar las limitaciones de los sistemas de representación de conceptos propuestos por la TGT. En los últimos cinco años, diversos proyectos terminológicos se han fijado en la ingeniería del conocimiento y en las formas de representación empleadas en la inteligencia artificial para dar cuenta de la organización conceptual con diversas aplicaciones. Una de estas formas de representación que en los últimos años ha cobrado auge son las ontologías. Dedicaremos todo el capítulo 4 a la revisión en profundidad de lo que se ha dado en llamar ingeniería ontológica (Gómez Pérez *et al*, 2004), pero de momento

definiremos ontología en un sentido amplio como un conjunto de conceptos organizados de manera jerárquica, representados en un sistema informático, que puede servir de soporte para diversas aplicaciones (Moreno Ortiz, 2002: 31).

En esta sección presentamos algunos de los proyectos terminológicos que han empleado ontologías para formalizar la información sobre los conceptos y sus relaciones. Se trata de los proyectos ONCOTERM (3.3.1), CAOS (3.3.2), el modelo de la termontografía (3.3.3) y el proyecto GENOMA-KB (3.3.4). Por último, dedicaremos el apartado 3.3.5 a analizar las posibles ventajas e inconvenientes del empleo de ontologías en el proceso terminográfico.

3.3.1 ONCOTERM (Faber y Jiménez, 2002)

El objetivo global del proyecto ONCOTERM, desarrollado en la Universidad de Granada bajo la dirección de la Dra. Pamela Faber, es la elaboración de un sistema de información específico del subdominio de la oncología (Grupo de Investigación ONCOTERM, 2002: 221). Sus objetivos operativos van desde la creación de un corpus de textos médicos hasta la elaboración de una base de datos terminológica articulada en torno a una ontología del evento médico desarrollada con el programa ONTOTERM®.

Una de las principales contribuciones del proyecto ONCOTERM es la especificación de **esquemas categoriales** (Faber *et al*, 2001: 169) que reflejan la organización interna de cada una de las categorías del evento médico. El empleo de dichos esquemas garantiza la consistencia de las definiciones terminográficas, de la ontología y de la base de datos terminográfica elaborada en el marco del proyecto.

Las bases teóricas sobre las que se asienta este proyecto (Faber y Tercedor, 2001) proceden de la Lingüística, y en concreto de la Lexicología. Son, fundamentalmente, el **Modelo Lexemático Funcional** (*Functional Lexematic Model*, MLF) de Martin Mingorance y en las funciones léxicas propuestas desde la **Teoría del Sentido-Texto** (*Meaning Text Theory*, MTT) propuesta por Mel'cuk, autor al que ya hemos hecho referencia en el apartado 3.2.5.

El MLF propone un enfoque del léxico basado en la distinción entre relaciones paradigmáticas y sintagmáticas. Desde el punto de vista paradigmático, los términos se organizan de modo onomasiológico en una jerarquía de dominios y subdominios. En este enfoque relacional, se acepta que existen propiedades

comunes que unen a los conceptos y propiedades que los diferencian. Esto permite una organización conceptual en forma de red, en la que los conceptos se ubican en nodos enlazados mediante distintos tipos de relaciones. En el MLF los conceptos se organizan a partir del análisis de definiciones terminográficas bien estructuradas propuestas por los especialistas. Así, las definiciones de los especialistas adquieren un papel relevante en la detección de las relaciones jerárquicas y no jerárquicas existentes entre los conceptos. Las unidades cuyas definiciones poseen el mismo *genus* pertenecen al mismo dominio o subdominio. Dicho *genus* debe mantenerse consistente en las jerarquías de términos relacionados.

El MLF permite sistematizar el conocimiento experto por medio de definiciones y formalizar las relaciones jerárquicas existentes entre términos (hiponimia y meronimia). Para la formalización de las relaciones no jerárquicas, el proyecto ONCOTERM se apoya en la Teoría del Sentido-Texto. Como hemos visto en el apartado 3.2.5, la Teoría del Significado del Texto propuesta por Mel'cuk permite codificar las relaciones paradigmáticas y sintagmáticas por medio de las funciones léxicas, a las que ya hemos definido como una correspondencia que asocia una expresión léxica con un conjunto de elementos léxicos que expresan un significado específico asociado con la función.

Entre las relaciones sintagmáticas (es decir, las que formalizan una relación semántica entre dos lexemas que se ejemplifica en una cadena textual de manera no predecible), Faber y Tercedor (2001: 200) mencionan las relaciones que permiten codificar grados y las que codifican la relación instrumental o la función, entre otras. Las funciones léxicas sintagmáticas pueden servir para describir las colocaciones y las frases en terminología y constituyen una base lingüística para la representación del conocimiento por medio de ontologías y bases de conocimiento, así como para la generación textual. Otras funciones léxicas especialmente útiles para la descripción de las unidades terminológicas son las de instrumento, lugar, modo, medio y resultado.

El proyecto ONCOTERM combina de estas dos propuestas teóricas para describir y estructurar los conceptos en varios niveles, reflejando su multidimensionalidad.

Otro fundamento teórico importante que se ha tenido en cuenta en ONCOTERM es la consideración del evento médico como una estructura

conceptual formada por macrocategorías conceptuales organizadas en estructuras de memoria más amplias: los los **marcos** (en el sentido de Fillmore, 1982)²². En el evento médico tal y como se concibe en ONCOTERM existen distintos tipos de agente que desencadenan distintos procesos, en los que se puede requerir el uso de diversos instrumentos y la realización de distintas actividades u operaciones en diversos lugares. Faber (2002: 7) representa el evento médico de forma gráfica tal y como se muestra en la Fig. 8:

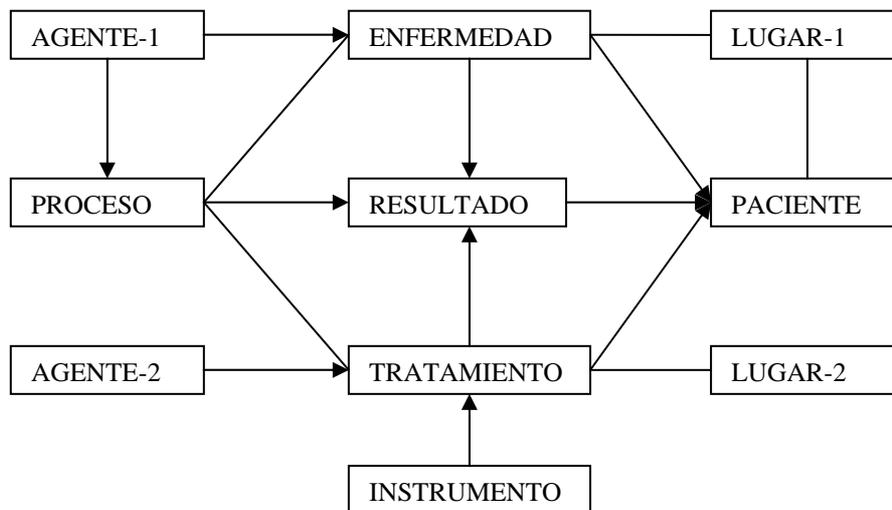


Fig. 8. Marco conceptual de referencia: representación del evento médico (Faber, 2002: 7).

Desde el punto de vista metodológico, el proyecto ONCOTERM defiende el análisis de los términos *in vivo* –esto es, tal y como aparecen en los textos–, *in vitro* –es decir, el concepto como suma de la información que potencialmente puede activar en el discurso– e *in situ*, esto es, el término se considera como parte de un evento, que forma parte de estructuras de conocimiento mayores. Por ello, su análisis se realiza a partir de un corpus de textos reales con distintos niveles de especialidad, combinando los enfoques *top-down*, es decir, tomando como punto de partida un inventario básico de términos y sus definiciones y *bottom-up*, es decir, a partir de las concordancias de los términos del corpus (Faber, 2002). Al

²² Un término evoca el marco conceptual completo, y lo mismo puede suceder con una imagen, de ahí que en el proyecto PUERTOTERM se esté estudiando la posibilidad de emplear imágenes que evocan marcos (Faber *et al*, 2006)

combinar estos dos enfoques el listado de términos no está cerrado *a priori*, sino que se combina la información procedente de las obras terminológicas con la información contenida en el corpus textual.

Una de las principales aportaciones del proyecto ONCOTERM es el empleo de ONTOTERM®, un sistema gestor de bases de datos terminológicas basado en el conocimiento. Moreno Ortiz (2002: 29) define en los siguientes términos el concepto para los fines de la representación del conocimiento por medio de una ontología:

Un concepto es una abstracción de un conjunto de objetos, propiedades o eventos existentes en el mundo real o un mundo posible, que puede poseer una realización física en una lengua natural o sistema de representación determinados, al cual se puede hacer referencia mediante un símbolo arbitrario, aunque necesariamente único, dentro de un sistema representacional. Como constructo, posee ciertas propiedades distintivas de los demás conceptos, con los que guarda diversos tipos de relaciones. Tanto sus propiedades intrínsecas como sus relaciones con los demás conceptos deben ser evidentes, y por tanto susceptibles de ser especificados de forma explícita.

(Moreno Ortiz, 2002: 29)

En esta definición destacan los siguientes aspectos:

- a) Un concepto es la abstracción de las propiedades de los objetos, propiedades y eventos de la experiencia.
- b) Los conceptos tienen un símbolo arbitrario que los representa en una lengua o sistema de representación
- c) Los conceptos tienen propiedades que los distinguen de otros, con los que establecen relaciones
- d) Tanto las propiedades como las relaciones pueden ser especificadas de manera explícita

ONTOTERM® hace posible relacionar de forma directa conceptos de la ontología con términos contenidos en una o varias bases de datos terminológicas.

Moreno Ortiz destaca las siguientes características de ONTOTERM® (Moreno Ortiz, 2002: 39):

- a) **Representación del conocimiento:** El gestor de ontologías de ONTOTERM® permite la creación y navegación por los elementos de conocimiento altamente estructurados susceptibles de ser utilizados en diversas tareas.

- b) **Escalabilidad:** ONTOTERM® está enfocado a la creación de ontologías de gran tamaño, en las que se pueda trabajar con un número muy elevado de nodos y relaciones entre nodos.
- c) **Flexibilidad:** El usuario es quien decide el nivel de complejidad que quiere otorgar a su sistema conceptual, por lo que la aplicación impone un número lo más reducido posible de restricciones.
- d) **Integridad:** La aplicación, y no el usuario, es la responsable de mantener en todo momento la integridad y la coherencia interna de la información que contiene.
- e) **Compatibilidad:** Tanto el esquema de representación utilizado como el formato de almacenamiento deben ser lo más estándar posible. El sistema permite la exportación de información a sistemas conceptuales extendidos.
- f) **Modularidad:** El sistema de representación conceptual es independiente del sistema de gestión terminológica, con el que interactúa de diversas maneras.
- g) **Facilidad de manejo:** El interfaz de usuario permite realizar todas las operaciones de forma fácil y cómoda.
- h) **Facilidad de acceso:** El sistema permite la navegación eficaz por medio de varios mecanismos.

Como hemos visto, la modularidad de ONTOTERM® queda establecida en dos módulos diferenciados dentro del sistema, el editor de ontologías y el editor de la base de datos terminológicos, que están interconectados.

Dentro del **editor de ontologías** pueden añadirse o modificarse nuevos conceptos, seleccionando el tipo de concepto de que se trata (objeto, evento, atributo o relación) y el tipo de relación que le une a su superordinado (IS-A o INSTANCE-OF). Cabe resaltar que los atributos y relaciones también se consideran conceptos dentro de la ontología, y que deben haber sido definidos previamente dentro de la misma. Al añadir un concepto, podremos indicar cuáles son sus propiedades (atributos) e indicar cuáles son los conceptos relacionados con él, explicitando el tipo de relación que los une.

ONTOTERM® permite tanto la **herencia exclusiva** como la **herencia acumulativa**. En la herencia exclusiva el concepto subordinado posee las

relaciones y atributos que pertenecen al propio concepto y también los heredados de sus conceptos superordinados directos, mientras que la herencia acumulativa contempla, además de la herencia exclusiva, todas las relaciones y atributos heredados de la ruta de todos sus hiperónimos, es decir, no se limita al superordinado directo, sino que se remonta a todos los conceptos superordinados en la jerarquía. Contempla también la herencia múltiple, que, como hemos visto en el apartado 2.3.1.7, consiste en asignar a un concepto más de un concepto superordinado, de los cuales hereda sus propiedades.

ONTOTERM® ofrece dos herramientas para la navegación y publicación de la ontología, que son el *Ontology Navigator* y el generador de informes.

El editor de la base de datos terminológica de ONTOTERM® (*Termbase Editor*) es el módulo en el que el terminólogo describe las categorías lingüísticas, de uso y administrativas de los términos. No puede introducirse un término en la base de datos terminológica si previamente no se ha introducido su concepto correspondiente en el editor de la ontología. De este modo se garantiza la orientación conceptual del trabajo terminológico.

Las categorías de datos que se incluyen en el editor de la base de datos se corresponden con las propuestas realizadas en el marco del CLS Framework (Concept-Oriented with Links and Shared references). Así, la base de datos se estructura en tres partes. La primera contiene información global sobre la base de datos (encabezado), la segunda contiene las entradas terminológico-conceptuales (cuerpo) y la tercera agrupa las referencias compartidas por varias entradas o partes de una entrada (información compartida). Dentro del cuerpo de las entradas, agrupamos los términos de una misma lengua en *langsets*. A diferencia de otros sistemas, en ONTOTERM® todos los atributos gozan del mismo estatus, por lo que podemos asignar atributos a los propios atributos.

En la Fig. 9 se muestra una captura de pantalla del término *abdominal-cancer* tal y como aparece recogido en la base terminológica de la oncología fruto del proyecto ONCOTERM. En ella se ilustran tanto la estructura conceptual como la información lingüística.

Generated by [OntoTerm](#). Updated: 28/08/2002 3:23:22

ABDOMINAL-CANCER

subject field: medicine: disease
 origination date: 10/08/2000
 originator: Clara
 URL: <http://www.mucos.cz/eng/onko/snbaen.html>

originator: Arianne
 origination date: 12/12/2000

check date: 25/02/2002
 checker: Pamela

Conceptual Structures	
ISA	MALIGNANT-NEOPLASM-BY-BODY-PART
ANCESTORS	ALL CHANGE-STATE CHANGE-STATE-ANIMAL DISEASE-EVENT EVENT HAVE-NEOPLASM LIVING-EVENT MALIGNANT-NEOPLASM MALIGNANT-NEOPLASM-BY-BODY-PART PHYSICAL-EVENT

English	
abdominal cancer	term type: main entry term part of speech: noun number: singular reliability code: 10 definition: different cancers originating in the area of the body between the diaphragm and pelvic bones. The main types of this disease are: colon/rectal cancer, bladder cancer, kidney cancer, stomach cancer, liver cancer, pancreatic cancer and small intestine cancer. <i>(en)</i>

Spanish	
cáncer abdominal	term type: main entry term part of speech: noun gender: masculine number: singular definition: cáncer que se origina en la cavidad formada entre el diafragma y la pelvis. Agrupa los siguientes tipos: cáncer colorrectal, cáncer de vejiga, tumores renales, cáncer de estómago, cáncer hepático, cáncer de páncreas y cáncer de intestino delgado. <i>(es)</i> reliability code: 9

Fig. 9. El concepto *abdominal-cancer* en la base de datos del proyecto ONCOTERM.

La interfaz de usuario del *Termbase Editor* resulta innovadora porque en ella están disponibles todas las categorías de datos, pero es el usuario quien decide

cuándo utilizarlas, siempre y cuando el sistema se lo permita. El usuario decide asimismo la complejidad de las entradas terminológicas.

Para finalizar, detallaremos las ventajas que, en opinión de su autor (Moreno Ortiz, 2002: 69), proporciona el sistema ONTOTERM®:

- a) El empleo de un sistema de representación estándar garantiza la reutilización de los datos y la compatibilidad.
- b) ONTOTERM® proporciona un interfaz gráfico de usuario que facilita las labores de compilación, administrativas y de publicación de datos.
- c) La modularidad del sistema permite trabajar de manera independiente con el sistema de representación conceptual o con el sistema de gestión terminológica.
- d) ONTOTERM® está orientado a la creación de bases de datos a gran escala.
- e) ONTOTERM® es un sistema flexible en el que el usuario decide la complejidad que deben tener su sistema conceptual y sus entradas terminológicas.
- f) ONTOTERM® garantiza la integridad de los datos, puesto que permite bloquear las operaciones inconsistentes.

Otra de las ventajas fundamentales de ONTOTERM®, señalada por Pérez (2002) es que las relaciones y atributos que se emplean para definir objetos y eventos son a su vez conceptos de la ontología. Esto, como indica Pérez, enriquece la estructura conceptual, ya que permite crear estructuras conceptuales ricas:

[...] enriquece enormemente la estructura conceptual, ya que, por medio de la identificación de relaciones y atributos, deja de ser una estructura jerárquica plana, como es, por ejemplo, el SEMANTIC NETWORK del UMLS o la clasificación de tipos de cáncer del ICD [...]. Esta estructura conceptual se ve, además, enriquecida por otras características: la obligatoriedad de definir una relación y su relación inversa (por ejemplo, la relación HAS-SYMPTOM se complementa con la inversa SYMPTOM-OF) y la posibilidad de implementar mecanismos de herencia no-monotónica (es decir, que un concepto hijo no herede determinadas propiedades de su concepto superior).

(Pérez, 2002: 5.3)

El sucesor del proyecto ONCOTERM es el proyecto PUERTOTERM, que, siguiendo las mismas bases teóricas que el anterior, se aplica al dominio de la ingeniería de costas. Este proyecto tiene como principales puntos de interés la organización conceptual subyacente en los recursos de conocimiento, la naturaleza multidimensional de las representaciones conceptuales y la extracción de conocimiento a partir de corpus multilingües (Faber *et al*, 2006). El aspecto más novedoso de este proyecto es la consideración de la información gráfica como elemento representativo de la información conceptual que puede servir para evocar un marco determinado.

Valoración

El proyecto ONCOTERM (y su sucesor PUERTOTERM) es, a nuestro juicio, una de las contribuciones más novedosas en el campo de la terminografía basada en corpus. Dentro de sus aportaciones destacamos las siguientes. En primer lugar, ONCOTERM se apoya en unos cimientos teóricos sólidos procedentes de la Lingüística y la Lexicografía. En segundo lugar, su enfoque metodológico, al combinar el trabajo sobre términos *in vivo*, *in vitro* e *in situ* garantiza la validez de los datos obtenidos. Tercero, la consideración de las macrocategorías que participan en un determinado evento permite sistematizar los tipos de relación que se establecen entre las diversas clases de conceptos. Cuarto, el empleo de una ontología como substrato conceptual de una base de datos terminológica garantiza que el producto resultante posea una orientación verdaderamente conceptual.

La aplicación del Modelo Lexemático Funcional en este proyecto nos parece interesante para el enfoque de arriba abajo. Sin embargo, no todas las disciplinas especializadas posean fuentes lexicográficas y terminológicas lo suficientemente desarrolladas como para utilizar sus definiciones como punto de partida para el análisis conceptual, y por tanto es posible que este modelo no sea aplicable en la práctica en todas las áreas temáticas.

3.3.2 El sistema informático CAOS (Madsen *et al*, 2004)

CAOS (Computer-Assisted Ontology Structuring) es un sistema informático desarrollado por el Departamento de Lingüística Computacional del

Copenhagen Business School con la colaboración de DANTERM²³. Su objetivo es desarrollar un sistema informático que permita la generación semiautomática de sistemas de conceptos, u ontologías. Dichos sistemas de conceptos generados semiautomáticamente por CAOS se integran posteriormente en la base de datos terminológicos DANTERM^{CBS}, que es una base de datos orientada al concepto. Para el desarrollo de este sistema es necesario formalizar el trabajo terminográfico.

Los principios aplicados en el sistema CAOS son los siguientes (Madsen *et al*, 2004):

1. **Concepto superior vacío:** El concepto superior en una ontología CAOS siempre es un concepto vacío denominado *top*. Esto permite introducir conceptos en la ontología que no están relacionados con ningún otro concepto.

2. **Estructura polijerárquica:** Un concepto en CAOS puede tener más de un superordinado directo²⁴.

3. **Principio de herencia:** Un concepto hereda automáticamente todas las características especificadas en sus conceptos superordinados. Así, en CAOS se distinguen dos tipos de especificaciones de características: primarias y heredadas. Las primarias se asignan directamente a un concepto y las heredadas proceden del superordinado. Las características primarias pueden modificarse, mientras que las heredadas no pueden cambiarse ni modificarse. Esto posibilita la economía en la construcción de la ontología y asegura su consistencia.

4. **Función de los atributos a los valores:** Un concepto se caracteriza por medio de un conjunto de especificaciones de las características, denominado *estructura de las características*. Una estructura de características se define formalmente como una función establecida entre un conjunto de atributos y el conjunto de valores que puede tener dicho atributo. Cada atributo debe tener una sola estructura de características, con el fin de evitar inconsistencias al estructurar la ontología.

²³ DANTERM es el centro nacional de terminología de Dinamarca, que se encarga tanto de la terminología como del desarrollo de la tecnología Lingüística y la organización conceptual. Para más información puede consultarse el sitio web <http://www.danterm.dk/forside.htm> (Fecha de consulta: junio de 2007).

²⁴ Esto coincide con lo que hemos denominado herencia múltiple en apartados anteriores.

5. Un concepto debe registrar los atributos y valores de todas las especificaciones de características de sus subordinados: Esto se hace por medio de la especificación de las dimensiones en el concepto superordinado. Una dimensión de un concepto es un atributo cuyos valores posibles permiten distinguir entre algunos de los conceptos subordinados al concepto en cuestión. Se puede representar del siguiente modo: (DIMENSIÓN: [valor1, valor2,...]). Para cada uno de los valores que aparecen en una especificación de la dimensión debe haber una especificación de la característica en uno de sus conceptos subordinados, o, a la inversa, cada vez que se añade una especificación de una característica primaria a un concepto, la información correspondiente se registra en una especificación de la dimensión en el concepto superordinado más próximo. Por ejemplo, la especificación de la dimensión CREATION en el concepto *mortgage* puede tener los valores *general words* y *legal charge* y se expresa del siguiente modo: CREATION: [general words, legal charge] (Madsen *et al*, 2004: 38).

6. Singularidad de las dimensiones: Una dimensión determinada sólo puede darse en uno de los conceptos de la ontología. Por consiguiente, las especificaciones de las características primarias con el mismo atributo deben estar presentes en los conceptos que están al mismo nivel (*sister concepts*). Siguiendo con el ejemplo anterior, en el sistema conceptual del concepto *mortgage* debe haber dos subordinados que adquieran el valor *general words* y *legal charge* para la dimensión CREATION, respectivamente.

7. Singularidad de las especificaciones de las características primarias: Una especificación de características sólo puede aparecer como primaria una vez en la ontología. Una especificación de una característica primaria sólo puede aparecer una vez en uno de los conceptos subordinados. Tanto este principio como el anterior contribuyen a la coherencia y simplicidad de la ontología.

8. Posibilidad de elegir una o más dimensiones para subdividir las ontologías: Al aplicar el criterio de subdivisión, es decir, al seleccionar una o más dimensiones para subdividir la ontología, cada concepto subordinado debe tener una única especificación de características cuyo atributo sea una de las dimensiones de subdivisión (*subdividing dimension*) del concepto superordinado. Una especificación de una característica primaria que contiene como atributo una dimensión de subdivisión se denomina *delimiting feature specification*. Por tanto,

se deben determinar siempre las características que se consideran esenciales porque son estas las que tienen valor para la definición.

9. Distinción entre los conceptos superordinados (*mother concept*) y subordinados (*daughter concepts*): Un concepto tiene que distinguirse de cada uno de sus conceptos superordinados más próximos al menos en una de las especificaciones de las características.

10. Distinción entre conceptos coordinados (*sister concepts*): Un concepto debe distinguirse de cada uno de sus conceptos hermanos al menos en una característica.

3.3.2.1 Gestión de la multidimensionalidad en el sistema CAOS

Madsen *et al* (2005) entienden la multidimensionalidad como los distintos criterios de subdivisión aplicados en los sistemas de conceptos. La noción de criterio de subdivisión se refiere al tipo de característica delimitadora, entendida como la condición necesaria y suficiente para la delimitación de un concepto en función de su posición en un sistema de conceptos.

Como hemos visto, una dimensión es un atributo que se da en una especificación de las características (no heredada) en uno de los conceptos subordinados. En CAOS, los criterios de subdivisión se expresan por medio de dimensiones formales, es decir, mediante la especificación de las dimensiones (DIMENSION: [value1, value2, ...]). Algunas dimensiones de un concepto se consideran dimensiones subdivisorias, y representan un criterio de subdivisión.

La representación gráfica de los criterios de subdivisión en CAOS 2²⁵ se realiza empleando una versión mejorada del lenguaje UML (Unified Modeling Language), que permite especificar los distintos valores de las dimensiones. Los conceptos que pertenecen a la misma dimensión subdivisora se agrupan, y dicha dimensión se muestra en los enlaces entre conceptos.

A la hora de seleccionar las dimensiones subdivisorias (especificaciones de características delimitadoras) que reflejan las características esenciales del concepto, se aplican tres restricciones formales, que son las siguientes:

²⁵ El sistema CAOS se denominó CAOS 1 entre 1998 y 2004. En 2005 se empezó a utilizar el sistema CAOS 2, que incluye una interfaz de usuario gráfica.

a) Un concepto puede contener como mucho una especificación delimitadora de las características (es decir, las dimensiones subdivisoras no pueden superponerse).

b) Todo concepto (excepto el concepto de nivel 1 *top*) debe contener al menos una especificación delimitadora de las características (es decir, las dimensiones subdivisoras deben cubrir todos los conceptos subordinados)

c) La definición de un concepto debe tener la estructura siguiente: ‘(Concept is) a SuperordinateConcept that has DelimitingCharacteristic’

Con el fin de ilustrar la gestión de la multidimensionalidad en CAOS, reproducimos en la Fig. 10 el ejemplo propuesto por Madsen *et al* (2005: 170):

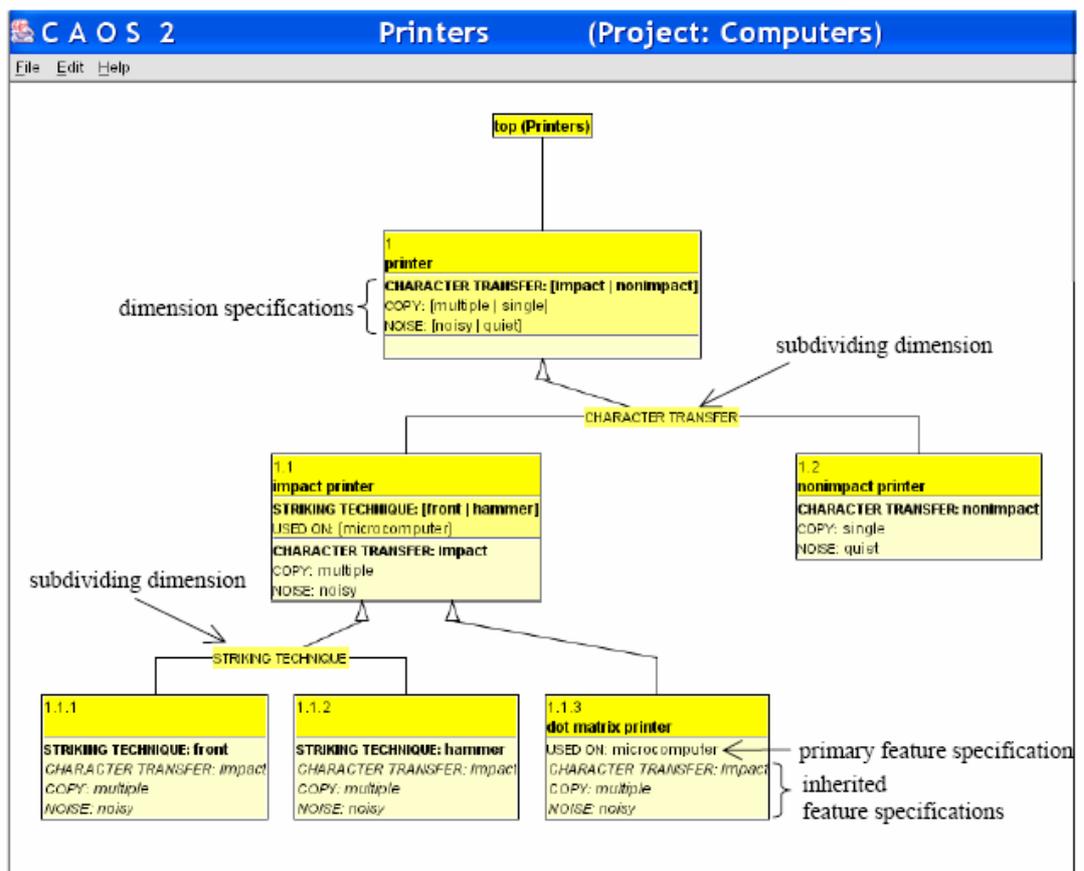


Fig. 10. Sistema de conceptos en CAOS 2 (Madsen *et al*, 2005).

En esta imagen, CHARACTER TRANSFER: [impact | nonimpact] y STRIKING TECHNIQUE: [front | hammer] representan especificaciones de las dimensiones, que son heredadas por los conceptos superordinados, mientras que la característica USED ON: microcomputer, que aparece en el concepto *dot matrix printer*, es una especificación primaria de las características.

Para Madsen *et al* (2005), esta concepción de la multidimensionalidad conlleva varias ventajas. En primer lugar, aumenta la simplicidad, la consistencia y la coherencia de la descripción. En segundo lugar, reduce el riesgo de omitir conceptos relevantes dentro de un ámbito temático. En tercer lugar, proporciona directrices claras y consistentes para la elaboración de definiciones. Por último, permite ofrecer una visión de conjunto del sistema de conceptos.

Valoración

La metodología para el desarrollo de sistemas de conceptos mediante ontologías que facilita el sistema CAOS toma como punto de partida los principios de la TGT, y mediante la distinción de las posibles dimensiones, permite convertir los sistemas de conceptos tradicionales, cuya estructura era informal, en ontologías formales (Madsen *et al*, 2002).

El aspecto más novedoso de este sistema es que permite el tratamiento semiautomático de los sistemas de conceptos gracias a las inferencias que el propio sistema permite hacer sobre las características de los conceptos.

Otra de las principales aportaciones de CAOS es la posibilidad de incluir varios criterios de subdivisión (dimensión subdivisora) en los sistemas de conceptos, lo que posibilita la gestión de la multidimensionalidad. Los criterios de subdivisión contribuyen asimismo a ofrecer un panorama más claro del ámbito temático.

3.3.3 Integración del enfoque sociocognitivo y las ontologías: Termontography (Kerremans y Temmerman, 2003)

La termontografía es un enfoque multidisciplinar en el que se combinan las teorías y métodos propuestos para el análisis terminológico multilingüe desde el enfoque sociocognitivo de la terminología (Temmerman, 2000, descrito en el apartado 2.2.3) con los métodos y directrices del análisis ontológico. Esta combinación se justifica por el gran número de coincidencias que Kerremans y Temmerman (2003) identifican entre las metodologías propuestas para la recopilación terminológica (Sager, 1990 y Cabré, 1993) y las del desarrollo de ontologías (ya se trate de ontologías basadas en texto u orientadas hacia tareas o aplicaciones). Se podría resumir como un enfoque terminológico en el cual el conocimiento multilingüe extraído a partir de un corpus textual se estructura de

manera independiente de la cultura y de la lengua, y posteriormente se va refinando en una capa en la que se consideran los aspectos culturales.

La termontografía está siendo aplicada en el marco del proyecto FF POIROT²⁶, un proyecto de investigación de la Unión Europea cuyo principal objetivo es desarrollar aplicaciones que permitan detectar el fraude en Internet. El proyecto busca integrar una ontología y una base de datos terminológica para el desarrollo de una aplicación que permita detectar e interceptar páginas web fraudulentas.

El enfoque de la termontografía es funcional, ya que el contenido y la estructura del diccionario resultante se basan en un análisis detallado de su finalidad, los requisitos de los usuarios y el alcance del dominio de especialidad objeto de estudio.

La termontografía contempla el estudio de las unidades de comprensión (*units of understanding, UoU*) en el sentido propuesto desde el enfoque sociocognitivo de la terminología propuesto por Temmerman (2000). Las UoU son categorías conceptuales independientes de la cultura y del lenguaje natural, que se organizan en una estructura de categorización (*categorisation framework*) en la que las relaciones intercategoriales están formalizadas (Kerremans, 2004). Este enfoque sirve como punto de partida para la extracción y la conexión (*mapping*) entre las unidades de comprensión y las unidades lexicalizadas extraídas a partir de un corpus textual multilingüe.

La estructura de categorización, que como hemos dicho es independiente de la cultura y de la lengua, constituye una primera capa que puede expandirse posteriormente con otra en la que se recojan las categorías específicas de cada cultura reflejadas en el corpus textual.

El método de trabajo *termontográfico* se define como un enfoque *middle-out*, puesto que combina los enfoques *top-down* y *bottom-up*. En los primeros compases de trabajo termontográfico se elabora –en colaboración con los expertos en el área temática– una estructura inicial de las categorías y de las relaciones intercategoriales que las vinculan. Esta estructura inicial se va refinando a medida que se extrae el conocimiento proporcionado por el material textual.

²⁶ Para más información puede consultarse la página del proyecto <http://www.ffpoirot.org/> (Fecha de consulta: junio de 2007).

En Temmerman y Kerremans (2003) se nos presenta el flujo de trabajo de la termontografía, que se organiza en seis fases que se describen a continuación:

- a) **Fase de análisis:** En esta primera fase se especifican la finalidad del proyecto, el alcance del dominio de especialidad y los requisitos de los futuros usuarios. Esta información se recogerá en el *Terminology Specification Report* (Informe de Especificación de la Terminología, TSR). En esta etapa se elabora también la estructura de categorización.
- b) **Fase de compilación de la información:** A la vista de los resultados de la fase de análisis, se lleva a cabo la búsqueda de material textual relevante.
- c) **Fase de búsqueda:** A partir del corpus recopilado en la fase anterior, el *termontógrafo* extrae los términos y categorías y los enlaza con la categoría correspondiente en la estructura de categorías. También se identifican patrones verbales que indiquen relaciones entre ellas. Al final de esta fase se obtiene una primera versión de la base de datos termontológica.
- d) **Fase de refinamiento:** La finalidad de esta fase es seguir completando la base de datos termontológica a partir de la información contextual extraída del corpus (concordancias).
- e) **Fase de verificación:** En esta fase se valora la consistencia y corrección de la base de datos termontológica, comprobando, por ejemplo, si todos los términos tienen asignada una categoría.
- f) **Fase de validación:** El *termontógrafo* comprueba que el contenido de la base de datos cumple los requisitos especificados en el TSR.

El recurso resultante de estas fases es un depósito de conocimiento terminológico estructurado de manera ontológica.

Valoración

La termontografía aporta una metodología para elaborar una base de datos en la que se combinan los postulados del enfoque sociocognitivo de la terminología con los principios necesarios para el desarrollo de ontologías.

Dado que nuestro objetivo empírico es la creación de una ontología que sirva como base para la organización de una terminología, sus aportaciones sobre

las fases del trabajo en el enfoque termontográfico nos pueden servir de ayuda para acometer la elaboración de nuestra base de datos conceptual empleando un editor de ontologías.

Sin embargo, el hecho de que esta metodología para el desarrollo de ontologías se apoye sobre un enfoque concreto de la terminología (el enfoque sociocognitivo), puede limitar en cierta medida su utilidad, ya que en este paradigma se consideran fundamentales aspectos como la prototipicidad de los términos.

3.3.4 Proyecto GENOMA-KB

El proyecto GENOMA-KB ha sido dirigido por la Dra. Teresa Cabré y desarrollado por el Grupo IULATERM del Institut Universitari de Lingüística Aplicada de la Universitat Pompeu Fabra. Su objetivo fundamental es la integración de las bases de datos terminológicas y las ontologías existentes en el dominio de la Genómica. De este modo se pretende proporcionar un acceso más amplio y más fácil a la información especializada contenida en la base de conocimiento.

La base de conocimiento GENOMA-KB²⁷ se estructura en cuatro módulos que convergen pero se mantienen independientes entre sí: la **ontología**, la **base terminológica**, el **corpus** y la **base de datos factual**. En la Fig. 11 se presenta la estructura general de la base de conocimiento GENOMA-KB (Cabré *et al*, 2004):

²⁷ Esta base de conocimiento se encuentra disponible para su acceso público mediante contraseña en la dirección <http://genoma.iula.upf.edu:8080/genoma/index.jsp> (Consulta: junio de 2007).

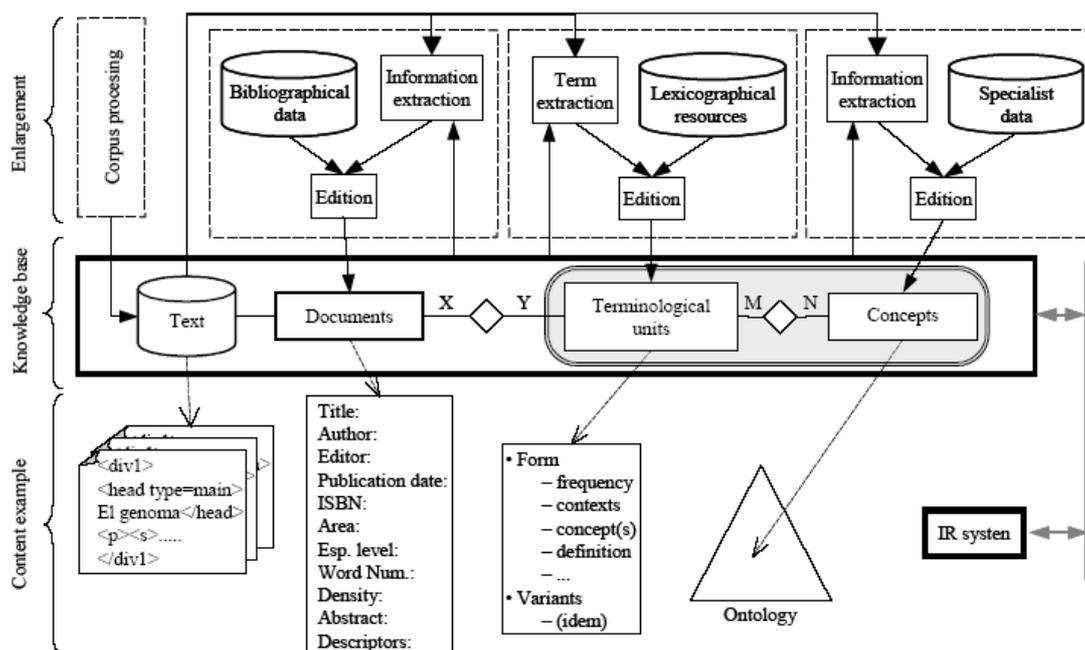


Fig. 11. Estructura general de la base de conocimiento GENOMA-KB (Cabré *et al*, 2004).

Para la edición conjunta de la ontología y la base terminológica se empleó el programa ONTOTERM®, descrito en el apartado 3.3.1. Como veíamos, este programa obliga a la constitución de una estructura conceptual en forma de ontología previa al desarrollo de la base terminológica. La ontología permite describir los conceptos de un modo formalizado, indicando en cada caso el lugar que ocupan en la jerarquía de conceptos, sus propiedades y las relaciones conceptuales que establecen. Recordemos que ONTOTERM® emplea sistemas de herencia, con lo que la construcción de la ontología se lleva a cabo de un modo más económico. Las relaciones conceptuales que se han tenido en cuenta en el proyecto GENOMA-KB son las propuestas por Feliu (2004), presentadas en el apartado 2.2.2.1: semejanza, hiponimia, secuencialidad (en el espacio y en el tiempo), causalidad, instrumentalidad, meronimia y relaciones asociativas.

La base terminológica está muy ligada a la ontología, tanto en el enfoque teórico como en su edición por medio de ONTOTERM®. En la base terminológica se ofrece la siguiente información para cada unidad terminológica: el término en catalán, castellano e inglés; la categoría gramatical, el género, el número, uno o varios contextos de uso y sus fuentes, la forma lematizada y algunos datos administrativos. En ocasiones también se añade una definición del término y su fuente, así como notas de uso cuando se consideren necesarias.

El corpus en el que se basa el trabajo terminográfico se ha almacenado y los documentos se han clasificado en función de la estructura conceptual del dominio validada por los expertos en Genómica. Todos los documentos se almacenan en una base de datos documental a la que también se puede acceder de forma independiente.

Por último, el módulo factográfico complementa la información de los corpus terminológico y textual. En él se almacenan referencias bibliográficas completas de todas las fuentes de información empleadas en los otros módulos, con la finalidad de completar la información.

Valoración

El proyecto GENOMA-KB aporta una nueva concepción de las bases de conocimiento, en la que se integran la información terminológica con la información conceptual, la bibliográfica y la información procedente del corpus textual. Con ello se crea un recurso terminográfico muy versátil y flexible, que puede cubrir las expectativas de un gran número de usuarios potenciales (traductores, redactores o especialistas, entre otros).

Estimamos que la concepción de las bases de conocimiento como un sistema modular con información conceptual, terminológica, textual y factual nos parece que supone un gran avance hacia recursos terminológicos dinámicos y flexibles, capaces de dar respuesta a la gran variedad de necesidades de sus usuarios potenciales.

Sin embargo, el manejo de la base de conocimiento y el modo de presentación de los datos todavía requieren algunas mejoras, sobre todo en cuanto a la presentación de los datos en la interfaz de la herramienta. No obstante, es probable que con el desarrollo de la plataforma para la construcción de bases terminológicas y ontologías TERMINTEGRAL (Cabré, 2004) estas deficiencias queden subsanadas.

3.3.5 Ventajas e inconvenientes del empleo de ontologías en el ámbito especializado

En este apartado se presentan las ventajas e inconvenientes del empleo de terminologías en el desarrollo de las terminologías identificadas por Lorente (2004).

El uso de ontologías en los ámbitos especializados permite mejorar la precisión y la relevancia en los sistemas de recuperación de la información (RI). Las principales ventajas de su empleo son la elevada granularidad de la descripción, la estabilidad de las ontologías con larga tradición académica y profesional (como UMLS en el ámbito de la medicina) y el hecho de que faciliten una comunicación controlada en los ámbitos especializados.

La **granularidad** se define como el grado de precisión con el que se describen las clases de conceptos y las relaciones conceptuales, resultado de la conceptualización y la categorización realizada en un ámbito de especialidad. Para Lorente (2004), las ontologías superan a otras herramientas, como los tesauros, que se limitan a hacer explícitas las relaciones jerárquicas y asociativas entre categorías determinadas *a priori*. Las ontologías permiten una visión más amplia de las relaciones conceptuales.

Una segunda ventaja de las ontologías es su **estabilidad**, al menos en los ámbitos especializados de larga tradición académica y profesional, como la medicina o la biología. Los avances y la evolución en estos ámbitos no invalidan la conceptualización anterior, sino que las novedades pueden incorporarse a la ontología sin invalidar completamente la organización conceptual previa. La estabilidad se debe al alto nivel de consenso internacional que muestran estas disciplinas.

Por último, las fuentes empleadas para la construcción de ontologías de ámbitos especializados suelen ser buenos ejemplos de **comunicación controlada**, es decir, los contenidos y su expresión suelen adecuarse al registro discursivo formal.

Por estos motivos, Lorente (2004) defiende la utilidad de las ontologías como recurso fiable para la recuperación de la información.

A pesar de estas ventajas, Lorente identifica también algunos inconvenientes del empleo de ontologías en los ámbitos especializados. Entre ellos destacan lo que ella denomina la “paradoja lingüística” y el dinamismo de los conceptos especializados.

La **paradoja lingüística** tiene que ver con la concepción lingüística de los lenguajes de especialidad en la que se enmarca el enfoque comunicativo de la terminología, al que Lorente se adscribe. Según este enfoque, los conceptos no existen como unidades independientes de la denominación o del conjunto de las

denominaciones que los vehiculan, sino que están condicionados por aspectos sociales y culturales. Generalmente, en las ontologías suele recurrirse a denominaciones en inglés para referirse a los conceptos, utilizando esta lengua como una especie de lengua franca, como si la representación conceptual fuera neutral con respecto a las lenguas. Lorente propone que en la Lingüística Aplicada se busquen soluciones metodológicas, como puede ser el empleo de caracteres alfanuméricos para referirse a los conceptos en lugar de términos asociados a una lengua concreta, o incluso la consideración de una jerarquía léxica propia para cada una de las lenguas implicadas.

En cuanto al **dinamismo de los conceptos**, Lorente (2004) considera la contradicción que se establece entre la pretendida estabilidad de la representación formalizada en la ontología y la constante evolución de los conceptos en los ámbitos especializados. Para superar este inconveniente, Lorente propone la constante actualización de la ontología, aun siendo consciente de los elevados costes asociados a esta labor de actualización.

Valoración

Si bien Lorente (2004) centra su valoración de las ontologías en su utilidad para la recuperación de la información, objetivo del proyecto RICOTERM-2 del IULA²⁸, todas sus consideraciones se hacen extensibles a la aplicación de las ontologías como base para una ontología terminológica. Destacamos como principales ventajas de las ontologías la granularidad de la representación y la estabilidad de los conceptos representados, sin perder de vista los riesgos que supone la evolución del conocimiento para dicha estabilidad.

Su reflexión sobre la “paradoja lingüística” arroja luz sobre el planteamiento de la representación de los conceptos mediante editores de ontologías en entornos multilingües.

3.4 Ontologías lingüísticas y bases de datos léxicas

En la sección anterior hemos podido comprobar que la terminología ha comenzado a utilizar las ontologías como medio para llevar a cabo una

²⁸ El proyecto RICOTERM-2. Control terminológico y discursivo para la recuperación de información en ámbitos comunicativos especializados, mediante recursos lingüísticos específicos y un reelaborador de consultas (HUM2004-05658-C02-01), financiado por el Ministerio de Educación y Ciencia, Plan Nacional de I+D+I (2004-07).

estructuración formalizada del conocimiento. La Documentación (como apuntábamos al comienzo de este capítulo) y la Lingüística también llevan algún tiempo considerando las posibilidades del empleo de ontologías. En este apartado se describen tres proyectos que, desde la Lingüística, han contemplado el uso de las ontologías o de bases de datos léxicas como medio para estructurar el léxico general.

Como veremos en el capítulo 4, dentro de las ontologías lingüísticas, Gómez Pérez *et al* (2004: 79-85) engloban las ontologías cuya finalidad es describir constructos semánticos en lugar de modelar un dominio específico, normalmente con la finalidad de aplicarlas al procesamiento del lenguaje natural. Su principal característica consiste en su descripción de la semántica de las unidades gramaticales (palabras, grupos nominales, adjetivos, etc.) y su estructuración de acuerdo con principios ontológicos. Gómez Pérez *et al* (2004) distinguen entre bases de datos léxicas en línea (como WordNet), ontologías para la traducción automática (como Mikrokosmos) y ontologías para la generación de lenguaje natural (GUM). Incluye entre las ontologías lingüísticas las siguientes: WordNet, EuroWordNet, GUM, Mikrokosmos y Sensus.

En esta sección se describen la base de datos léxica WordNet, así como su extensión en varias lenguas europeas EuroWordNet (3.4.1) y el enfoque de la semántica ontológica propuesto en el marco del desarrollo de la ontología Mikrokosmos (3.4.2). El tercer apartado de esta sección está dedicado a la presentación del proyecto FrameNet que, si bien no tiene como objetivo la creación de una ontología en sentido estricto, plantea las posibilidades de estructuración del conocimiento sobre la lengua general basándose en la teoría de marcos de Fillmore (3.4.3).

3.4.1 WordNet

WordNet es una gran base de datos léxica en lengua inglesa, desarrollada en la Universidad de Princeton y fundamentada en las teorías psicolingüísticas sobre la memoria léxica humana (Fellbaum, 1998). El hecho de que sea una ontología basada en el léxico implica que las palabras son la base de la organización conceptual. Su disponibilidad en varias lenguas (a través de su extensión EuroWordNet) y su amplia cobertura hacen que sea un recurso utilizado en múltiples aplicaciones de procesamiento del lenguaje natural.

En origen, WordNet se concibió como un banco de pruebas para las teorías cognitivas de organización del léxico, pero posteriormente se ha utilizado con éxito en distintas aplicaciones relacionadas con la recuperación de la información, la desambiguación de significados o la ingeniería del conocimiento²⁹.

Las unidades básicas que componen WordNet son los *synsets*, que se definen como conjuntos de sinónimos formados por todas las palabras que representan un mismo concepto léxico subyacente. Entre los *synsets* se establecen una serie de relaciones paradigmáticas, como la sinonimia, la antonimia, la hiperonimia e hiponimia (*Subclass-Of* y *Superclass-Of*), o la meronimia y holonimia (*Part-Of* y *Has-a*). WordNet distingue claramente entre los niveles conceptual y léxico, y por tanto entre relaciones semántico-conceptuales (las que se dan entre los *synsets*) y relaciones léxicas (las que se establecen entre palabras).

Los *synsets* están claramente divididos según su categoría gramatical y pueden pertenecer a las categorías de nombres, verbos, adjetivos y adverbios. Los nombres se organizan en jerarquías temáticas, mientras que los verbos se organizan según varias relaciones de implicación (*backward implication*, *temporal inclusion*). Entre los adjetivos y los adverbios se establecen relaciones de sinonimia y antonimia.

La relación básica en WordNet es la sinonimia (Miller, 1998), que es una relación semántica entre las formas de las palabras. Sin embargo, la relación más importante en términos de organización de los sustantivos es la relación de hiponimia, que permite desarrollar jerarquías léxicas o diagramas arbóreos que conforman un sistema de herencia en el que los subordinados heredan todas las propiedades de sus superordinados.

WordNet organiza los nombres en jerarquías independientes desarrolladas a partir de un *unique beginner*, un nombre sin superordinado. Estas jerarquías múltiples permiten la organización en campos semánticos relativamente diferenciados, en cada uno de los cuales los subordinados heredan las características de su *unique beginner*. Miller (1998: 29-30) propone una lista de veinticinco *unique beginners*, organizados en once niveles, algunos de ellos con subniveles. En la Fig. 12 se enumeran los once *unique beginners* propuestos por

²⁹ En la tercera parte de Fellbaum (1998) se ofrece información sobre estas y otras aplicaciones de WordNet.

Miller (marcados tipográficamente con negrita y cursiva), con indicación de los subniveles iniciales:

entity:
organism:
 { animal, fauna }
 { person, human being }
 { plant, flora }
object:
 { artifact }
 { natural object }:
 { body }
 { substance }:
 { food }
abstraction:
 { attribute }
 { quantity, amount }
 { relation }:
 { communication }
 { time }
psychological feature:
 { cognition, knowledge }
 { feeling, emotion }
 { motivation, motive }
natural phenomenon:
 { process }
activity
event
group
location
possession
shape
state

Fig. 12. Los once *unique beginners* propuestos en WordNet (Miller: 1998, 30).

La organización de los *synsets* WordNet se señala por medio de punteros que representan las relaciones semánticas de hiponimia e hiperonimia. La información sobre los rasgos semánticos o características en WordNet se encuentra en las definiciones (o glosas) explicativas que acompañan a los *synsets*, es decir, la información sobre los rasgos semánticos no define la hiponimia ni contribuye a la jerarquía de forma explícita.

Además de las relaciones básicas de sinonimia –que agrupa los lexemas en *synsets*– e hiponimia –que permite estructurar los *synsets*– de manera jerárquica, en WordNet se contemplan otras relaciones paradigmáticas, como la

meronimia³⁰, los atributos y los modificadores entre sustantivos, y la implicación, la troponimia³¹ y las relaciones causales entre verbos. Los adjetivos y adverbios, por su parte, se organizan en función de las relaciones de sinonimia y antonimia.

A modo de ejemplo, a continuación se reproducen los resultados de la búsqueda del lexema *tile* (baldosa) en WordNet 2.1, con indicación de su significado, sus relaciones de hiponimia (hiperónimos e hipónimos) y sus relaciones meronímicas (merónimos y holónimos).

Significado

The noun tile has 2 senses (first 1 from tagged texts)

1. (1) tile -- (a flat thin rectangular slab (as of fired clay or rubber or linoleum) used to cover surfaces)
2. tile, roofing tile -- (a thin flat slab of fired clay used for roofing)

The verb tile has 1 sense (no senses from tagged texts)

1. tile -- (cover with tiles; "tile the wall and the floor of the bathroom")

Hiperónimos

Sense 1

tile -- (a flat thin rectangular slab (as of fired clay or rubber or linoleum) used to cover surfaces)

=> slab -- (block consisting of a thick piece of something)

=> block -- (a solid piece of something (usually having flat rectangular sides); "the pyramids were built with large stone blocks")

=> artifact, artefact -- (a man-made object taken as a whole)

=> whole, unit -- (an assemblage of parts that is regarded as a single entity; "how big is that part compared to the whole?"; "the team is a unit")

=> object, physical object -- (a tangible and visible entity; an entity that can cast a shadow; "it was full of rackets, balls and other objects")

=> physical entity -- (an entity that has physical existence)

=> entity -- (that which is perceived or known or inferred to have its own distinct existence (living or nonliving))

Sense 2

tile, roofing tile -- (a thin flat slab of fired clay used for roofing)

=> roofing material -- (building material used in constructing roofs)

=> building material -- (material used for constructing buildings)

=> artifact, artefact -- (a man-made object taken as a whole)

³⁰ En WordNet se consideran tres tipos de meronimia: componente funcional-objeto, miembro-colección y material-objeto, coincidentes con algunas de las descritas en el apartado 2.4.2.

³¹ La troponimia hace referencia a la relación de modo entre verbos (Fellbaum, 2002). No se trata de una relación homogénea, sino que implica tres subrelaciones: modo, función y resultado.

=> whole, unit -- (an assemblage of parts that is regarded as a single entity; "how big is that part compared to the whole?"; "the team is a unit")
=> object, physical object -- (a tangible and visible entity; an entity that can cast a shadow; "it was full of rackets, balls and other objects")
=> physical entity -- (an entity that has physical existence)
=> entity -- (that which is perceived or known or inferred to have its own distinct existence (living or nonliving))

Hipónimos

Sense 1

tile -- (a flat thin rectangular slab (as of fired clay or rubber or linoleum) used to cover surfaces)

=> tessera -- (a small square tile of stone or glass used in making mosaics)

=> tessella -- (a small tessera)

Sense 2

tile, roofing tile -- (a thin flat slab of fired clay used for roofing)

=> hip tile -- (a tile shaped so as to cover the hip of a hip roof)

=> pantile -- (a roofing tile with a S-shape; laid so that curves overlap)

=> ridge tile -- (a decorative tile that is bent in cross section; used to cover the ridge of a roof)

Holónimos

Sense 1

tile -- (a flat thin rectangular slab (as of fired clay or rubber or linoleum) used to cover surfaces)

PART OF: tile roof -- (a roof made of fired clay tiles)

Merónimos

Sense 2

tile, roofing tile -- (a thin flat slab of fired clay used for roofing)

HAS SUBSTANCE: clay -- (a very fine-grained soil that is plastic when moist but hard when fired)

Valoración

WordNet es una gran base de datos léxica en la que los lexemas y sus relaciones se estructuran de manera formalizada. Su empleo en múltiples aplicaciones para el procesamiento del lenguaje natural da cuenta de la relevancia que tiene la formalización de las estructuras léxicas en las aplicaciones informáticas.

Sin embargo, creemos que WordNet aún podría tener un mayor valor si la estructuración formal se extendiera a los rasgos semánticos de las unidades léxicas. En WordNet, esta información no se encuentra formalizada, sino que se halla dispersa en las definiciones, que, por otra parte, no se incluyen para todos los *synsets*.

Una de las deficiencias que se suele achacar a WordNet es su granularidad asimétrica, es decir, el hecho de que las distintas áreas de la ontología (dominios)

no se encuentren desarrolladas en la misma medida (Feliu *et al*, 2002: 14). Además, al tratarse de un recurso de conocimiento general, en el mismo *synset* aparecen términos y palabras con distinto grado de especialidad.

3.4.1.1.1 EuroWordNet (Vossen, 1999)

EuroWordNet (EWN) es una base de datos léxica multilingüe creada en Europa siguiendo las directrices propuestas para la creación de WordNet, disponible sólo en inglés. EuroWordNet contiene las bases léxicas individuales en las siguientes lenguas: alemán, checo, español, estonio, francés, inglés, italiano y neerlandés.

Al igual que WordNet, cada una de las bases léxicas contiene información sobre sustantivos, verbos, adjetivos y adverbios organizados en *synsets* y relacionados entre sí por medio de relaciones semántico-conceptuales (hiperonimia, hiponimia, meronimia, etc.). Sin embargo, en EWN la presencia de adjetivos y adverbios está limitada, y sólo se incluyen en la medida en que están relacionados con sustantivos y verbos. El vocabulario abarca las palabras generales y básicas de cada lengua. Se ha incluido además un subapartado de terminología informática, con el fin de evaluar la posibilidad de integrar la terminología en un lexicón con fines generales.

En EuroWordNet, las relaciones semánticas se consideran relaciones intralingüísticas, es decir, se producen entre *synsets* dentro de una misma lengua. Además de las relaciones intralingüísticas, cada *synset* en una lengua se asocia con el de significado más próximo de WordNet, de modo que los *synsets* contenidos en WordNet funcionan como índice interlingüístico y convierten la base de datos en multilingüe. La base léxica multilingüe puede emplearse para la recuperación de la información en varias lenguas, para la transferencia de información entre recursos o como base para el estudio de recursos léxico-semánticos.

La descripción detallada de la compleja estructura de la base léxica excede los límites de este trabajo; no obstante, existe un aspecto diferenciador de EuroWordNet con respecto a WordNet que reclama nuestra atención, y es el tratamiento de las relaciones semántico-conceptuales.

En efecto, mientras que WordNet sólo recogía relaciones semánticas entre *synsets* pertenecientes a la misma clase de palabras (sustantivos, verbos, adjetivos

o adverbios), EWN permite relaciones entre clases de palabras diferentes. Además, EuroWordNet utiliza etiquetas para las relaciones que permiten una descripción más formal y precisa. Dichas etiquetas permiten diferenciar las implicaciones semánticas de las relaciones de forma precisa. Se distinguen cuatro tipos de etiquetas (Vossen, 1999: 13-15):

- a) **Conjunción o disyunción** de las relaciones múltiples del mismo tipo relacionadas con un *synset* (por ejemplo, todas las piezas juntas que constituyen un coche);
- b) **“Factividad”(factivity) o no** de las relaciones causales. Por ejemplo, *matar implica morir* (causalidad factiva), mientras que *buscar puede implicar encontrar* (causalidad no factiva);
- c) **Reversividad de las relaciones:** En EWN es necesario que cada relación tenga su relación inversa; la etiqueta *reversed* se emplea en las relaciones que son fruto de la inversión de la relación, aunque no se cumpla la implicación en los dos sentidos.
- d) **Negación de las relaciones:** Esta etiqueta permite bloquear de manera explícita algunas implicaciones. Por ejemplo, los *macacos* suelen tener cola, pero existe un macaco específico, el *Barbary ape*, que no tiene. El empleo de la etiqueta *negative* permite bloquear esa inferencia.

EuroWordNet añade algunas relaciones a las formalizadas en WordNet. En la Tabla 10 reproducimos las relaciones establecidas entre *synsets* en EWN (Vossen, 1999: 17-18). En ella se indican las clases de palabras relacionadas, las etiquetas que pueden asociarse a la relación y los tipos de datos (en este caso sólo *synsets*, instancias o variantes de los *synsets*).

Tipo de relación	Clases de palabra	Etiquetas	Tipos de datos
NEAR_SYNONYM	N<>N, V<>V		Syn <>Syn
XPOS_NEAR_SYNONYM	N<>V, N<>AdjAdv, V<>AdjAdv		Syn <>Syn
HAS_HYPERONYM	N>N, V>V	dis, con	Syn <>Syn
HAS_HYPONYM	N>N, V>V	dis	Syn <>Syn
HAS_XPOS_HYPERONYM	N>V, N>AdjAdv, V>AdjAdv, V>N, AdjAdv>N, AdjAdv>V	dis, con	Syn <>Syn
HAS_XPOS_HYPONYM	N>V, N>AdjAdv, V>AdjAdv, V>N, AdjAdv>N, AdjAdv>V	dis	Syn <>Syn
HAS_HOLONYM	N>N	dis, con, rev, neg	Syn <>Syn
HAS_HOLO_PART	N>N	dis, con, rev, neg	Syn <>Syn
HAS_HOLO_MEMBER	N>N	dis, con, rev, neg	Syn <>Syn
HAS_HOLO_PORTION	N>N	dis, con, rev, neg	Syn <>Syn
HAS_HOLO_MADEOF	N>N	dis, con, rev, neg	Syn <>Syn
HAS_HOLO_LOCATION	N>N	dis, con, rev, neg	Syn <>Syn
HAS_MERONYM	N>N	dis, con, rev, neg	Syn <>Syn
HAS_MERO_PART	N>N	dis, con, rev, neg	Syn <>Syn
HAS_MERO_MEMBER	N>N	dis, con, rev, neg	Syn <>Syn
HAS_MERO_MADEOF	N>N	dis, con, rev, neg	Syn <>Syn
HAS_MERO_LOCATION	N>N	dis, con, rev, neg	Syn <>Syn
ANTONYM	N<>N, V<>V		Syn <>Syn
NEAR_ANTONYM	N<>N, V<>V		Syn <>Syn
XPOS_NEAR_ANTONYM	N<>V, N<>AdjAdv, V<>AdjAdv		Syn <>Syn
CAUSES	V>V, N>V, N>N, V>N, V>AdjAdv, N>AdjAdv	dis, con, non-f, rev, neg	Syn <>Syn
IS_CAUSED_BY	V>V, N>V, N>N, V>N, AdjAdv>V, AdjAdv>N	dis, con, non-f, rev, neg	Syn <>Syn
HAS_SUBEVENT	V>V, N>V, N>N, V>N	dis, con, rev, neg	Syn <>Syn
IS_SUBEVENT_OF	V>V, N>V, N>N, V>N	dis, con, rev, neg	Syn <>Syn
ROLE	N>V, N>N, AdjAdv>N, AdjAdv>V	dis, con, rev, neg	Syn <>Syn
ROLE_AGENT	N>V, N>N	dis, con, rev, neg	Syn <>Syn
ROLE_INSTRUMENT	N>V, N>N	dis, con, rev,	Syn <>Syn

		neg	
ROLE_PATIENT	N>V, N>N	dis, con, rev, neg	Syn <>Syn
ROLE_LOCATION	N>V, N>N, AdjAdv>N, AdjAdv>V	dis, con, rev, neg	Syn <>Syn
ROLE_DIRECTION	N>V, N>N, AdjAdv>N, AdjAdv>V	dis, con, rev, neg	Syn <>Syn
ROLE_SOURCE_DIRECTION	N>V, N>N, AdjAdv>N, AdjAdv>V	dis, con, rev, neg	Syn <>Syn
ROLE_TARGET_DIRECTION	N>V, N>N, AdjAdv>N, AdjAdv>V	dis, con, rev, neg	Syn <>Syn
ROLE_RESULT	N>V, N>N	dis, con, rev, neg	Syn <>Syn
ROLE_MANNER	AdjAdv>N, AdjAdv>V	dis, con, rev, neg	Syn <>Syn
INVOLVED	V>N, N>N, V>AdjAdv, N>AdjAdv	dis, con, rev, neg	Syn <>Syn
INVOLVED_AGENT	V>N, N>N	dis, con, rev, neg	Syn <>Syn
INVOLVED_PATIENT	V>N, N>N	dis, con, rev, neg	Syn <>Syn
INVOLVED_INSTRUMENT	V>N, N>N	dis, con, rev, neg	Syn <>Syn
INVOLVED_LOCATION	V>N, N>N, V>AdjAdv, N>AdjAdv	dis, con, rev, neg	Syn <>Syn
INVOLVED_DIRECTION	V>N, N>N, V>AdjAdv, N>AdjAdv	dis, con, rev, neg	Syn <>Syn
INVOLVED_SOURCE_DIRECTION	V>N, N>N, V>AdjAdv, N>AdjAdv	dis, con, rev, neg	Syn <>Syn
INVOLVED_TARGET_DIRECTION	V>N, N>N, V>AdjAdv, N>AdjAdv	dis, con, rev, neg	Syn <>Syn
INVOLVED_RESULT	V>N, N>N	dis, con, rev, neg	Syn <>Syn
CO_ROLE	N>N	rev	Syn <>Syn
CO_AGENT_PATIENT	N>N	rev	Syn <>Syn
CO_AGENT_INSTRUMENT	N>N	rev	Syn <>Syn
CO_AGENT_RESULT	N>N	rev	Syn <>Syn
CO_PATIENT_AGENT	N>N	rev	Syn <>Syn
CO_PATIENT_INSTRUMENT	N>N	rev	Syn <>Syn
CO_PATIENT_RESULT	N>N	rev	Syn <>Syn
CO_INSTRUMENT_AGENT	N>N	rev	Syn <>Syn
CO_INSTRUMENT_PATIENT	N>N	rev	Syn <>Syn
CO_INSTRUMENT_RESULT	N>N	rev	Syn <>Syn
CO_RESULT_AGENT	N>N	rev	Syn <>Syn
CO_RESULT_PATIENT	N>N	rev	Syn <>Syn
CO_RESULT_INSTRUMENT	N>N	rev	Syn <>Syn
IN_MANNER	V>AdjAdv, N>AdjAdv	dis, con, rev, neg	Syn <>Syn
MANNER_OF	AdjAdv>N, AdjAdv>V	dis, con, rev, neg	Syn <>Syn
BE_IN_STATE	N>AdjAdv,	dis, con, rev,	Syn <>Syn

	V>AdjAdv	neg	
STATE_OF	AdjAdv>N, AdjAdv>V	dis, con, rev, neg	Syn <>Syn
FUZZYNYM	N<>N, V<>V		Syn <>Syn
XPOS_FUZZYNYM	N<>V, V<>AdjAdv, N<>AdjAdv		Syn <>Syn

Tabla 10. Relaciones intralingüísticas entre *synsets* en EuroWordNet (Vossen, 1999: 17).

Las relaciones pueden agruparse en los siguientes subtipos, que se definen y describen a continuación:

- a) **Sinonimia:** La sinonimia en EWN se considera una relación simétrica que relaciona dos palabras semánticamente equivalentes, es decir, que denotan el mismo rango de entidades, sin tener en cuenta las diferencias morfosintácticas, de registro, estilo o dialecto, ni las diferencias pragmáticas. Así, se permite la sinonimia entre distintas clases de palabra. Los tipos de sinonimia presentes en EWN son NEAR_SYNONYM y XPOS_NEAR_SYNONYM.
- b) **Hiponimia:** La hiponimia es la relación fundamental que permite estructurar las distintas bases léxicas. Es una relación asimétrica y transitiva. En este subtipo de relación se incluyen la hiperonimia y la hiponimia, que, al igual que en la relación de sinonimia, pueden darse entre distintas clases de palabras. Los tipos de hiponimia presentes en EWN son HAS_HYPERONYM, HAS_HYPONYM, HAS_XPOS_HYPERONYM y HAS_XPOS_HYPONYM.
- c) **Antonimia:** La antonimia es una relación simétrica que se establece entre formas de las palabras, pero no entre los *synsets*. Para que se dé, las palabras relacionadas tienen que tener un cierto grado de similitud. En EWN están presentes los siguientes tipos de relación de antonimia: ANTONYM, NEAR_ANTONYM y XPOS_NEAR_ANTONYM.
- d) **Meronomia:** EWN considera la relación meronímica como una familia de relaciones parte-todo, en la que se pueden distinguir cinco tipos de relación fundamentales (Díez Orzas, 1999). Son los siguientes:

- a. Relación entre el todo y las partes constituyentes: HAS_HOLO_PART y su inversa, HAS_MERO_PART.
- b. Relación entre una porción y el todo del que se ha obtenido: HAS_HOLO_PORTION y HAS_MERO_PORTION.
- c. Relación entre un lugar y el lugar más grande en el que se ubica: HAS_HOLO_LOCATION y HAS_MERO_LOCATION.
- d. Relación entre un conjunto y sus miembros: HAS_HOLO_MEMBER y HAS_MERO_MEMBER.
- e. Relación entre un objeto y la sustancia de la que está hecho: HAS_HOLO_MADEOF y HAS_MERO_MADEOF

Además de estos cinco tipos, se definen dos relaciones inversas generales HAS_HOLONYM y HAS_MERONYM.

- e) **ROLE e INVOLVED:** Estos dos tipos de relación se establecen entre eventos y el papel que ciertas entidades materiales e inmateriales desempeñan en dicha acción. La relación entre una entidad concreta o mental y los verbos o sustantivos que denotan eventos se denomina ROLE, y su inversa, es decir, la que se da entre el verbo y las entidades que participan en la acción, se denomina INVOLVED. Además de las relaciones generales, se distinguen varios tipos de relación ROLE / INVOLVED, que hacen referencia a los distintos participantes en un evento: agente, paciente, resultado, dirección (origen y destino) y modo. Las relaciones presentes en EWN son las siguientes:

- a. ROLE / INVOLVED_AGENT,
- b. ROLE / INVOLVED_PATIENT,
- c. ROLE / INVOLVED_RESULT,
- d. ROLE / INVOLVED_INSTRUMENT,
- e. ROLE / INVOLVED_LOCATION,
- f. ROLE / INVOLVED_DIRECTION, con los subtipos ROLE / INVOLVED_SOURCE_DIRECTION y ROLE / INVOLVED_TARGET_DIRECTION,
- g. ROLE / INVOLVED_MANNER

- f) **CO_ROLE:** Este tipo de relación permite representar pares de relaciones del tipo **ROLE / INVOLVED** entre entidades mentales o concretas en los casos en los que el evento no se menciona de forma explícita. De este modo se relacionan los coparticipantes en un evento. Además de una relación general **CO_ROLE** (que es bidireccional), se identifican los siguientes subtipos de relaciones:
- a. **CO_AGENT_PATIENT**
 - b. **CO_AGENT_INSTRUMENT**
 - c. **CO_AGENT_RESULT**
 - d. **CO_PATIENT_AGENT**
 - e. **CO_PATIENT_INSTRUMENT**
 - f. **CO_PATIENT_RESULT**
 - g. **CO_INSTRUMENT_AGENT**
 - h. **CO_INSTRUMENT_PATIENT**
 - i. **CO_INSTRUMENT_RESULT**
 - j. **CO_RESULT_AGENT**
 - k. **CO_RESULT_PATIENT**
 - l. **CO_PATIENT_INSTRUMENT**
- g) **CAUSES e IS_CAUSED_BY:** La relación causal en EWN se utiliza para relacionar dos *synsets* que se refieren a eventos (pueden ser sustantivos, verbos o adjetivos). Se distingue entre distintos tipos de causalidad dependiendo de si la causa siempre provoca el efecto (relación causal factiva) o no (relación causal no factiva).
- h) **HAS_SUBEVENT e IS_SUBEVENT_OF:** Estas dos relaciones permiten dar cuenta de la relación entre eventos (verbos o nombres que se refieren a eventos o procesos), de un modo similar a como lo hace la relación meronímica cuando nos referimos a entidades concretas en lugar de eventos.
- i) **IN_MANNER y MANNER_OF:** Este tipo de relación recoge la noción de troponimia propuesta en WordNet para los verbos que indican modos de hacer algo, por ejemplo la relación que se establece entre *comer* y *sorber* (que sería un modo de comer). En EWN se ha optado por no diferenciar entre hiponimia y troponimia, pero se ha utilizado la relación **IN_MANNER**

/MANNER_OF para expresar el modo en el que se realiza una acción de forma explícita.

- j) **BE_IN_STATE** y **STATE_OF**: Esta relación se emplea para expresar el vínculo que existe entre nombres que se refieren a un estado particular expresado por medio de un adjetivo (por ejemplo, un *pobre* (como sustantivo) se encuentra en el estado expresado mediante el adjetivo *pobre*).
- k) **Relaciones derivacionales**: Se identifican dos tipos de relación derivacional: la relación DERIVED y la relación PERTAINS. La primera es una relación puramente morfológica, que debe darse junto con otro tipo de relación semántica (sinonimia, antonimia, causa, rol). La relación PERTAIN se emplea para relacionar los adjetivos que sólo pueden aparecer asociados a sustantivos y el sustantivo del que se derivan (por ejemplo, *atómico/átomo*).
- l) **Instancia y clase**: Las relaciones BELONGS_TO_CLASS y HAS_INSTANCE se emplean para relacionar las instancias con las clases a las que pertenecen. En realidad se trata de la relación de hiponimia que se establece entre instancias (entidades individuales) y clases de entidades.
- m) **Relaciones indefinidas: “fuzzynyms”**: Esta relación permite asociar significados de palabras cuando el tipo de relación que las une no se ajusta a ninguno de los tipos anteriores. Puede darse entre palabras de la misma clase (FUZZYNYM) y palabras de clases distintas (XPOS_FUZZYNYM).

Siguiendo a Cruse (1986), en EWN se han desarrollado tests que permiten comprobar la validez de cada relación en cada una de las lenguas³². Además, se han seguido otros dos principios para la codificación de las relaciones (Vossen, 1999: 13), como el **principio de economía**, según el cual una palabra no debe definirse por medio de palabras más generales si existen palabras más específicas que permitan su definición. Así, para definir un hipónimo utilizaremos el

³² En Vossen (1999) se detallan los distintos tests que permiten escoger el tipo de relación que se establece entre palabras en inglés. En Alonge (1996) y Climent *et al* (1996) se detallan los tests para el resto de las lenguas presentes en EWN.

hiperónimo inmediatamente superior, y no otro situado más arriba en la jerarquía.

El segundo principio es el **principio de compatibilidad**, según el cual si dos significados están unidos por un tipo de relación particular, entonces no pueden estar vinculados por medio de otra relación. Este principio en ocasiones es difícil de mantener, por ejemplo cuando hablamos de nombres colectivos (como por ejemplo *cutlery* o *furniture*) en los que se pueden identificar relaciones tanto de hiponimia como de meronimia con *fork* o *table*. En ese caso, algunas relaciones se consideran prioritarias sobre otras, es decir, que la hiponimia prevalece sobre la meronimia.

Valoración

EuroWordNet amplia y perfecciona el modo en el que se representan las relaciones semánticas y conceptuales entre unidades léxicas. Estas relaciones se formalizan de manera explícita, con lo que la expresividad y las posibilidades de recuperación de la información se ven sensiblemente mejoradas con respecto a WordNet.

EWN formaliza de un modo coherente con las teorías procedentes de la Semántica Léxica, que constituyen una base sólida para el modelo de formalización.

3.4.2 La ontología Mikrokosmos y el paradigma de la semántica ontológica

La semántica ontológica (*ontological semantics*) es un enfoque teórico propuesto por Nirenburg y Raskin (2004) que pretende afrontar de forma exhaustiva y sistemática el tratamiento por ordenador del significado de los textos. Se trata de una propuesta que sistematiza las ideas sobre la descripción semántica y la representación y manipulación del significado por medios informáticos (Nirenburg y Raskin, 2004: xii). La semántica ontológica se presenta como una teoría del significado en lenguaje natural y un enfoque del procesamiento del lenguaje natural que utiliza la ontología, entendida como un modelo del mundo construido, como recurso fundamental para extraer y representar el significado de los textos en lenguaje natural y para razonar sobre el conocimiento que emana de los textos:

Ontological semantics is a theory of meaning in natural language and an approach to natural language processing (NLP) that uses a constructed world model, or ontology, as the central resource for extracting and representing the meaning of natural language texts and for reasoning about knowledge derived from texts.

(Nirenburg y Raskin, 2004: 6)

Lo que distingue al paradigma de la semántica ontológica es que propone un enfoque de urna (*glass-box*) en lugar de un enfoque de caja negra (*black-box*). Es decir, aspira a obtener hipótesis sobre qué procesos y qué conocimiento se necesitan para recrear la capacidad humana de procesar el lenguaje, con el fin de procesarlo por medios informáticos, en lugar de dar cuenta del significado en función de los fenómenos observables (Nirenburg y Raskin, 2004: 7).

La arquitectura típica de una implementación de la semántica ontológica comprende un conjunto de fuentes de conocimiento estáticas, diversos lenguajes de representación del conocimiento y una serie de módulos de procesamiento (Nirenburg y Raskin, 2004: 6):

- a) Entre las fuentes de conocimiento estáticas se encuentran una ontología, un repositorio de hechos, un lexicón y un onomasticon:
 1. La **ontología** contiene un modelo del mundo físico y de los participantes en el discurso, junto con conocimiento sobre la situación comunicativa. La ontología proporciona un metalenguaje que permite describir el significado de las unidades léxicas de una lengua, así como la especificación del significado codificado en las representaciones del significado del texto (TMR, *text meaning representations*). Contiene definiciones de los conceptos, que se corresponden con clases de objetos y eventos del mundo. Desde el punto de vista del formato, la ontología es un conjunto de marcos (*frames*), o colecciones nombradas de pares propiedad-valor.
 2. El **repositorio de hechos** (*fact repository, FR*) contiene ejemplos (*instances*) de los objetos y eventos definidos como conceptos en la ontología. Por ejemplo, si en la ontología tenemos un concepto CIUDAD, en el repositorio de hechos tendremos entradas para Londres, París o Roma.
 3. El **lexicón** conecta la ontología con el lenguaje natural. Existe un lexicón por cada una de las lenguas implicadas en la aplicación. No

contiene únicamente información semántica, si bien especifica qué concepto/s o propiedad/es de los conceptos definidos en la ontología deben instanciarse en la TMR para dar cuenta del significado de una unidad léxica de entrada.

4. El *onomasticon*, del que también hay uno por lengua implicada, contiene nombres propios, que señalan directamente a elementos del repositorio (FR). Las entradas en el *onomasticon* se indizan por el nombre, mientras que en el FR se identifican por un número único que nombra al concepto correspondiente.

Estos cuatro elementos de las fuentes estáticas de conocimiento en la semántica ontológica sirven para representar, por un lado, el conocimiento (a través de los conceptos de la ontología y el lexicón) y las instancias de los conceptos (a través del FR y del *onomasticon*). Por otro lado, la ontología y el FR nos permiten representar el conocimiento del mundo, mientras que en el lexicón y el *onomasticon* se representan los elementos del lenguaje natural (Nirenburg y Raskin, 2004: 192).

- b) Los lenguajes de representación del conocimiento permiten especificar las estructuras de sentido, las ontologías y los lexicones.
- c) El conjunto de módulos de procesamiento dinámico está formado por, al menos, un analizador semántico y un generador de textos semántico.

La semántica ontológica se acoge a la tesis de la inteligencia artificial débil (*weak AI*), que sugiere la equivalencia funcional, es decir, que los programas informáticos pueden conseguir resultados similares a los de los humanos por métodos distintos de los que estos emplean.

La teoría de la semántica ontológica se compone de una serie de **microteorías** (Nirenburg y Raskin, 2004: 27 y ss.) determinadas por las necesidades concretas que van surgiendo en el proceso de desarrollo de una aplicación. Las microteorías se ocupan de aspectos concretos dentro de todos los fenómenos posibles. Pueden referirse a la lengua en general o a un idioma en particular; a las clases de palabras, a las construcciones sintácticas, a los fenómenos semánticos y pragmáticos o al conocimiento del mundo (ontológico), así como a los procesos involucrados en el análisis y generación del lenguaje por medio del ordenador. Los métodos son por tanto heterogéneos, y dependen de la tarea que haya de llevarse a término. Algunos ejemplos de microteorías pueden

ser las preposiciones en español, la negación o la voz pasiva, entre otras. La hipótesis que mantienen Nirenburg y Raskin es que todas estas microteorías, que a veces se solapan, pueden combinarse en un solo sistema computacional que tenga en cuenta todos los fenómenos del lenguaje para los que deber servir como modelo. Las microteorías permiten combinar aportaciones de distintos ámbitos, como la Lingüística Computacional, la Inteligencia Artificial, la Filosofía del Lenguaje, la Lingüística de Corpus o las Ciencias Cognitivas.

La integración de las distintas microteorías se hace posible gracias a diversos modelos de arquitectura computacional planos (*flat architectural models*). Cada microteoría se define por medio de un conjunto de reglas que permiten rellenar un espacio concreto en una **representación del significado de un texto** (TMR, *text meaning representation*).

Las TMR (Nirenburg y Raskin, 2004: 166) son un formato que permite representar el significado del texto en la semántica ontológica. Los componentes esenciales de cada TMR son los significados de las unidades léxicas más básicas, que constituyen el punto de partida para, por medio de un proceso semántico composicional, llegar a la representación del significado de un texto, es decir, a una TMR básica del texto.

Las TMR son el resultado del proceso de análisis llevado a cabo por los distintos módulos de la arquitectura computacional del procesador semántico-ontológico (análisis ecológico, morfológico, sintáctico y semántico) y sirve como *input* para el procesamiento especializado en una aplicación concreta. Podría decirse que las TMR representan el significado del texto con el que una aplicación determinada puede trabajar. Las TMR no contienen únicamente información léxica, sino que se componen también de los resultados de los análisis morfológico y sintáctico del texto de entrada. Los elementos principales de una TMR en semántica ontológica son la ejemplificación de conceptos ontológicos y la ejemplificación de parámetros semánticos no relacionados con la ontología.

3.4.2.1 Descripción de la ontología Mikrokosmos

Como ya hemos indicado, la ontología es el módulo de conocimiento estático central dentro de los sistemas de semántica ontológica. Dado que además es el aspecto de la semántica ontológica que más nos interesa en el presente

trabajo, dedicamos este apartado a describir en detalle el metalenguaje de la ontología Mikrokosmos.

El vocabulario de dicho metalenguaje asigna etiquetas a las representaciones de eventos, objetos y sus propiedades. La sintaxis del metalenguaje facilita la expresión de significados léxicos complejos. La ontología organiza las representaciones del significado de los objetos y eventos individuales, así como de sus propiedades.

In ontological semantics, such lexical meanings [the meanings of most open-class lexical items] are represented as expressions in a special metalanguage whose vocabulary labels representations of events, objects, and their properties and whose syntax is specially designed to facilitate expressing complex lexical meanings. The representations of the meaning of individual events, objects, and their properties are organized in a structure called an ontology.

(Nirenburg y Raskin, 2004: 192-193)

La principal diferencia entre la semántica ontológica y la representación del conocimiento en IA radica en que, mientras que en IA la cobertura de los fenómenos semánticos es práctica y más general, en la semántica ontológica se pone el acento en que el sistema de representación formal sea completo desde el punto de vista teórico y en que no resulte contradictorio.

Un concepto importante dentro de la semántica ontológica es el concepto de **instanciación** (Nirenburg y Raskin, 2004: 191). Las instancias o ejemplos de los conceptos ontológicos se producen durante las fases de análisis de los textos en lenguaje natural, y se manipulan durante la síntesis de textos. Se utilizan también en los procesos de inferencia que permiten alcanzar conclusiones basadas en el análisis de una entrada determinada, sin necesidad de que toda la información esté explícita en el texto de entrada. El FR facilita el acceso a las TMR producidas por las distintas aplicaciones y permite seguir procesándolas en función de las necesidades.

En la ontología se produce un equilibrio constante entre la simplicidad y el detalle en la representación, con el fin de que el conocimiento ontológico sea fácilmente comprensible tanto para las personas que lo emplean como para los programas que la manipulan:

[...] ontological (and other static) knowledge in an ontological semantic system must be readily comprehensible to people who acquire and inspect it as well as to computer programs that are supposed to manipulate it.

(Nirenburg y Raskin, 2004: 193)

En la semántica ontológica, las primitivas³³ son las **propiedades**, que pueden ser tanto **atributos** de los conceptos como **relaciones** entre conceptos. Las propiedades se entienden como funciones que relacionan sus dominios (es decir, los conjuntos de elementos ontológicos cuya semántica ayuda a describir las propiedades) mediante conjuntos de valores.

Este conjunto de propiedades debe ser un conjunto relativamente pequeño, que al mismo tiempo dé cuenta de un gran número de elementos para la representación y manipulación del significado léxico.

En lo que se refiere a los tipos de categorías que la ontología tiene que ser capaz de representar, deben incluirse:

- a) Categorías que permitan referirse a los estados del mundo exterior.
- b) Categorías que permitan codificar el conocimiento entre los agentes, intenciones, planes, acciones y creencias.
- c) Categorías que permitan describir el metaconocimiento (es decir, el conocimiento sobre el propio conocimiento y su manipulación).
- d) Medios para codificar las categorías generadas mediante la aplicación de inferencias.

Las decisiones que se toman pueden dar lugar a distintas topologías de la jerarquía ontológica. Por ello es necesario cuestionarse cuáles son las necesidades concretas para decidir la granularidad de la descripción semántica, es decir, el grado de especificación de la información que contiene la ontología.

A continuación se describen los elementos que componen una ontología creada siguiendo los principios de la semántica ontológica, en concreto, la ontología Mikrokosmos.

La ontología Mikrokosmos es una ontología lingüística independiente del idioma en el que se trabaje. Ha sido elaborada como parte del proyecto de traducción automática Mikrokosmos. Esta ontología está construida en función de

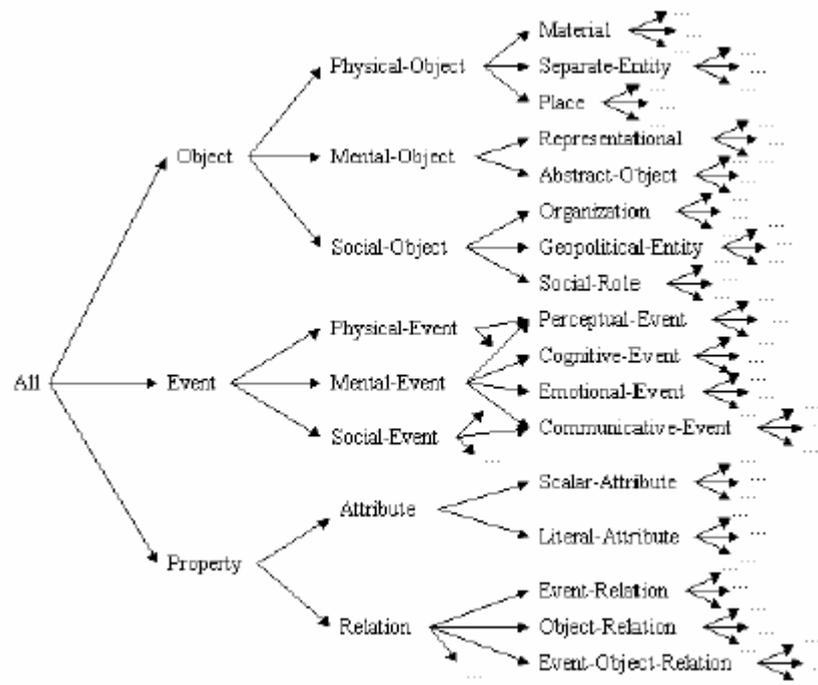
³³ Las primitivas en informática son tipos de datos que se consideran elementales en un lenguaje de programación.

consideraciones prácticas, y el principal criterio para su diseño es la distinción cuidadosa entre el conocimiento específico de una lengua (representado en un lexicón) y el conocimiento independiente de la lengua, representado en la ontología. Cada concepto en Mikrokosmos se representa mediante un marco que posee un nombre en inglés.

Una ontología se organiza como un conjunto de **conceptos**, cada uno de los cuales está formado por una colección de propiedades cuyos valores se especifican parcialmente. Los conceptos se conciben como símbolos primitivos que permiten la representación del significado por medio de atributos y relaciones con otros conceptos bien definidos. Las etiquetas empleadas para los conceptos no son palabras asociadas a ninguna lengua, sino nombres asociados a conceptos ontológicos.

Los conceptos se representan por medio de marcos (*frames*) y las propiedades de los conceptos por medio de *slots*³⁴.

En la Fig. 13 se reproduce la estructura superior de los conceptos contenidos en la ontología Mikrokosmos, que se describen a continuación:



³⁴ El concepto de *slot* volverá a manejarse al hablar del editor de ontologías *Protégé*. A falta de una traducción válida y reconocida para este término, en este trabajo se mantendrá la denominación en inglés.

Fig. 13. Estructura superior de los conceptos en la ontología Mikrokosmos (Mahesh y Nirenburg, 1995).

Los conceptos pueden aparecer en tres formatos sintácticos diferentes (ROOT, OBJECT-OR-EVENT y PROPERTY), que aparecen como un conjunto organizado en una jerarquía de herencia, lo que garantiza la economía de la ontología.

El concepto raíz es un concepto único de la ontología, que no hereda sus propiedades de ningún otro, porque es el nodo superior de la jerarquía. El concepto raíz en Mikrokosmos y en otras implementaciones de la semántica ontológica se denomina ALL.

Los **eventos** y **objetos** ontológicos son los principales instrumentos para representar los elementos esenciales del modelo del mundo. Las **propiedades**, por su parte, permiten distinguir formalmente entre los distintos objetos y eventos.

Los eventos y objetos van acompañados de nombres y definiciones. Siempre proceden de otro objeto o evento, cuya filiación se indica por medio del *slot* IS-A. Algunos de ellos poseen instancias almacenadas en el FR. Las propiedades de los conceptos poseen valores particulares que los diferencian de otros conceptos.

Las **propiedades** aparecen en la ontología de dos modos distintos, a saber, como tipos de conceptos definidos en la rama de las propiedades y como *slots* en las definiciones de objetos y eventos. Por tanto, las propiedades se representan en la ontología como conceptos y también como *slots*.

Un *slot* es el mecanismo básico utilizado para representar las relaciones entre conceptos. Los *slots* son el predicado metaontológico fundamental, y en ellos se basa la descripción axiomática de la ontología. Los distintos valores que pueden tener las propiedades se describen introduciendo facetas en el lenguaje de representación para poder tener en cuenta las distintas restricciones.

La diferencia entre atributos y relaciones viene marcada por el tipo de valores que tengan: las relaciones contienen referencias a conceptos, mientras que los atributos pueden contener elementos de conjuntos de valores específicos (valores individuales, conjuntos o rangos).

Además de las relaciones y atributos descritos hasta ahora, existe un grupo de relaciones que poseen una semántica especial. Estas relaciones describen las conexiones entre eventos y objetos o con otros eventos acerca de los que tratan los

eventos considerados “principales”. Es decir, algunos eventos contribuyen a la descripción de la semántica de proposiciones a través de la especificación de sus argumentos. Estos argumentos se denominan **roles de casos** y hacen referencia a las funciones típicas que puede desempeñar un predicado³⁵.

En la gramática de Fillmore (1982) un caso se define como un rol semántico que un argumento (generalmente un nombre) puede tener cuando se asocia con un predicado concreto (generalmente un verbo). La gramática de casos busca una descripción semántica adecuada para los verbos de una lengua.

Algunas implementaciones de la semántica ontológica (en concreto CAMBIO/CREST) han definido un conjunto de roles de casos (*case roles*) que recogen este tipo de relaciones entre los conceptos ontológicos. En este trabajo no entraremos a fondo en la descripción de estos roles, simplemente los enumeraremos y definiremos de manera sucinta:

- a) **Agente:** La entidad que causa una acción o es responsable de ella.
- b) **Tema:** La entidad manipulada por una acción.
- c) **Paciente:** La entidad afectada por una acción
- d) **Instrumento:** El objeto o evento utilizado para llevar a cabo la acción.
- e) **Fuente:** Punto de partida de varios tipos de movimiento y transferencia.
- f) **Destino:** Punto de llegada de varios tipos de movimiento y transferencia.
- g) **Ubicación:** Lugar en el que un evento tiene lugar o en el que existe un objeto.
- h) **Ruta:** Ruta a lo largo de la cual se desplaza, físicamente o de otro modo, una entidad (es decir, un tema).
- i) **Modo:** Estilo en el que se hace algo.

Las decisiones tomadas al crear la ontología Mikrokosmos pueden resumirse en 34 axiomas. La explicación de cada uno de ellos excede los límites de este trabajo, pero consideramos que constituyen una base fundamental a la hora

³⁵ Más adelante volveremos a incidir en la naturaleza de los roles semánticos al contemplar la propuesta de formalización de FrameNet (3.4.3)

de elaborar cualquier ontología. Para obtener una visión detallada de los axiomas formales para la construcción de Mikrokosmos, remitimos a Nirenburg y Raskin (2004: 224-227).

Valoración

La semántica ontológica es sin duda un marco teórico relevante para nuestro estudio, puesto que estructura y organiza los principios y axiomas que rigen la creación de ontologías orientadas a aplicaciones lingüísticas, como es nuestro caso.

En el trabajo empírico utilizaremos los niveles superiores de la ontología Mikrokosmos como punto de partida para la estructuración de nuestra propia ontología. Esta ontología, por su concepción destinada a crear ontologías útiles para el procesamiento del lenguaje natural, posee una estructura superior que tiene en cuenta los principios de la semántica, hecho éste que la hace adecuada para la organización de los conceptos de un ámbito especializado con vistas a la representación de la estructura semántica de los conceptos especializados.

Asimismo, la organización de la ontología mediante clases y *slots* es totalmente compatible con el tipo de herramienta que utilizamos en nuestro trabajo para el desarrollo de la ontología, el editor de ontologías *Protégé*.

3.4.3 La base de datos léxica FrameNet

El proyecto FrameNet de la Universidad de Berkeley tiene por finalidad la creación de una base de datos léxica en lengua inglesa, apoyándose en la semántica de marcos (*frame semantics*) y en el análisis de corpus. En principio, FrameNet se desarrolló únicamente para la lengua inglesa, pero en la actualidad se están llevando a cabo proyectos en otras lenguas, entre ellas el español, el japonés y el alemán. Por otro lado, también se estudia la posibilidad de diseñar y crear una base léxica multilingüe utilizando los marcos y los elementos que los componen como estructura común a varias lenguas.

El objetivo del proyecto FrameNet es documentar todas las *valencias* (posibilidades combinatorias sintácticas y semánticas) de cada palabra y cada uno de sus sentidos. Al hablar de valencias se hace referencia tanto a valencia semántica (roles semánticos) como a valencia sintáctica (funciones gramaticales). Este proyecto está basado en el empleo de corpus textuales y, en la medida de lo

posible, utiliza descripciones gramaticales pertenecientes a teorías neutras, puesto que se espera que los resultados del proyecto puedan ser útiles para investigadores con diferentes puntos de partida teóricos.

La información en FrameNet se presenta por medio de **unidades léxicas** (LU, *lexical units*), que consisten en la unión de una palabra y uno de sus significados. Cada sentido de una palabra polisémica corresponde a un **marco semántico** (*semantic frame*). Los marcos semánticos se definen (Fillmore *et al*, 2003; Ruppenhofer *et al*, 2005) como representaciones esquemáticas de las estructuras conceptuales y patrones de las creencias, prácticas, instituciones, imágenes, etc. que constituyen la base para la interacción significativa en una comunidad de habla. Un marco es una estructura que describe un tipo de situación, objeto o evento, así como los participantes y la utilería (*props*) implicados. Los marcos son, en definitiva, una representación esquemática de una situación en la que están implicados varios participantes.

Los marcos están formados por unidades básicas, los elementos de marco o **argumentos semánticos** (*frame elements, FE*), que son roles semánticos definidos dentro de un marco determinado. Normalmente, las unidades léxicas que evocan un marco son verbos y los elementos del marco son sus dependientes sintácticos, aunque también pueden ser sustantivos o adjetivos.

Cada unidad léxica está asociada a un marco, y por tanto al resto de las palabras que dicho marco evoca. En este sentido, FrameNet se asemeja a un tesoro, en el que se agrupan las palabras semánticamente próximas. Los marcos permiten de este modo caracterizar las relaciones semánticas existentes entre las palabras.

Además de los marcos y elementos de los marcos, FrameNet recurre también a los denominados **tipos semánticos** (*semantic types*), que constituyen un mecanismo para capturar hechos semánticos sobre marcos individuales, elementos de los marcos y unidades léxicas que no encajan en la jerarquía de marcos desarrollada en FrameNet. Los tipos semánticos permiten hacer generalizaciones semánticas que no se pueden realizar por medio de los marcos y las relaciones establecidas entre ellos. Para reflejar esta información se ha creado una tabla de tipos semánticos que pueden asociarse a unidades léxicas, marcos o elementos de los marcos. De este modo, algunos FE pueden asociarse con el tipo semántico 'human', cuando el elemento del marco tiene que ser humano. Otro ejemplo sería

el marco *Calendric_units*, en el que se incluyen los nombres de los doce meses del año, pero se utiliza el tipo semántico ‘month name’ para no tener que anotar todos los meses.

En FrameNet se estudian además las construcciones sintácticas en las que aparecen los predicados, tratando de identificar cómo las características semánticas que se definen en un marco adquieren una forma sintáctica en una o más construcciones sintácticas diferentes.

FrameNet incluye una red de relaciones entre los marcos (Baker *et al*, 2003). Las más importantes son las relaciones de herencia (IS-A), las relaciones de uso (*using*) entre marcos y las relaciones con los marcos subordinados (*subframes*), que se explican a continuación:

- a) **Herencia entre marcos:** Las relaciones de herencia suponen que un marco “hijo” es un subtipo del marco “padre”, y que cada elemento del “padre” está ligado a un elemento en el marco “hijo” (por ejemplo, el *Revenge Frame* hereda los elementos del *Rewards_and_punishment Frame*). La herencia entre marcos se define como completa y monotónica, permitiendo la posibilidad de herencia múltiple. Que la herencia sea completa significa que si un marco **B** hereda de un marco **A**, en el marco **B** tiene que haber un FE que se corresponda con un FE de **A**, si bien el marco **B** puede tener elementos que no estén presentes en **A**. El hecho de que sea monotónica supone que si un marco padre o un FE posee un tipo semántico, el tipo semántico del marco o del FE hijo debe ser el mismo. La posibilidad de que exista herencia múltiple implica que un marco hijo (y sus FE) pueden tener más de un padre.
- b) **Relación Subframe:** Muchos marcos expresan conceptos que pueden subdividirse en partes bien definidas, como por ejemplo los eventos complejos. En los marcos subordinados (*subframes*), el marco “hijo” es un evento subordinado de un evento complejo representado por el “padre”. Por ejemplo, el *Criminal_process frame* tiene submarcos como *Arrest*, *Arraignment*, *Trial* y *Sentencing*. Dichos submarcos son a su vez marcos, que comparten muchos de sus FE.

- c) **Relación Using:** La relación *Using* es similar a la herencia, pero su definición es menos estricta. En este tipo de relación, no todos los elementos del marco “padre” tienen que estar presentes en el “hijo”.
- d) **Relación See also:** Esta relación posibilita la remisión entre marcos, de forma similar a como se hace en los diccionarios tradicionales. Esta relación es un puntero establecido desde un marco a otro en el que se explican en detalle las diferencias existentes entre diversos marcos. Por ejemplo, la palabra *load* aparece en los marcos *Filling* y *Placing*. La distinción entre ambos se explica en la definición de *Filling*, por lo que desde *Placing* se remite a este marco.

La información en FrameNet se estructura en dos bases de datos distintas, que son la **base de datos léxica** (*Lexical Database*) y la **base de datos de anotación** (*Annotation Database*). La primera contiene la información sobre los marcos y los elementos que los componen, así como los lemas, lexemas, formas de palabras y partes de la oración, es decir, la información necesaria para caracterizar las unidades léxicas. Las relaciones entre marcos y elementos de los marcos también están contenidas en esta base de datos.

Por su parte, la *Annotation Database* almacena oraciones anotadas junto con los subcorpus a partir de los cuales éstas se extraen. Para cada palabra anotada existen una serie de capas (*annotation layers*) que corresponden a los distintos tipos de información (*frame elements*, *phrase types* y *grammatical functions*).

En cuanto a las posibles aplicaciones de FrameNet, la primera y más evidente es su utilidad en lexicografía, sobre todo para quienes emprendan la elaboración de un diccionario desde cero. Además, se ha utilizado en aplicaciones de procesamiento del lenguaje natural para crear un analizador semántico (*semantic parser*) que produce roles de FrameNet para frases de entrada. Dentro del PLN se están investigando sus posibilidades como recurso semántico para la extracción de información y comprensión del lenguaje natural. Por último, FrameNet resulta útil para la investigación semántica de aspectos como la polisemia, las colocaciones o las discrepancias entre las estructuras sintáctica y semántica.

Valoración

Como apuntábamos al principio de esta sección, FrameNet no puede considerarse una ontología de objetos en sentido estricto, puesto que se centra en la descripción semántica de palabras, pero carece de una estructura organizada de los objetos del mundo.

Sin embargo, FrameNet es un buen recurso para observar el comportamiento de las unidades léxicas y sus relaciones, y proporciona una buena plataforma para observar las posibilidades combinatorias de las unidades léxicas. No obstante, FrameNet no permite desplegar directamente la red semántica: se considera que las unidades léxicas están relacionadas por el hecho de pertenecer al mismo marco, y por las relaciones entre marcos, pero las relaciones no se hacen explícitas ni se pueden consultar directamente.

3.5 *Recapitulación y conclusiones*

En este capítulo nos planteábamos como objetivo principal la revisión de los distintos modos de formalización de las relaciones conceptuales que se han propuesto desde la Terminología clásica, la Terminología moderna y la Lingüística.

En la sección 3.1 hemos repasado la representación de las relaciones conceptuales en la TGT por medio de sistemas de conceptos. Estos sistemas de conceptos se estructuran casi exclusivamente por medio de relaciones jerárquicas, cuya expresividad es limitada, incluso cuando se proponen formas para su representación gráfica. Como hemos visto, los sistemas conceptuales tradicionales apenas dan cuenta de las relaciones no jerárquicas, que también resultan útiles para la estructuración del conocimiento especializado.

En la sección 3.2 hemos presentado distintos modelos de formalización de las relaciones que, desde enfoques de la Terminología más próximos a la Lingüística, tratan de superar las deficiencias de los sistemas conceptuales tradicionales. Así, hemos revisado de forma cronológica las aportaciones del proyecto COGNITERM (3.2.1), las redes semántico-terminológicas (3.2.2), la propuesta de estructuración conceptual mediante grafos (3.2.3), los esquemas relacionales (3.2.4) y la aplicación de las funciones léxicas a la elaboración de diccionarios terminológicos creados a partir de corpus (3.2.5). Todas ellas tienen en cuenta relaciones que van más allá de las jerárquicas, y plantean aspectos

interesantes para la estructuración y recuperación de la información terminológica, como la gestión de la multidimensionalidad.

Sin embargo, estas nuevas aportaciones se hacen –a excepción de COGNITERM– en un nivel teórico, sin entrar a detallar las herramientas informáticas que podrían facilitar estas estructuraciones.

En la sección 3.3 se han presentado nuevas propuestas para la gestión de la terminología, que se apoyan en la creación de ontologías, en las que las relaciones conceptuales se representan de manera explícita y formalizada. Algunas de ellas sugieren metodologías bastante detalladas para llevar a cabo la estructuración de la información terminológica mediante ontologías (CAOS y la termontografía, por ejemplo). En la valoración de cada apartado se han detallado las ventajas y los inconvenientes de cada uno de estos sistemas.

Otra aportación que queremos destacar de esta sección es la presentación del sistema integrado de gestión de la terminología ONTOTERM®, empleado en los proyectos ONCOTERM (3.3.1) y GENOMA-KB (3.3.4). Este programa permite crear bases terminológicas teniendo en cuenta la estructuración conceptual subyacente, contenida en la ontología. De este modo, se supera el problema de la falta de formalización de que adolecen los sistemas tradicionales.

Por último, en la sección 3.4 hemos revisado tres propuestas desde la lingüística cognitiva (3.4.1), la lingüística orientada al procesamiento del lenguaje natural (3.4.2) y la semántica de marcos (3.4.3). Aunque todas estas bases léxicas y ontologías se centran en la estructuración de la lengua general, sus principios y métodos son también válidos para la estructuración de la terminología de un dominio de especialidad. Destacamos especialmente el enfoque teórico de la semántica ontológica, en cuyo seno se enmarca la ontología Mikrokosmos. En nuestro trabajo empírico reutilizaremos la estructura superior de esta ontología para crear nuestra base de conocimiento con el editor de ontologías.

En conclusión, las herramientas y los métodos de la ingeniería del conocimiento constituyen una buena base para la representación del conocimiento especializado, posibilitando la superación de los enfoques clásicos de la terminología. Observamos, sin embargo, que el uso de las herramientas de la ingeniería del conocimiento no está generalizado en el trabajo terminológico. De hecho, en algunos proyectos se desarrollan herramientas específicas para el

tratamiento de la terminología, sin considerar las posibilidades de adaptación y reutilización de las herramientas existentes.

Compartimos con Antia (2000) la opinión de que se hace indispensable optimizar y rediseñar las herramientas y los recursos existentes para adaptarlos a las necesidades del proceso terminográfico. Con vistas a lograr este objetivo, en el capítulo siguiente analizaremos las herramientas que la ingeniería ontológica pone a nuestro alcance. Será nuestra misión en el estudio empírico proponer modos para la optimización y adaptación de dichas herramientas al trabajo terminológico automatizado.

CAPÍTULO 4: REPRESENTACIÓN DEL CONOCIMIENTO EN LA INGENIERÍA ONTOLÓGICA

4 Representación del conocimiento en la ingeniería ontológica

El objetivo de este capítulo es presentar los principios y herramientas de la denominada ingeniería ontológica. Esta rama de la ingeniería del conocimiento, surgida en los últimos años, se encarga del estudio del conjunto de actividades implicadas en el desarrollo de ontologías, y en el ciclo de vida de las ontologías, así como de las diversas metodologías, lenguajes y herramientas empleados en la construcción de ontologías (Gómez Pérez *et al*, 2004: 5).

La razón por la que nos interesa el estudio de la ingeniería ontológica radica en nuestra convicción de que las herramientas y metodologías empleadas en esta nueva disciplina pueden ser reutilizadas en el proceso terminológico automatizado, con ciertas adaptaciones. Como veíamos en el capítulo anterior, nos interesa optimizar y rediseñar los recursos y herramientas disponibles, con el fin de acercar dos ámbitos, el de la ingeniería ontológica y el de la terminología, que tienen en común la organización del conocimiento con vistas a su recuperación y reutilización en diversas aplicaciones informáticas.

El capítulo se ha estructurado del siguiente modo. En la sección 4.1 se revisan las distintas definiciones de ontología y las técnicas de modelado y componentes básicos que integran una ontología. Esta definición será el punto de partida para presentar los principios seguidos para el diseño de ontologías, las metodologías y los métodos empleados y los lenguajes ontológicos (sección 4.2).

En la sección 4.3 se presenta la clasificación de las ontologías atendiendo a diversos criterios y se describen de manera sucinta algunas de las ontologías más relevantes. La sección 4.4 está dedicada a la descripción de las principales herramientas para el desarrollo de ontologías, haciendo especial hincapié en los editores de ontologías. Por último, en el apartado 4.5 se recogen las conclusiones de este capítulo.

4.1 Definición de ontología

La palabra *ontología* se ha tomado prestada de la Filosofía, donde se emplea para referirse a una explicación sistemática del ser. La Ontología (con mayúscula) es “una rama de la metafísica que se ocupa del estudio de la naturaleza de la existencia, de los seres y de sus propiedades transcendentales”

(Moreno Ortiz, 2002: 30). Etimológicamente, *ontología* significa “tratado del ser”. La Ontología como disciplina trata, desde hace siglos, de desentrañar los misterios de la naturaleza, la “esencia” de los conceptos.

En el ámbito de la ingeniería del conocimiento, este término se emplea para referirse a un conjunto de conceptos organizados jerárquicamente, representados en algún sistema informático que sirven como soporte a diversas aplicaciones que requieren de conocimiento específico sobre la materia que la ontología representa (Moreno Ortiz, 2002: 31).

Son muchas las definiciones de ontología que se han propuesto en los últimos años (Neches *et al*, 1991; Gruber, 1993; Guarino y Giaretta, 1995; Swartout *et al*, 1997; Studer *et al*, 1998; Gómez Pérez *et al*, 2004). A continuación ofrecemos de manera cronológica algunas de estas definiciones, subrayando en cada una de ellas los rasgos que consideramos más relevantes³⁶:

An ontology defines the basic terms and relations comprising the vocabulary of a topic area as well as the rules for combining terms and relations to define extensions to the vocabulary. (Una ontología define los términos y relaciones básicos que abarca el vocabulario de un área temática, así como las reglas que permiten combinar los términos y las relaciones para definir ampliaciones del vocabulario).

(Neches *et al*, 1991: 40)

An ontology is an explicit specification of a conceptualization. (Una ontología es una especificación explícita de una conceptualización).

(Gruber, 1993: 199)

A logical theory which gives an explicit, partial account of a conceptualization. (Una teoría lógica que explica una conceptualización de forma explícita y parcial).

(Guarino y Giaretta, 1995)

An ontology is a hierarchically structured set of terms for describing a domain that can be used as a skeletal foundation for a knowledge base. (Una ontología es un conjunto de términos estructurado de forma jerárquica que describen un dominio, que pueden utilizarse como esqueleto para una base de conocimiento).

(Swartout *et al*, 1997: 138)

³⁶ Ofrecemos entre paréntesis nuestra propia traducción de las definiciones.

An ontology is a formal, explicit specification of a shared conceptualization. Conceptualization refers to an abstract model of some phenomenon in the world by having identified the relevant concepts of that phenomenon. Explicit means that the type of concepts used, and the constraints on their use are explicitly defined. Formal refers to the fact that the ontology should be machine-readable. Shared reflects the notion that an ontology captures consensual knowledge, that is, it is not private of some individual, but accepted by a group. (Una ontología es una especificación formal y explícita de una conceptualización compartida. La conceptualización es un modelo abstracto de un fenómeno del mundo a través de la identificación de los conceptos relevantes de dicho fenómeno. Explícito quiere decir que el tipo de conceptos empleados y las restricciones de su uso se definen de forma explícita. Formal se refiere al hecho de que una ontología debe ser comprensible por un ordenador. Compartido refleja el hecho de que una ontología refleja el conocimiento consensuado, es decir, no es el propio de un individuo, sino que es aceptado por un grupo).

(Studer *et al*, 1998)

Ontologies aim to capture consensual knowledge in a generic way, and that they may be reused and shared across software applications and by groups of people. They are usually built cooperatively by different groups of people in different locations. (Las ontologías pretenden capturar el conocimiento consensuado de forma genérica, pueden ser reutilizadas y compartidas por diferentes aplicaciones de software o grupos de personas. Normalmente se construyen de forma colaborativa entre grupos de personas diferentes ubicados en distintos lugares).

(Gómez Pérez *et al*, 2004: 8-9)

Todas ellas coinciden en señalar que una ontología trata de especificar una conceptualización explícita de los fenómenos del mundo. Una **conceptualización** se define como un modelo abstracto de algún fenómeno del mundo en el que se han identificado los conceptos relevantes de dicho fenómeno. El hecho de que sea **explícita** se refiere a que los conceptos y las restricciones que se les aplican se definen de una manera predeterminada. Casi todas ellas abundan también en la necesidad de que la conceptualización ha de ser **consensuada**, es decir, que surge del acuerdo entre un grupo de expertos, lo que permite que sea **reutilizada** y **compartida** por distintos grupos.

4.1.1 Técnicas de modelado y componentes básicos de las ontologías

Las ontologías pueden modelarse utilizando distintas técnicas que se implementan en diversos lenguajes ontológicos, como veremos en el apartado

4.2.2. Dependiendo del grado de formalidad, las ontologías puede dividirse en **ontologías ligeras** (*lightweight ontologies*) y **ontologías pesadas** (*heavyweight ontologies*). Mientras que las primeras incluyen conceptos, taxonomías de conceptos, relaciones y propiedades que permiten definirlos, las ontologías pesadas incluyen además axiomas y restricciones que aclaran el significado de los conceptos recogidos. Definimos a continuación cuáles son los principales componentes de una ontología dependiendo del paradigma de representación del conocimiento empleado.

Dos son las principales técnicas de modelado de ontologías pesadas: las que emplean marcos y lógica de primer orden y las que emplean lógica descriptiva.

4.1.1.1 Marcos y lógica de primer orden

Gruber (1993) propone el modelado de ontologías utilizando marcos y lógica de primer orden. Un **marco** es una estructura de datos a la que se le asigna un nombre, que se utiliza para representar un concepto dentro de un dominio. La lógica de primer orden permite describir conceptos, propiedades y relaciones mediante expresiones lógicas. Los distintos componentes que puede tener un marco (también denominados *primitivas*) son clases, relaciones, funciones, axiomas formales e instancias:

- a) **Clases:** Una clase es la representación de un concepto en sentido amplio. Las clases suelen organizarse en taxonomías en las que funcionan los mecanismos de herencia. Representan conceptos tanto abstractos como concretos, y pueden organizarse en metaclases, que permiten gradaciones de significado, estableciendo diferentes capas de clases en la ontología en la que se definen. Las clases poseen atributos que las describen.
- b) **Relaciones:** Una relación representa un tipo de asociación entre conceptos de un dominio. Se definen formalmente como cualquier subconjunto de un producto de **n** conjuntos. Las relaciones contenidas en las ontologías suelen ser binarias. En ellas, el primer argumento de la relación recibe el nombre de *dominio* y el segundo argumento se denomina *rango*. Las relaciones binarias también permiten conectar diferentes taxonomías. Se utilizan en ocasiones para expresar atributos de los conceptos (*slots*). Los atributos se distinguen de las relaciones por el tipo de valor de su rango,

que en el caso de las relaciones es un concepto o una instancia, mientras que en los atributos puede tener como valor una cadena de caracteres, un número o un símbolo.

- c) **Funciones:** Las funciones son un tipo especial de relaciones en las cuales el n-avo elemento de la relación es único para los n-1 elementos precedentes. Es decir, una función permite deducir el valor de un atributo de una clase en función de los elementos anteriores. Un ejemplo de función es *pays*, que deduce el precio de una habitación después de aplicar un descuento.
- d) **Axiomas formales:** Los axiomas formales permiten modelar oraciones que siempre son verdaderas. Suelen emplearse para representar el conocimiento que no puede definirse formalmente por medio de otros componentes o para verificar la consistencia de la ontología o del conocimiento almacenado en ella. Su utilidad estriba en su capacidad para inferir conocimiento nuevo a partir del conocimiento contenido en la ontología.
- e) **Instancias:** Las instancias se emplean para representar elementos o individuos que ejemplifican una clase o una relación. Las instancias remiten a objetos concretos de la realidad.

4.1.1.2 Lógica descriptiva (Description Logics, DL)

La lógica descriptiva es un formalismo utilizado para la representación del conocimiento que consiste en una ampliación de los marcos equipada con una semántica basada en la lógica de primer orden. Una teoría de lógica descriptiva se divide en dos partes: la Tbox y la Abox. La Tbox contiene conocimiento intensional (terminológico) y se construye por medio de declaraciones que describen las propiedades generales de los conceptos.

La Abox, por su parte, contiene conocimiento extensional (*assertional*), que es específico de los individuos del dominio del discurso. Es decir, la TBox contiene las definiciones de los conceptos y sus funciones, y la ABox contiene las definiciones de los individuos (*instances*).

La lógica descriptiva permite la representación del conocimiento mediante tres tipos de componentes, a saber, conceptos, roles e individuos:

- a) **Conceptos:** Los conceptos representan clases de objetos. Pueden ser *primitivos*, si se definen mediante condiciones necesarias, o *definidos*, si especifican las relaciones necesarias y suficientes que deben satisfacer los individuos pertenecientes a una clase.
- b) Los **roles** describen relaciones binarias entre conceptos, y por tanto permiten la representación de sus propiedades. También pueden ser primitivos o definidos (también llamados en este caso *derivados*).
- c) Los **individuos** representan ejemplos de conceptos y los valores de sus *roles*, es decir, sus propiedades.

Los conceptos y los roles se formalizan mediante descripciones lógicas a partir de términos preexistentes que se combinan empleando un conjunto de operadores lógicos (denominados en inglés *constructors*), entre los que cabe mencionar la conjunción, la disyunción, la negación, y las restricciones (restricción del valor, cuantificación existencial y restricción existencial, entre otras).

Los axiomas formales en la lógica descriptiva se encuentran insertados en las definiciones de roles o conceptos, y utilizan los operadores que acabamos de mencionar.

La principal ventaja de este paradigma con respecto a los marcos es su potencial para el empleo de razonadores, que permiten clasificar los conceptos automáticamente, y también comprobar la consistencia de la ontología.

Otras técnicas para la construcción de ontologías que no entraremos a considerar en detalle aquí son las técnicas de ingeniería del software (mediante el lenguaje UML, Unified Modeling Language) y el empleo de la tecnología de bases de datos, que permiten el desarrollo de ontologías ligeras³⁷.

Los formalismos empleados para modelar el conocimiento y los lenguajes que implementan estas técnicas limitan el tipo de conocimiento que puede representarse en una ontología.

³⁷ Remitimos a Gómez Pérez *et al* (2004 : 21-25) para una presentación más completa de estos formalismos.

Valoración

En esta sección hemos definido qué se entiende por ontología en la ingeniería del conocimiento y hemos analizado cuáles son sus componentes básicos.

La definición de ontología como una conceptualización explícita y consensuada de los fenómenos del mundo, que puede ser reutilizada y compartida por distintos grupos de personas y aplicaciones informáticas, sitúa las ontologías como un instrumento válido para la representación del conocimiento especializado. La inclusión de una ontología en el proceso terminográfico proporciona a los recursos creados un armazón conceptual formalizado en el que los conceptos se estructuran en torno a sus atributos y relaciones.

Las ontologías pueden desarrollarse empleando diversos formalismos. En esta sección hemos presentado los marcos y la lógica de primer orden, así como la lógica descriptiva. Estos formalismos condicionan los componentes que tiene la ontología. Sea cual sea la técnica de modelado elegida, estos componentes de las ontologías nos permiten, en su aplicación a la terminología, representar los principales elementos del análisis conceptual. Así, las clases nos permiten representar conceptos; mediante las relaciones podremos hacer explícitas las características y relaciones conceptuales, tanto jerárquicas como no jerárquicas. Las instancias y los individuos pueden representar objetos individuales o, como veremos en nuestro estudio empírico, se pueden utilizar para representar las distintas denominaciones existentes para un concepto.

4.2 Principios para el diseño de ontologías

Gruber (1993) propone cinco criterios básicos para facilitar el diseño de ontologías compartidas reutilizables. El objetivo es construir representaciones formales del conocimiento susceptibles de intercambiarse independientemente de la plataforma de hardware, el lenguaje de programación y los protocolos de red empleados, y que además puedan intercambiar conocimiento entre distintas aplicaciones (*interoperability*). Los cinco principios propuestos por Gruber, generalmente aceptados por los expertos en ingeniería ontológica, son los siguientes:

- a) **Claridad:** Una ontología debe definir de forma efectiva el significado de los conceptos mediante definiciones objetivas,

independientes del contexto social o computacional. El *formalismo* es un medio para conseguir este fin. Es preferible una definición *completa* (un predicado definido por las condiciones necesarias y suficientes) frente a una parcial (mediante condiciones necesarias). Las definiciones deben documentarse en lenguaje natural.

- b) **Coherencia:** Una ontología debe sancionar inferencias consistentes con las definiciones. La coherencia debe darse al menos en los axiomas defintorios, aunque también en los conceptos definidos de manera informal.
- c) **Posibilidad de ampliación** (*extendibility*): Una ontología debe ser diseñada de modo que anticipe los usos del vocabulario compartido. Debe permitir la definición de nuevos términos para usos especiales a partir del vocabulario existente, de tal forma que no sea necesaria la revisión de las definiciones disponibles.
- d) **Sesgo de codificación mínimo** (*minimal encoding bias*): La conceptualización debe especificarse de forma que no dependa de una codificación determinada. Se produce un sesgo en la codificación cuando las decisiones que se toman para la representación se basan únicamente en lo que resulta adecuado desde el punto de vista de la notación o de la implementación. Conviene limitar al máximo el sesgo porque, si el conocimiento ha de ser compartido, es necesario evitar que se empleen sistemas o estilos de representación distintos.
- e) **Compromiso ontológico mínimo** (*minimal ontological commitment*): El compromiso ontológico se define como la función que relaciona los términos del vocabulario de la ontología con una conceptualización. En otras palabras, el *compromiso ontológico* permite seleccionar cuál de los significados posibles del término es el que se aplica en la ontología. Una ontología debería ejercer el menor número posible de restricciones sobre la realidad que representa.

Gómez Pérez *et al* (2004: 43-44) proponen tres criterios adicionales a los propuestos por Gruber, que son: la representación de conocimiento disjunto y

exhaustivo, la reducción al mínimo de la distancia sintáctica entre conceptos hermanos y la normalización de los nombres:

- f) **La representación de conocimiento disjunto y exhaustivo:** Si las subclases de un concepto forman conjuntos disjuntos, es decir, si las instancias de las clases subordinadas a una determinada clase no pueden pertenecer a varias subclases simultáneamente, se puede definir una descomposición disjunta. Por ejemplo, los vuelos con un número determinado no pueden ser ejemplos de vuelos con otro número. La descomposición es exhaustiva si define completamente el concepto superordinado, es decir, si contempla todos los casos posibles.
- g) **Reducción de la distancia sintáctica entre conceptos hermanos:** Los conceptos hermanos se representan utilizando las mismas primitivas. De este modo se mejora la facilidad de comprensión y de reutilización de la ontología. Se recomienda emplear el mismo patrón para definir conceptos hermanos.
- h) **Normalización de los nombres:** Con el fin de facilitar la comprensión de una ontología deben emplearse las mismas convenciones para asignar los nombres a los términos relacionados. En definitiva, se trata de nombrar los conceptos siempre del mismo modo.

4.2.1 Metodologías y métodos para el desarrollo de ontologías

La elaboración de una ontología es una tarea que requiere un gran esfuerzo. Por ello es importante decidir en los primeros pasos del proceso de desarrollo las necesidades que esperan cubrirse con la ontología, y asimismo conviene definir claramente los objetivos que se pretende conseguir y los pasos que han de darse para la consecución de dichos objetivos. Los objetivos determinarán el tipo de herramientas empleadas para el desarrollo y el lenguaje ontológico empleado.

La IEEE (Institute of Electrical and Electronics Engineers, Inc.) define **metodología** como “a comprehensive, integrated series of techniques or methods creating a general systems theory of how a class of thought-intensive work ought to be performed” (IEEE, 1990, citado en Gómez Pérez, 2004: 108). Aplicada al

desarrollo de ontologías, una metodología proporciona un conjunto de técnicas y métodos que indican cómo debe desarrollarse. Un **método** es un proceso o procedimiento ordenado empleado en el desarrollo de un producto y una **técnica** se define como un procedimiento técnico y de gestión que permite conseguir un objetivo en una determinada aplicación. Un **proceso**, por su parte, es una función que debe llevarse a cabo en el ciclo de vida del software, compuesto por actividades. Por su parte, una **actividad** es cada una de las tareas que constituyen un proceso, y a su vez una **tarea** es una misión bien definida para uno o más miembros de un proyecto.

En el proceso de desarrollo de las ontologías, una metodología indica qué actividades deben llevarse a cabo para construir una ontología. En el marco de la metodología METHONTOLOGY (Gómez Pérez *et al*, 2004: 109) se identifican tres categorías de actividades, que son las **actividades de gestión**, las **actividades orientadas al desarrollo de la ontología** y las **actividades de soporte** de las ontologías.

1. Dentro de las **actividades de gestión** se encuentra, en primer lugar, la **planificación temporal** (*scheduling*), que identifica las tareas que han de llevarse a cabo, así como la asignación del tiempo y de los recursos necesarios. En segundo lugar, las actividades de **control** garantizan el cumplimiento de las tareas programadas y, por último, se procede a la **comprobación de la calidad** (*quality assurance*) del resultado.

2. Las **actividades orientadas al desarrollo de la ontología** se agrupan en actividades de **pre-desarrollo**, **desarrollo** y **post-desarrollo**. En la actividad de pre-desarrollo se realiza un estudio ambiental para conocer en qué plataformas y aplicaciones se integrará la ontología. Además, se lleva a cabo un estudio de viabilidad para saber si la construcción de la ontología es posible, y –en caso afirmativo– si es aconsejable.

En la fase de desarrollo propiamente dicha se identifican cuatro actividades importantes, que son:

- a) **Especificación:** Se indica el motivo por el que se construye la ontología y cuáles son los usos y usuarios para los que va destinada.
- b) **Conceptualización:** Se estructura el conocimiento del dominio de la ontología en forma de modelos significativos.

- c) **Formalización:** Se transforma el modelo conceptual en un modelo formal semi-computable.
- d) **Implementación:** Se construyen modelos computables en un lenguaje de ontologías.

En la fase de **post-desarrollo** se llevan a cabo actividades para el mantenimiento (actualizaciones y correcciones) y utilización (o reutilización) de la ontología en otras aplicaciones u ontologías.

3. Las **actividades de soporte** de la ontología engloban la adquisición del conocimiento, la evaluación, integración, documentación, fusión o alineación de varias ontologías y la gestión de la configuración (gestión de las distintas versiones). Todas estas actividades se llevan a cabo simultáneamente con las de desarrollo que acabamos de explicar.

El proceso de desarrollo de las ontologías identifica las tareas que deben llevarse a cabo, pero no el orden en el que deben ejecutarse. El ciclo de vida de la ontología identifica el conjunto de las etapas (y el orden) por las que atraviesa la ontología durante su vida, y describe las actividades de cada etapa y cómo se relacionan entre sí.

Existen diversas metodologías para el desarrollo de ontologías. Sin embargo, su presentación exhaustiva excede los límites de esta tesis. A modo de ejemplo, y puesto que consideramos que es una de las metodologías más completas, presentaremos la metodología METHONTOLOGY, desarrollada en el grupo de investigación sobre ontologías de la Universidad Politécnica de Madrid. Otros ejemplos de metodologías son el método Cyc (Lenat y Guha, 1990), el método de Uschold y King (1995), el enfoque KACTUS (KACTUS, 1996), la metodología de Grüninger y Fox (1995), el método basado en SENSUS y la metodología desarrollada en el seno del proyecto On-To-Knowledge (Staab *et al*, 2001). Todas ellas detallan los distintos procesos que se deben llevar a cabo para la construcción de una ontología desde cero, y para la reutilización de ontologías previas.

4.2.1.1 METHONTOLOGY

La metodología METHONTOLOGY (Gómez Pérez *et al*, 2004: 125-142) identifica el proceso de desarrollo de ontologías descrito en el apartado anterior, el ciclo de vida de la ontología, basado en la evolución de prototipos y las técnicas

necesarias para llevar a cabo cada actividad en las actividades de gestión, desarrollo y soporte.

Hemos elegido esta metodología puesto que permite su utilización con editores de ontologías como *Protégé* y *OntoEdit*, a pesar de haberse desarrollado como soporte tecnológico para las herramientas *ODE* y *WebODE*.

El ciclo de construcción de ontologías en *METHONTOLOGY* se basa en la evolución de prototipos. Las actividades de gestión y soporte (control de la calidad, adquisición del conocimiento, integración, evaluación, etc.) se desarrollan en paralelo con las de desarrollo (especificación, conceptualización, formalización, implementación y mantenimiento).

El aspecto que más nos interesa para nuestro trabajo es el modelado conceptual en *METHONTOLOGY*, ya que esta actividad determina el resto de la construcción de la ontología. Por ello en este apartado se presentan las tareas de que consta esta actividad. La conceptualización nos permite transformar nuestra visión inicial sobre un dominio en una especificación semiformal, utilizando un conjunto de representaciones intermedias en forma de tablas y grafos que son fácilmente comprensibles tanto por los expertos en el dominio como por los desarrolladores de la ontología. Estas representaciones intermedias resultan muy expresivas y permiten la conceptualización de los principales componentes de las ontologías (conceptos, atributos, relaciones, axiomas formales y reglas) en los lenguajes ontológicos más habituales.

A continuación presentamos las tareas que deben llevarse a cabo para la elaboración del modelo conceptual de una ontología.

Tarea 1. Construcción de un glosario de términos: Esta tarea consiste en la creación de una tabla en la que se consignan todos los conceptos relevantes del dominio (conceptos, instancias, atributos que representan las características de los conceptos, relaciones entre conceptos, etc.). A esta lista de términos se añade su descripción en lenguaje natural y sus sinónimos y siglas si los hubiere. En la Tabla 11 se reproduce un extracto (adaptado) del ejemplo propuesto por Gómez Pérez *et al* (2004: 133). Incluimos un ejemplo de cada uno de los tipos de términos que podemos definir en el glosario inicial:

Nombre	Sinónimos	Siglas	Descripción	Tipo
Location	Place	--	A position or site occupied or available for occupancy or marked by some distinguishing feature.	Concept
arrival Date	--	--	Date of arrival of the trip.	Attribute Instance
departure Place (<i>Travel, Location</i>)	--	--	The location where the travel departs from.	Relation
maximum Number of Travelers in a Plane	--	--	The maximum number of travelers in a plane at the same time.	Constant

Tabla 11. Glosario de términos para la ontología de viajes (Gómez Pérez *et al* (2004: 133).

Nótese que el término “arrival Date” puede ser tanto un atributo como una instancia, y que en el caso de la relación presentada (*departure Place*) se han indicado ya el dominio y rango de la relación.

Tarea 2. Construcción de taxonomías conceptuales que permitan clasificar los conceptos

Cuando el número de términos contenidos en el glosario sea suficientemente representativo, empiezan a construirse una o varias taxonomías de los conceptos que permiten definir la jerarquía conceptual. El enfoque para construir estas taxonomías puede ser de tres tipos:

- a) **Top-down:** Se identifican en primer lugar los conceptos más abstractos, y se van especializando en conceptos más específicos.
- b) **Bottom-up:** Se identifican en primer lugar los conceptos más específicos y se generalizan para obtener conceptos más abstractos.
- c) **Middle-out:** Se identifica en primer lugar un grupo de términos básicos, y se van especializando y generalizando convenientemente.

En METHONTOLOGY no se recomienda ningún tipo de enfoque, sino que en cada caso se aplicará el que resulte más apropiado.

Para elaborar las taxonomías de conceptos se seleccionan términos que correspondan a conceptos en el glosario. Es importante distinguir en la taxonomía

conjuntos de conceptos disjuntos, es decir, conceptos que no pueden tener instancias en común. METHONTOLOGY propone utilizar las cuatro relaciones taxonómicas definidas en la ontología OKBC (ver apartado 4.3.2.1.1), que son: *Subclass-Of*, *Disjoint-Decomposition*, *Exhaustive Decomposition* y *Partition*. La primera relación taxonómica es la más conocida, pero veamos brevemente qué se entiende por cada una de ellas:

- a) ***Subclass-Of***: Es una relación taxonómica que indica que una clase conceptual más específica es “hija” de una clase situada en un nivel superior de la taxonomía. Por ejemplo, *Iberia flight* es una subclase de *Flight* porque todas las instancias de *Iberia flight* pertenecen también a la clase *Flight*.
- b) ***Disjoint-Decomposition***: Se dice que un conjunto de clases son disjuntas si todas ellas son subclases de la misma clase. Esta clasificación no es necesariamente completa, es decir, puede haber instancias de la clase que no lo sean de ninguna de las subclases definidas. Por ejemplo, si definimos las clases *BA0068*, *BA0066* y *BA0069* como subclases de la clase *British Airways Flight*, ningún vuelo de British Airways puede ser al mismo tiempo *BA0068*, *BA0066* y *BA0069*. Sin embargo, puede haber otros vuelos de British Airways con otros códigos.
- c) ***Exhaustive-Decomposition***: Esta relación taxonómica define un conjunto de clases que son subclases de la misma clase de manera completa, es decir, que todos los ejemplos de la superclase deben pertenecer a alguna de las subclases en que se descompone la clase, si bien las subclases pueden tener ejemplos en común. Por ejemplo, los conceptos *Economy Trip*, *Business Trip* y *Luxury Trip* conforman una descomposición exhaustiva del concepto *Travel Package*, es decir, todas las subclases de *Travel Package* deben pertenecer al menos a una de esas tres subclases. Sin embargo, pueden pertenecer a dos de ellas simultáneamente (un viaje de negocios puede ser tanto un viaje económico como un viaje muy caro).
- d) ***Partition***: Esta relación taxonómica define el conjunto de clases disjuntas como subclases de la superclase. Esta clasificación es

completa, es decir, la superclase resulta de la unión lógica de todas las clases que pertenecen al conjunto. En este caso las subclases no pueden tener ninguna instancia en común. Por ejemplo, el concepto *Flight* tiene dos subclases *International Flight* y *Domestic Flight*, que cubren todos los posibles conceptos dependientes de *Flight*. Sin embargo, ningún vuelo puede ser a la vez un vuelo internacional y un vuelo nacional.

El producto final de esta tarea será una taxonomía de conceptos organizados en función de estas cuatro relaciones taxonómicas. Todos los lenguajes ontológicos y herramientas de desarrollo permiten tener en cuenta la relación taxonómica *SubclassOf*. Sin embargo, no todas las herramientas permiten tener en cuenta las clases disjuntas, la descomposición exhaustiva y las particiones. En función del tipo de lenguaje y herramienta que se pretenda emplear, habrá que determinar qué tipos de relaciones taxonómicas se van a utilizar.

Antes de proseguir conviene evaluar el resultado de la taxonomía, ya que pueden producirse errores como circularidad en la jerarquía o instancias comunes en una partición, entre otros.

Tarea 3. Construcción de diagramas de relaciones binarias *ad hoc*

En esta tarea se identifican las relaciones entre los conceptos de la ontología y con conceptos de otras ontologías. Una vez construida y evaluada la taxonomía de conceptos, hemos de construir diagramas de relaciones binarias *ad hoc* que permitan observar otros tipos de relación no taxonómica entre conceptos pertenecientes a la misma o a distintas taxonomías. METHONTOLOGY propone la creación de diagramas que reflejen estas relaciones, que posteriormente se pulirán en la tarea 6. Un ejemplo de relación binaria sería la que se entabla entre un viaje y el lugar de destino. En la ontología de ejemplo propuesta por Gómez Pérez *et al* (2004: 135) se definen como *departure Place* y *arrival Place*, y se establecen entre las clases *Travel* y *Location*. Gráficamente esta relación puede representarse del modo siguiente:

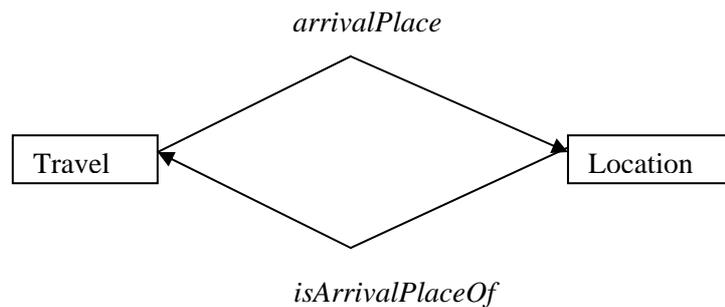


Fig. 14. Representación gráfica de la relación *arrivalPlace* y su inversa *isArrivalPlaceOf* (Gómez Pérez *et al*, 2004: 135).

En esta tarea se tiene en cuenta la posible existencia de relaciones inversas (en este caso *is ArrivalPlaceOf*). Antes de pasar a la siguiente tarea, conviene evaluar que no existan errores en la especificación de los dominios y rangos de las relaciones binarias.

Tarea 4. Construcción de un diccionario de conceptos. Una vez generadas las taxonomías y los diagramas de las relaciones, deben especificarse las propiedades y relaciones que permitan describir cada concepto de la taxonomía. Esta información se recoge en un diccionario de conceptos, que es una tabla que contiene todos los conceptos del dominio, sus relaciones, sus instancias y sus atributos de clase o de instancia. A continuación ofrecemos la fila del diccionario de conceptos que recoge el concepto *Travel* en la ontología que utilizan como ejemplo (Gómez Pérez *et al*, 2004: 136):

Nombre del concepto	Atributos de clase	Atributos de las instancias	Relaciones
Travel	--	arrival Date company Name departure Place return Fare single Fare	arrival Place departure Place

Tabla 12. El concepto *Travel* en el diccionario de conceptos de la ontología de viajes (Gómez Pérez *et al*, 2004: 136).

Tarea 5. Definición detallada de las relaciones binarias *ad hoc*

El objetivo de esta tarea es describir en detalle todas las relaciones binarias *ad hoc* incluidas en el diccionario de conceptos, y recoger los resultados de esta descripción en una tabla. Para cada relación binaria se especifica su nombre, los nombres de los conceptos de partida y llegada (dominio y rango), su cardinalidad (es decir, el número mínimo y/o máximo de elementos de partida y llegada que

puede tener una relación), su relación inversa, si es que existe, y sus propiedades matemáticas (simetría y transitividad, fundamentalmente). A continuación recogemos dos ejemplos de las relaciones binarias presentadas para la ontología de viajes. La relación *sameFlights* es una relación de equivalencia o sinonimia, y la relación *arrivalPlace* es la que hemos presentado en la Fig. 14:

Nombre de la relación	Concepto de partida (dominio)	Cardinalidad del la fuente (máxima)	Concepto de llegada (rango)	Propiedades matemáticas	Relación inversa
Same Flight as	Flight	N	Flight	Simétrica Transitiva	--
Arrival Place	Travel	1	Location	--	is Arrival Place of

Tabla 13. Extracto de la tabla de relaciones binarias *ad hoc* para la ontología de viajes (Gómez Pérez *et al*, 2004: 137).

Tarea 6. Definición detallada de los atributos de las instancias

El objetivo de esta tarea es describir exhaustivamente todos los atributos de las instancias definidos en el diccionario conceptual mediante una tabla de atributos. Antes de continuar, cabe aclarar la distinción entre atributos de instancia y atributos de clase. Los primeros hacen referencia a las propiedades (lo que en terminología solemos denominar *características*) que poseen las instancias pertenecientes a una clase determinada, que pueden ser distintos para cada una de las instancias del concepto. Los atributos de clase describen conceptos y adquieren su valor en el nivel de la clase. Los atributos de clase no son heredados por las subclases ni por las instancias de una clase determinada.

La tabla detallada de los atributos de las instancias debe contener los siguientes datos: el nombre del atributo, el nombre del concepto al que se asocia, el tipo de valor que puede tener, la unidad de medida, la precisión y el rango de valores (en el caso de valores numéricos), así como la cardinalidad máxima y mínima. Otros tipos de información que puede contener esta tabla son los atributos de instancia y de clase y las constantes, fórmulas o reglas que permiten inferir sus valores, así como las referencias utilizadas para definir el atributo.

En la Tabla 14 se muestran dos ejemplos de atributos de instancias de la ontología de viajes. Se trata de un atributo cuyo tipo de valor es una cifra con decimales (*return fare*) y de otro cuyo tipo de valor es una cadena de caracteres (*company Name*).

Nombre del atributo de instancia	Nombre del concepto	Tipo de valor	Unidad de medida	Precisión	Rango de valores	Cardinalidad
return Fare	Travel	Float	Currency Quantity	0.01	--	(0,1)
company Name	Travel	String	--	--	--	(0,N)

Tabla 14. Extracto de la tabla de atributos de las instancias para la ontología de viajes (Gómez Pérez *et al*, 2004: 138).

Tarea 7. Definición detallada de los atributos de clase

Como hemos visto en la tarea anterior, los atributos de clase describen conceptos y adquieren su valor en la clase en la que se definen, y no en sus instancias. Los atributos de clase no son heredados por las subclases ni por las instancias. La información sobre los atributos de clase recoge los datos siguientes: nombre del atributo, concepto en el que se define, tipo de valor que puede tener, unidad de medida, precisión, cardinalidad y valores que posee. En la Tabla 15 se ofrecen dos ejemplos, el atributo *company Name* de la clase *Iberia Flight* y el atributo *number of Stars* de la clase *Three-star Hotel*:

Nombre del atributo	Definido en el concepto	Tipo de valor	Unidad de medida	Precisión	Cardinalidad	Valores
company Name	Iberia Flight	String	--	--	(1,1)	IB
number of Stars	Three-star Hotel	Integer	star	1	(1,1)	3

Tabla 15. Extracto de la tabla de atributos de clase para la ontología de viajes (Gómez Pérez *et al* 2004: 138).

Tarea 8. Definición detallada de las constantes

El objetivo de esta tarea es describir de manera detallada cada una de las constantes definidas en el glosario de términos. La información consignada en esta tabla es la siguiente: nombre de la constante, tipo de valor, valor y unidad de medida. También pueden indicarse los atributos que pueden inferirse utilizando esta constante. El único ejemplo de constantes que se ha definido en la tabla de términos es el número máximo de viajeros de un avión, que se presenta en la Tabla 16:

Nombre	Tipo de valor	Valor	Unidad de medida
Maximum number of Travelers in a plane	Integer	200	person

Tabla 16. Extracto de la tabla de constantes para la ontología de viajes (Gómez Pérez *et al* 2004: 139).

Las tareas que se describen a continuación permiten la definición de los componentes de las ontologías para el modelado de ontologías pesadas. Se trata de la definición de axiomas formales (expresiones lógicas que son siempre verdaderas y que se emplean generalmente para especificar restricciones en la ontología) y reglas (utilizadas para inferir conocimiento en la ontología). METHONTOLOGY propone que se definan en paralelo una vez que se hayan definido los conceptos, atributos, constantes y relaciones *ad hoc*.

Tarea 9. Definición de axiomas formales

METHONTOLOGY propone que se recojan en una tabla las siguientes informaciones sobre axiomas formales: nombre, descripción en lenguaje natural, expresión lógica (en lógica de primer orden) que describe formalmente el axioma, conceptos, atributos y relaciones a los que se refiere, y variables utilizadas. El ejemplo de axioma formal propuesto por Gómez Pérez *et al* (2004: 139) para la ontología de viajes es el axioma *Train inside Europe*, que establece que los trenes que parten de un lugar situado en Europa deben llegar a otro lugar ubicado en Europa.

Tarea 10. Definición de reglas: En esta tarea se definen las reglas de inferencia necesarias para la ontología y se describen por medio de una tabla de reglas. La información recogida para cada regla es la siguiente: nombre de la regla, descripción de la regla en lenguaje natural, expresión que describe la regla formalmente, los conceptos, atributos y relaciones a los que se aplica, y las variables empleadas en la expresión. El ejemplo propuesto por Gómez Pérez *et al* (2004) es la regla que establece que todos los cruceros cuyo punto de partida esté situado en Europa serán operados por la compañía Costa Cruises, lo que permite inferir el nombre de la compañía que gestiona un crucero que parte de Europa.

Tarea 11. Definición de instancias

La última tarea descrita para el proceso de conceptualización en METHONTOLOGY es la definición de instancias, que –como los pasos anteriores– tiene como objetivo la construcción de una tabla con las instancias que aparecen en el diccionario de la ontología. Para cada instancia de la ontología se ha de definir la siguiente información: su nombre, el nombre del concepto al que se asocia y los valores de sus atributos, si se conocen. Ofrecemos en la Tabla 17 un ejemplo de la ontología de viajes, correspondiente a un vuelo concreto de la compañía American Airlines para la fecha 8 de febrero de 2002:

Nombre de la instancia	Nombre del concepto	Atributo	Valores
AA7462_Feb08_2002	AA7462	company Name	American Airlines
		departure Date	08/02/2002
		arrival Date	08/02/2002
		single Fare	300

Tabla 17. Extracto de la tabla de atributos de las instancias para la ontología de viajes (Gómez Pérez *et al* 2004: 141).

Valoración

En este apartado hemos repasado los principios que rigen el diseño de ontologías, y hemos presentado la metodología para el desarrollo de ontologías METHONTOLOGY. Como hemos apuntado, el desarrollo de una ontología es una tarea ardua, por lo que es muy importante tomar las decisiones sobre el proceso de trabajo de una manera estructurada. Con esa finalidad se han desarrollado metodologías que organizan el flujo de trabajo.

En nuestro estudio nos hemos fijado en la metodología METHONTOLOGY porque las tareas que propone se asemejan bastante al proceso de trabajo terminológico. Como vemos, el proceso de conceptualización en la metodología METHONTOLOGY consiste en definir una serie de representaciones intermedias en forma de tablas para cada uno de los componentes de la ontología, que posteriormente permiten llevar a cabo las actividades de formalización e implementación en un lenguaje de ontologías. Conviene recordar que cada ontología de dominio tiene sus propias necesidades, y que el modelo conceptual empleado dependerá en cada caso de dichas necesidades de representación del conocimiento.

En nuestro trabajo empírico tenemos previsto elaborar una base de conocimiento que structure los conceptos correspondientes a los productos cerámicos acabados. Dado que nuestro principal objetivo es la formalización de las relaciones entre conceptos, las tareas 3 y 5 propuestas por METHONTOLOGY nos permitirán crear una tabla que recoja las características y la definición de las relaciones en nuestra ontología.

No obstante, el nombre asignado a esta tarea del desarrollo de la ontología –conceptualización– no nos parece muy adecuado para el trabajo terminológico, puesto que hablar de “conceptos” tiene otro sentido en Terminología. Por ello, en nuestro trabajo empírico esta tarea se denominará “elaboración de un catálogo de

relaciones”, puesto que así evitamos la posible ambigüedad de la *conceptualización*.

4.2.2 Lenguajes para la creación de ontologías

Una de las principales decisiones a la hora de desarrollar una ontología es la selección del lenguaje o conjunto de lenguajes ontológicos en los que ésta se implementará. Un lenguaje ontológico se define como un medio para describir formalmente un dominio de conocimiento, con el objetivo de permitir que un ordenador pueda proporcionar distintos tipos de razonamiento sobre el dominio y sobre el conocimiento descrito en la ontología (Clark, 2003). No todos los lenguajes ontológicos tienen la misma expresividad ni llevan a cabo el razonamiento del mismo modo. Mientras que algunos lenguajes permiten la creación de ontologías tanto ligeras como pesadas, otros no contemplan la posibilidad de definir axiomas formales y funciones o reglas, por lo que su uso queda restringido al desarrollo de ontologías ligeras. Los lenguajes ontológicos emplean diversos paradigmas de representación del conocimiento. En el apartado 4.1.1 hemos descrito los paradigmas de marcos y lógica descriptiva, pero existen otros, que –como vimos en el capítulo 3– también se emplean para la formalización. A continuación resumimos en qué consisten todos ellos:

- a) **Marcos** (*Frames*): Un marco es una estructura de datos (u objeto) que recibe un nombre y se emplea para representar un concepto en un dominio. Un marco permite agrupar afirmaciones relacionadas con un concepto. Cada marco se asocia con un conjunto de *slots*, facetas y valores de los mismos.
- b) **Lógica de primer orden**: Permite describir conceptos, propiedades y relaciones mediante expresiones lógicas.
- c) **Lógica descriptiva** (*Description Logics, DL*): La lógica descriptiva es una evolución de las redes semánticas y los marcos. Es un fragmento de la lógica de primer orden que permite el razonamiento automático.
- d) **Redes semánticas** (*semantic networks*): Una red semántica es un esquema de representación del conocimiento en el que los objetos o los conceptos se almacenan como nodos de un gráfico y se enlazan entre sí por medio de relaciones etiquetadas.

- e) **Grafos conceptuales** (*conceptual graphs*): Un grafo conceptual es un sistema de representación del conocimiento en el que los conceptos se representan como rectángulos y las relaciones entre ellos se representan como círculos u óvalos. Los conceptos y relaciones se enlazan mediante flechas que indican la direccionalidad de la relación.

Esta diversidad de paradigmas de representación del conocimiento motiva que los distintos lenguajes ontológicos sean bastante dispares entre sí, si bien muchos de ellos se construyen como evoluciones de los anteriores. Todos ellos pueden describirse teniendo en cuenta dos dimensiones: el modelo de representación y los mecanismos de razonamiento. En cuanto al modelo de representación, hemos de tener en cuenta cómo se implementan los diferentes componentes de las ontologías en función del paradigma de representación del conocimiento subyacente. Estos componentes son, como ya hemos visto en el apartado 4.1.1, los conceptos y sus atributos, las relaciones entre conceptos, las funciones, los axiomas formales y las instancias. En cuanto a los mecanismos de razonamiento, los aspectos que deben considerarse son: los motores de inferencia asociados a cada lengua, la presencia o no de un clasificador automático, la gestión del tipo de herencia entre los conceptos y sus atributos (simple o múltiple) y la posibilidad de incluir excepciones y restricciones (*constraints*).

Dentro de los lenguajes ontológicos podemos distinguir entre **lenguajes tradicionales** y **lenguajes de marcado**. Los lenguajes tradicionales emplean los paradigmas de representación del conocimiento mencionados anteriormente. Los lenguajes de marcado (*ontology markup languages*) han establecido los fundamentos de la Web Semántica³⁸, y permiten anotar semánticamente los recursos web apoyándose en una ontología. Su sintaxis se basa en los lenguajes de marcado empleados en la Web (principalmente XML y HTML). Estos lenguajes son menos expresivos que los tradicionales, si bien poco a poco se les van añadiendo nuevas características lógicas.

³⁸ La Web Semántica es una extensión de la World Wide Web actual, en la que la información posee un significado definido, que permite a los ordenadores procesarlo. Su principal característica es su universalidad y su carácter descentralizado (Berners-Lee *et al*, 2001).

A continuación enumeramos los principales lenguajes ontológicos, indicando entre paréntesis el paradigma de representación del conocimiento que emplean y una referencia al recurso web en el que se describen en detalle. Su descripción exhaustiva excede el alcance de esta tesis, por lo que para obtener una visión más amplia remitimos a dichas referencias³⁹.

Lenguajes ontológicos tradicionales:

- a) **Ontolingua** (basado en KIF⁴⁰ y en la ontología de marcos):
<http://ontolingua.stanford.edu>
- b) **LOOM** (Lógica descriptiva):
<http://www.isi.edu/isd/LOOM/LOOM-HOME.html>
- c) **OKBC** (basado en marcos): Más que un lenguaje es un protocolo para acceder a bases de conocimiento almacenadas en distintos sistemas de representación del conocimiento:
<http://www.ai.sri.com/~okbc/>
- d) **OCML** (basado en marcos): <http://kmi.open.ac.uk/projects/ocml/>
- e) **FLogic**: basado en marcos y lógica de primer orden):
http://www.ontoprise.de/documents/tutorial_flogic.pdf

Lenguajes ontológicos de marcado:

- a) **SHOE**: <http://www.cs.umd.edu/projects/plus/SHOE/>
- b) **XOL**: <http://www.ai.sri.com/~pkarp/xol/>
- c) **RDF(S)**: <http://www.w3.org/RDF/>
- d) **OIL**: <http://www.ontoknowledge.org/oil/>
- e) **DAML+OIL**: <http://www.w3.org/Submission/2001/12/>
- f) **OWL**: <http://www.w3.org/2004/OWL/>

³⁹ Para obtener una visión exhaustiva de las características de los lenguajes ontológicos remitimos al cuadro comparativo de Gómez Pérez *et al* (2004: 287).

⁴⁰ KIF (Knowledge Interchange Format) es un lenguaje basado en la lógica de predicados con extensiones para definir términos, metaconocimiento, conjuntos, razonamientos no monotónicos, etc; y pretende ser un lenguaje capaz de representar la mayoría de los conceptos y distinciones actuales de los lenguajes más recientes de representación del conocimiento. Se trata de un lenguaje diseñado para intercambiar conocimiento entre sistemas de computación distintos, diferentes lenguas, etc; y no para la interacción entre seres humanos (Lamarca Lapuente, 2007).

Todos los lenguajes ontológicos permiten representar conceptos y sus atributos (también llamados *slots*, roles y propiedades, dependiendo del lenguaje). En cuanto a las facetas (restricciones de los atributos), todos los lenguajes permiten algún tipo de restricción de las facetas (cardinalidad y tipo de valor, entre otros).

La estructuración de las taxonomías de conceptos se realiza en todos los lenguajes mediante la relación *SubclassOf*, pero no todos ellos permiten elaborar taxonomías de conceptos disjuntos, particiones y descomposiciones exhaustivas de conceptos.

En cuanto a las relaciones, todos los lenguajes permiten representar relaciones binarias entre conceptos. Mientras que en algunos lenguajes los atributos y las relaciones se representan por medio de las mismas primitivas con distintos rangos, en otros están separados. Las relaciones con una *aridad* superior (con más de dos argumentos) se representan generalmente mediante reificación de las relaciones (es decir, considerando las relaciones como conceptos). En los lenguajes basados en la lógica descriptiva se pueden establecer jerarquías de relaciones.

Las funciones son, como hemos visto, un tipo específico de relación, si bien existen bastantes diferencias entre los lenguajes con respecto a la aridad de las funciones.

Todos los lenguajes permiten definir instancias de los conceptos y de las relaciones definidas en la ontología.

En lo que se refiere a axiomas formales y reglas, los lenguajes ontológicos de marcado no permiten su definición, si bien el desarrollo de las capas lógicas añadidas a estos lenguajes facilitará su definición en un futuro próximo.

Valoración

En este apartado hemos presentado de manera muy sucinta los distintos lenguajes ontológicos tradicionales y de marcado. Consideramos que no es nuestra misión en este trabajo detallar todos y cada uno de los lenguajes ontológicos, si bien es conveniente conocer su filosofía porque en ellos se basan las herramientas para el desarrollo de ontologías que veremos en el apartado 4.4 y que reutilizaremos para el trabajo terminográfico en la parte empírica de esta tesis.

Los aspectos más destacables de los lenguajes ontológicos, de los que sin duda hemos de ser conscientes al desarrollar nuestra ontología, son los siguientes:

en primer lugar, todos los lenguajes ontológicos permiten representar conceptos y sus atributos. En segundo lugar, la estructuración de los conceptos siempre se hace de manera jerárquica por medio de la relación básica *SubclassOf*, que, traducida a la terminología, sería la relación que hemos denominado relación genérica. Esta relación, que vertebra toda la ontología, permite aprovechar los mecanismos de herencia, que facilitan el desarrollo de ontologías de un modo más económico.

Por último, los lenguajes ontológicos tradicionales contemplan la posibilidad de definir axiomas y funciones, que dan lugar a ontologías pesadas. Si bien la base de conocimiento que se desarrollará en el marco de este trabajo no contempla la definición de axiomas y funciones, una vez construida dicha base podremos considerar el empleo de axiomas y funciones para llevar a cabo procesos de razonamiento, lo que nos permitirá explotar al máximo las posibilidades de la misma.

4.3 Tipos de ontologías y ontologías más representativas

En el apartado 4.3.1 de esta sección presentamos la clasificación de las ontologías propuesta por Gómez Pérez *et al* (2004). En el apartado 4.3.2 se ofrece una breve descripción panorámica de las ontologías más representativas de los tipos propuestos.

4.3.1 Tipología de las ontologías

Las ontologías pueden clasificarse atendiendo a diversos criterios. No es nuestra intención explicar exhaustivamente todos los tipos de ontología, pero sí que creemos necesario establecer una serie de distinciones básicas. De nuevo recurrimos a la propuesta de clasificación de Gómez Pérez *et al* (2004), por su exhaustividad y porque unifica distintas clasificaciones propuestas en la bibliografía sobre ingeniería ontológica. Estos autores clasifican las ontologías atendiendo a dos criterios: la riqueza de la estructura interna y la temática de la conceptualización.

4.3.1.1 Ontologías según la riqueza de la estructura interna

Como ya hemos apuntado en apartados anteriores de este capítulo, atendiendo a la riqueza de su estructura interna, es decir, a su grado de formalidad

descriptiva, se puede distinguir entre ontologías “**ligeras**” y “**pesadas**”. Las primeras incluyen conceptos, taxonomías de conceptos, relaciones entre conceptos y propiedades para describirlos. Las ontologías “pesadas” añaden a estos componentes básicos axiomas formales y restricciones. Los distintos tipos de ontologías se organizan en un *continuum*, que va desde los vocabularios controlados (muy informales) hasta las ontologías que expresan restricciones lógicas generales (muy formales). Todas ellas se definen brevemente a continuación:

- a) **Vocabularios controlados:** Consisten en una lista finita de términos. Un ejemplo típico son los catálogos.
- b) **Glosarios:** Consisten en una lista de términos cuyo significado se especifica en lenguaje natural, es decir, contienen una definición.
- c) **Tesauros:** Listas de términos que proporcionan información sobre relaciones semánticas, como las relaciones de sinonimia, pero no proporcionan una jerarquía explícita.
- d) **Jerarquías informales del tipo IS-A:** Este tipo de jerarquías agrupa conceptos básicos sobre un tema, si bien los elementos situados en un nivel inferior de la jerarquía no son necesariamente subclases estrictas del elemento situado en un nivel superior. Un ejemplo sería la ordenación del directorio Yahoo.
- e) **Jerarquías formales del tipo IS-A:** Son sistemas en los que se establece la transitividad y se explota el sistema de herencia.
- f) **Jerarquías formales del tipo IS-A que incluyen ejemplos del dominio:** Estos sistemas son similares a los descritos en el apartado anterior, pero en este caso incluyen ejemplos (instancias) para las distintas clases.
- g) **Marcos:** En este tipo de ontología se incluyen las clases y sus propiedades, que pueden ser heredadas por las subclases de la taxonomía.
- h) **Ontologías que expresan restricción de valores:** En estas ontologías puede restringirse el tipo de valor de un atributo, por ejemplo, se puede especificar que el valor de la propiedad *fecha de llegada* debe ser una fecha.

- i) **Ontologías que expresan restricciones lógicas generales:** Este tipo de ontologías son las más expresivas. Sus creadores pueden establecer restricciones en lógica de primer orden entre los conceptos utilizando los distintos lenguajes para el modelado de ontologías.

4.3.1.2 *Ontologías según la temática de la conceptualización*

Atendiendo a la temática de la conceptualización, Gómez Pérez *et al* (2004: 29-34) distinguen ocho tipos de ontologías:

- a) **Ontologías de representación del conocimiento:** Este tipo de ontologías, también denominadas en ocasiones **metaontologías**, capturan la representación de las primitivas utilizadas para formalizar el conocimiento según un paradigma de representación determinado. Los ejemplos más representativos de este tipo de ontologías son la *Frame Ontology* y la *OKBC Ontology*. Estas ontologías proporcionan definiciones formales de las primitivas de representación generalmente empleadas en los lenguajes basados en marcos (clases, subclasses, atributos, valores, relaciones y axiomas). Más adelante volveremos sobre este tipo de ontologías, puesto que son la base que permite la construcción de otras ontologías.
- b) Las **ontologías generales o comunes** se utilizan para representar el conocimiento general de manera que sea reutilizable en diversos dominios. Incluyen vocabulario sobre objetos, eventos, tiempo, espacio, causalidad, comportamiento, función, meronimia, etc. Un ejemplo de este tipo de ontologías es la *Mereology Ontology*, que define la relación parte-todo y sus propiedades.
- c) **Ontologías de nivel superior (*top-level o upper-level ontologies*):** Este tipo de ontologías describe conceptos muy generales y proporciona nociones que permiten relacionar todos los conceptos con la raíz en una ontología. Algunos ejemplos de ontologías de nivel superior son la ontología *Mikrokosmos* (descrita en detalle en el apartado 3.4.2.1), la ontología *Sensus* o la ontología *Cyc*. Dado que cada una de ellas sigue un criterio diferente para la definición

de los conceptos más generales, se ha creado un grupo de trabajo que está tratando de unificarlas en una ontología estándar, la SUO (Standard Upper Ontology), a partir de la cual puedan estructurarse todas las ontologías de dominio.

- d) Las **ontologías de dominio** son reutilizables en un dominio específico (medicina, ingeniería, derecho, etc.). Estas ontologías proporcionan vocabularios sobre conceptos y sus relaciones específicas, sobre las actividades que se desarrollan y sobre las teorías y principios elementales que rigen en ese dominio concreto. Los conceptos de las ontologías de dominio suelen ser especializaciones de los conceptos ya definidos en las ontologías de nivel superior. Algunos ejemplos de ontologías de dominio son la ontología UNSPSC (United Nations Standard Products and Services Codes) en el ámbito del comercio electrónico, la ontología UMLS en medicina (Unified Medical Language System) y la ontología Chemicals en el ámbito de la química.
- e) Las **ontologías de tareas** describen el vocabulario asociado a una tarea o actividad genérica por especialización de los términos de una ontología de nivel superior. Proporcionan un vocabulario sistemático de los términos empleados para resolver problemas asociados con tareas. Un ejemplo sería la *Scheduling Task Ontology*.
- f) Las **ontologías de dominio-tarea** son ontologías de tareas asociadas a un dominio específico. Son independientes de la aplicación y no son reutilizables en todos los dominios. Un ejemplo serían las que propone Timpf (2002) para la ontología del dominio de la Geología GIS.
- g) **Ontologías de método:** Este tipo de ontologías proporciona definiciones de los conceptos relevantes y de las relaciones que se aplican para especificar un proceso de razonamiento que permite llevar a cabo una determinada tarea.
- h) Las **ontologías de aplicación** contienen todas las definiciones que permiten modelar el conocimiento necesario para una aplicación determinada. Generalmente amplían y especializan el vocabulario

de una ontología de dominio o de tareas para una aplicación determinada. Por ejemplo, podría crearse una ontología de aplicación para las agencias de viajes españolas especializadas en viajes a Norteamérica.

4.3.2 Descripción de algunas ontologías representativas

La descripción exhaustiva de todos los componentes de las ontologías excede los límites de este trabajo, pero aun así resumimos en este apartado cómo se formalizan las ontologías más relevantes, haciendo hincapié en la formalización de las relaciones conceptuales. Para obtener una visión más completa sobre estos aspectos, recomendamos la lectura del capítulo 2 de Gómez Pérez *et al* (2004: 47-106), en el que se presentan algunas ontologías representativas, bien porque se utilizan en una gran cantidad de proyectos, bien porque sus aportaciones teóricas son muy relevantes para el campo de la ingeniería ontológica.

En nuestro trabajo tomaremos como punto de partida esta descripción de las ontologías, completada con la consulta de cada una de ellas siempre que sea posible.

4.3.2.1 Ontologías de representación del conocimiento (metaontologías)

Como hemos visto, las ontologías de representación del conocimiento (*KR ontologies*) contienen la descripción formal de las primitivas empleadas para formalizar el conocimiento mediante un paradigma de representación.

En las ontologías de representación del conocimiento los conceptos aparecen representados como clases, en las que se especifican sus atributos. Un rasgo común a la mayoría de las ontologías es la organización de las clases y de los ejemplos mediante taxonomías. Esta organización se realiza en función de la relación lógica que se establece entre una clase situada en un nivel superior de abstracción y otra situada en un nivel inferior.

Las ontologías de representación del conocimiento más relevantes son: la *Frame Ontology* y la ontologías OKBC, la ontología RDF y RDF Schema, la ontología OIL, la ontología DAML+OIL y la ontología OWL. A continuación describimos brevemente cada una de ellas.

4.3.2.1.1 *Frame Ontology* y ontología OKBC

La *Frame Ontology* (desarrollada por Gruber, 1993) está escrita en formato KIF (*Knowledge Interchange Format*), y recoge las convenciones más habituales para la organización del conocimiento utilizadas en las representaciones basadas en marcos. Su objetivo es unificar la semántica de las primitivas más empleadas en el paradigma de marcos y permitir el desarrollo de las ontologías que emplean este paradigma.

En 1997 algunas de las primitivas se cambiaron a la ontología OKBC (*Open Knowledge Base Connectivity*) y en la actualidad la *Frame Ontology* (FO) contiene la ontología OKBC.

La FO contiene 23 clases, 31 relaciones y 13 funciones y la OKBC contiene 8 clases, 36 relaciones y tres funciones.

Las primitivas más representativas en estas dos ontologías son las siguientes:

- a) **Clases, particiones de clases e instancias:** Las dos primitivas identificadas en estas dos ontologías son *Class*, término que define a un conjunto de individuos o ejemplos e *Individual*, que define a un ejemplo de una clase.
- b) **Taxonomías de clases:** Las taxonomías permiten organizar las clases y las instancias que aparecen en las taxonomías. Las relaciones taxonómicas más importantes son *SubclassOf*, *SuperclassOf*, *Disjoint-Decomposition*, *Exhaustive-Decomposition*, *Partition* e *InstanceOf*
- c) **Relaciones y sus propiedades:** En la FO y en la ontología OKBC las relaciones se expresan como clases que se ordenan en una jerarquía de relaciones que poseen distintas propiedades (simetría, transitividad y reflexividad). Algunas de las primitivas que definen relaciones son: *Relation*, *Subrelation-Of*, *Reflexive-Relation*, *Irreflexive-Relation*, *Symmetric-Relation*, *Antisymmetric-Relation*, *Asymmetric-Relation*, *Transitive-Relation*, *Equivalence-Relation*, *Partial-Order-Relation* y *Total-Order-Relation*.
- d) **Atributos:** Definen una característica de una clase, que a su vez es heredada por sus subclases. Las dos primitivas que los definen en la ontología OKBC son *Template-Slot-Of* y *Slot-Of*

- e) **Facetas y tipos de facetas:** las facetas son propiedades de los atributos. En la FO se definen como relaciones ternarias entre un marco (que puede ser una clase o un individuo), un atributo y la faceta. Algunas de las primitivas que se refieren a facetas son las que definen la cardinalidad de un *slot*, el tipo de *slot* y los valores por defecto.

4.3.2.1.2 RDF (Resource Description Framework) y RDF Schema

El *Resource Description Framework* (RDF) es una recomendación del Consorcio W3C⁴¹ desarrollada para describir recursos web por medio de metadatos.

El modelo de datos del RDF es equivalente al paradigma de representación del conocimiento en una red semántica (*semantic network*). Una red semántica es un grafo etiquetado compuesto por un conjunto de nodos y un conjunto de arcos unidireccionales, cada uno de los cuales tiene un nombre. Los nodos representan conceptos, instancias de conceptos y valores de las propiedades. Los arcos representan propiedades de los conceptos y relaciones entre ellos. Este paradigma es menos expresivo que el basado en marcos, ya que no permite representar valores por defecto ni restricciones en la cardinalidad de los atributos.

El modelo de datos del RDF consta de tres componentes:

- a) **Recursos** (*resources*): Cualquier tipo de datos, descritos con expresiones en RDF y denominados URIs (*Uniform Resource Identifiers*)
- b) **Propiedades** (*properties*, también llamados *predicates*): Atributos y relaciones utilizados para describir un recurso.
- c) **Afirmaciones** (*statements*): Las afirmaciones asignan un valor a una propiedad en un recurso específico. Los *statements* en RDF poseen sujetos, propiedades y objetos.

⁴¹ El Consorcio World Wide Web (W3C) es un consorcio internacional cuya misión es guiar la Web hacia su máximo potencial a través del desarrollo de protocolos y pautas que aseguren el crecimiento futuro de la Web. Para más información, puede consultarse la siguiente dirección: <http://www.w3.org/>.

Los recursos pueden ser objetos de una afirmación RDF, y las afirmaciones pueden convertirse a su vez objetos mediante un proceso de reificación⁴².

La ontología de representación del conocimiento RDF está escrita en el lenguaje RDFS y contiene quince primitivas (siete clases, siete propiedades y una instancia). Este modelo de representación no proporciona primitivas para definir las relaciones entre propiedades y recursos. Esta limitación se supera por medio de RDFSschema, que es el lenguaje de descripción del vocabulario RDF, que añade 16 primitivas nuevas (seis clases y nueve propiedades).

Las primitivas RDFS se agrupan en clases y propiedades centrales (*core classes and properties*), clases y propiedades contenedor (*container classes and properties*), colecciones, vocabulario de reificación y propiedades de utilidad (*utility properties*). Las propiedades centrales permiten establecer, entre otras cosas, taxonomías de clases y de propiedades, el rango y el dominio a los que se aplica una propiedad.

RDF y RDFS definen las relaciones por medio de propiedades. En principio no parecen ontologías muy expresivas en cuanto a las relaciones entre clases se refieren.

4.3.2.1.3 Ontología de representación del conocimiento OIL (Ontology Inference Layer)

El lenguaje OIL es una extensión del RDF(S), que añade más primitivas de representación del conocimiento basado en marcos, evitando el mecanismo de reificación. OIL se ha desarrollado según una estructura de capas, en la que cada nueva capa añade complejidad y funcionalidad a la capa inferior. Las capas de OIL son: CoreOIL, StandardOIL, InstanceOIL y HeavyOIL. En CoreOIL se agrupan las primitivas que tienen una correspondencia directa en RDF(S), pero excluyendo el mecanismo de reificación. En StandardOIL se añaden primitivas basadas en marcos. InstanceOIL permite definir instancias de conceptos y roles.

⁴² La reificación consiste en tratar a un concepto abstracto, una relación o una proposición como si fuera un objeto de pleno derecho. Esta técnica se emplea con frecuencia en los lenguajes de representación del conocimiento para referirse a entidades abstractas como si fueran objetos concretos, con el fin de facilitar el razonamiento (definición extraída de http://boxmind.leeds.ac.uk/lectures/frank_van_harmelen/glossary.htm#reification).

Por último, HeavyOIL está reservado para futuras ampliaciones (reglas, metaclasses, etc.).

La ontología StandardOIL consiste en 37 clases y 19 propiedades, la mayoría de las cuales están definidas en RDF(S), agrupadas en seis grupos de primitivas, que son:

- a) Clases para definir expresiones de tipo concretas.
- b) Clases para definir expresiones de clases.
- c) Clases para definir características matemáticas de las propiedades (transitividad, funcionalidad⁴³ y simetría).
- d) Clases para definir axiomas que permiten definir el conocimiento disjunto y exhaustivo en las taxonomías de clases.
- e) Clases para definir tipos de datos.
- f) Clases predefinidas: Se trata de las clases *Top* y *Bottom*, que corresponden a la clase más general que contiene el resto de las clases (*Top*) y a una clase vacía (*Bottom*) que está contenida en todas las demás clases.

Las 19 propiedades contenidas en la ontología StandardOIL permiten relacionar las clases entre sí y asignar valores a las propiedades de las clases.

4.3.2.1.4 Ontología de representación del conocimiento DAML+OIL

DAML+OIL es también una extensión de RDF(S), pero a diferencia de OIL, este lenguaje no está dividido en capas, sino que amplía directamente la ontología RDF(S) mediante extensiones en lógica de descripción. Esta ontología contiene 53 primitivas (14 clases, 38 propiedades y una instancia), que se dividen en los siguientes grupos:

- a) Clases para definir clases, restricciones y tipos de datos
- b) Clases para definir propiedades (transitividad, ambigüedad, propiedad de objeto o de tipo de datos)
- c) Clases para definir contenedores
- d) Clases predefinidas
- e) Clases para definir valores literales

⁴³ Se dice que una relación es funcional si sólo puede tener un valor para cada instancia de su dominio.

f) Clases para describir ontologías

Las expresiones para definir clases en DAML+OIL se construyen con las siguientes primitivas, que son propiedades:

- a) Conjunción, disyunción y negación
- b) Colección de individuos
- c) Restricciones de las propiedades, que puede expresar restricciones de los valores, valores de los roles (*role fillers*), restricciones existenciales y restricciones de la cardinalidad (*qualified number restriction*).

Las propiedades no sólo se emplean para crear expresiones de clases, sino también para definir otras relaciones entre los componentes de las ontologías.

4.3.2.1.5 Ontología de representación del conocimiento OWL (Ontology Web Language)

El lenguaje OWL has sido desarrollado en el marco del grupo de trabajo sobre ontologías web del consorcio W3C (WebOnto). Al igual que DAML+OIL, OWL también se divide en capas: OWL Lite, OWL DL, y OWL Full. La primera permite crear taxonomías de clases y restricciones básicas. La segunda, basada en la lógica descriptiva, incluye todo el vocabulario de OWL. Por último, la capa OWL Full proporciona una mayor flexibilidad a la hora de representar ontologías.

En la ontologías OWL DL KR hay 40 primitivas, que corresponden a 16 clases y 24 propiedades, y se pueden agrupar del siguiente modo:

- a) Clases para la definición de clases y restricciones
- b) Clases para definir propiedades: Las propiedades en OWL pueden ser propiedades de objeto (que conectan dos clases entre sí), propiedades de datos (se emplean para definir propiedades que conectan clases con un *datatype*), propiedades funcionales (y sus inversas), propiedades transitivas, propiedades simétricas y propiedades de anotación.
- c) Clases que indican la diferencia (falta de igualdad) entre dos individuos
- d) Clases para definir enumeraciones de tipos de datos (*datatypes*)
- e) Clases predefinidas
- f) Clases para definir ontologías

g) Clases para describir versiones de las ontologías

Las expresiones de clases en OWL se construyen con primitivas de representación del conocimiento que son propiedades. Se agrupan en las que utilizan primitivas definidas en OWL Lite y las que utilizan primitivas definidas en OWL DL. Las primeras incluyen la conjunción (intersección) y las restricciones de las propiedades (de valor, existencial y numérica). Las definidas en DL incluyen la conjunción, la disyunción y la negación, las colecciones de individuos y las restricciones de propiedades (*role fillers, number restriction*)

OWL no difiere mucho de DAML+OIL. Las relaciones entre clases se expresan mediante propiedades que pueden ser funcionales, simétricas o transitivas, mientras que las propiedades de tipos de datos se reservan para los atributos de una clase.

4.3.2.2 Ontologías de nivel superior

Las ontologías de nivel superior (*top-level o upper-level ontologies*) describen conceptos generales que son comunes a varios dominios y proporcionan nociones a las que deberían poder asociarse los términos existentes en las ontologías de dominio. En ocasiones se construyen ontologías de dominio basándose en ellas, pero lo más habitual es que las ontologías de dominio se creen de forma independiente y más adelante se enlacen con una ontología de nivel superior.

Una ontología de nivel superior debe ser **universal** y **articulada**, es decir, cualquier concepto imaginado en una ontología específica debe ser susceptible de enlazarse a una ontología de nivel superior, independientemente de lo especializado que sea y del origen del creador de la ontología. El hecho de que sea articulada implica que existe una justificación para cada concepto dentro de la ontología y que la ontología contiene suficientes conceptos como para permitir distintos tipos de aplicación.

A continuación describimos brevemente cuatro ontologías de nivel superior: la ontología de universales y particulares, la ontología de nivel superior de Sowa, la ontología Cyc y la ontología SUO.

4.3.2.2.1 Ontologías de universales y particulares (Guarino y Welty, 2000)

Guarino y sus colaboradores han desarrollado dos ontologías siguiendo la metodología OntoClean: una ontología de universales y otra de particulares. Un universal corresponde a un concepto y un particular a un individuo. La relación entre particulares y universales se establece por medio de la relación *Instance-Of*.

La ontología de nivel superior de universales contiene conceptos cuyas instancias son universales. Se ha construido siguiendo los principios de rigidez, identidad y dependencia, que son atributos de las propiedades.

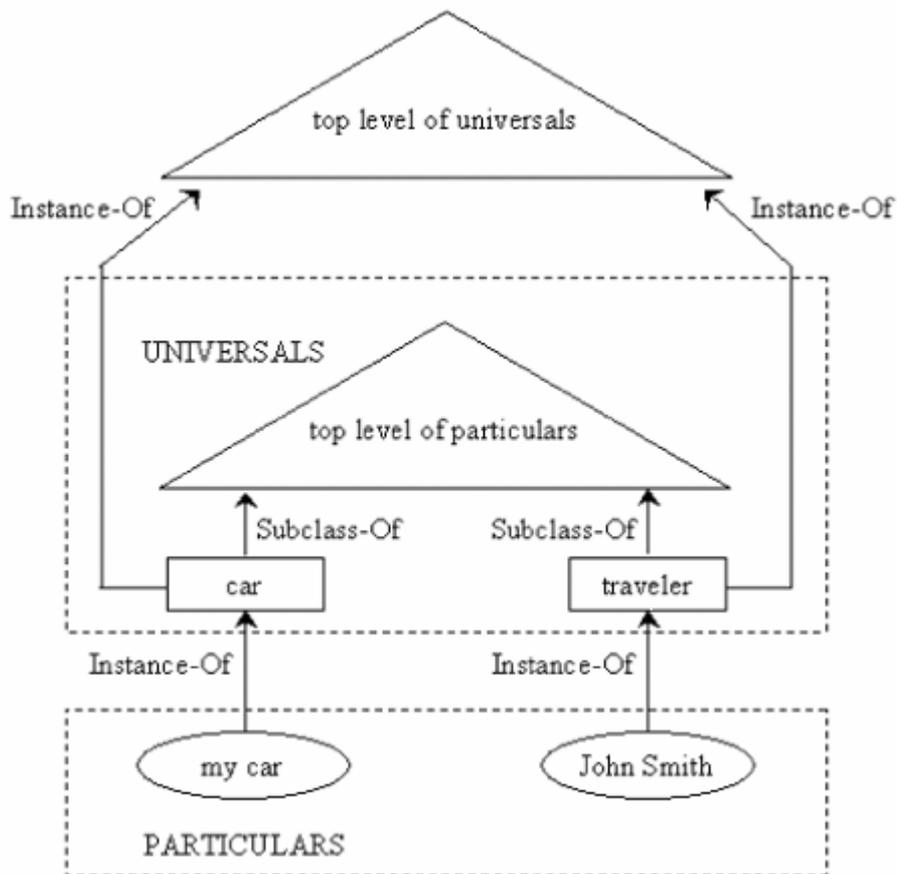


Fig. 15. Relación entre universales y particulares en la concepción de Guarino y Welty (2000).

Cada concepto de la ontología de universales tiene cuatro atributos: rigidez, proporciona identidad, conlleva identidad y dependencia.

Algunos conceptos definidos en la ontología de universales son: *property*, *role property*, *sortal property*, *non-sortal property*.

Por su parte, la ontología de nivel superior de particulares contiene conceptos generales que pueden relacionarse con conceptos de dominio mediante

la relación *Subclass-Of*. La ontología de particulares contiene tres raíces, a saber: *abstract*, *concrete* y *relation*. Esta última raíz puede dividirse en *inter-categorical relation* e *intra-categorical relation*.

En la Fig. 15 se reproduce un esquema que ilustra la relación que se establece entre particulares y universales en la concepción de Guarino y Welty (2000).

4.3.2.2 Ontología de nivel superior de Sowa

Esta ontología de nivel superior incluye las categorías y distinciones básicas derivadas de la Lógica, la Lingüística, la Filosofía y la Inteligencia Artificial. Contiene 27 conceptos ordenados en una estructura reticular en la que el concepto de nivel superior es el *universal type* (que contiene todas las posibles instancias de la ontología) y el de nivel inferior es el *absurd type* (que no tiene instancias y es una subclase de todos los conceptos de la taxonomía). Las subclases directas de las clase universal son las siguientes: *independent*, *relative*, *mediating*, *continuant*, *physical*, *abstract* y *occurrent*. Cada par de conceptos de la taxonomía tiene al menos una subclase común directa o indirecta. En esta ontología, la ubicación de nuevos conceptos se obtiene mediante la combinación de conceptos de los niveles superiores.

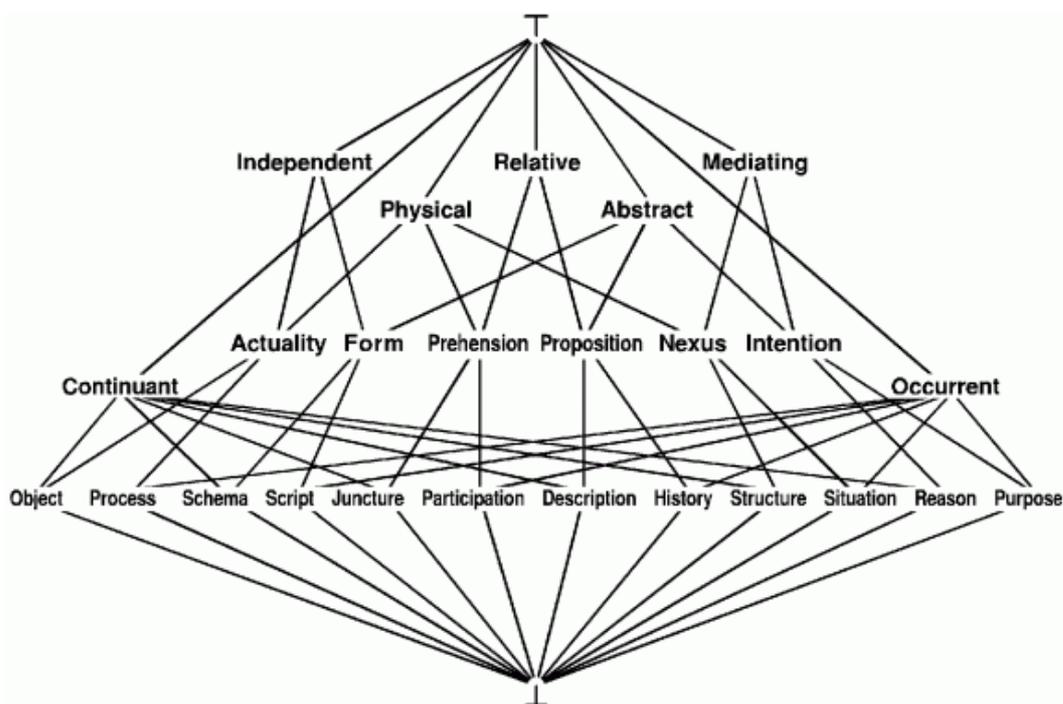


Fig. 16. Ontología de nivel superior de Sowa (2000).

En la Fig. 16 ofrecemos la representación gráfica de la jerarquía de los niveles superiores de la ontología de nivel superior propuesta por Sowa (2000):

4.3.2.2.3 **Ontología de nivel superior Cyc**

Esta ontología está contenida en la base de conocimiento Cyc (Lenat y Guha, 1990), que incorpora una gran cantidad de conocimiento común. La ontología superior Cyc contiene unos 3000 términos agrupados en 43 grupos temáticos (fundamentales, tiempo y fechas, relaciones espaciales, etc.). La raíz de la ontología es la clase *Thing*. Esta ontología de nivel superior se ha creado mediante la subdivisión del conjunto universal en pares opuestos (tangible/intangible, estático/dinámico, etc.), que se va refinando a medida que se añade nuevo conocimiento a la ontología. La ontología se divide en cientos de microteorías (grupos de afirmaciones pertenecientes al mismo dominio) y está implementada en el lenguaje CycL.

4.3.2.2.4 **Standard Upper Ontology (SUO)**

La *Standard Upper Ontology* es una gran ontología formal con fines generales, resultado de un esfuerzo común por armonizar la diversidad de ontologías de nivel superior existentes. En la actualidad se está trabajando a partir de dos metaontologías diferenciadas: IFF (*Information Flow Framework*) y SUMO (*Suggested Upper Merged Ontology*). El objetivo de SUMO es construir una ontología de nivel superior de gran alcance derivada a partir de algunas fuentes públicas como la ontología de Sowa, los axiomas temporales y las teorías de planes y procesos, entre otras.

Se trata de una ontología modular, es decir, dividida en dos subontologías: la ontología estructural y la ontología de base. En la Fig. 17 se pueden apreciar las dependencias entre ambas ontologías.

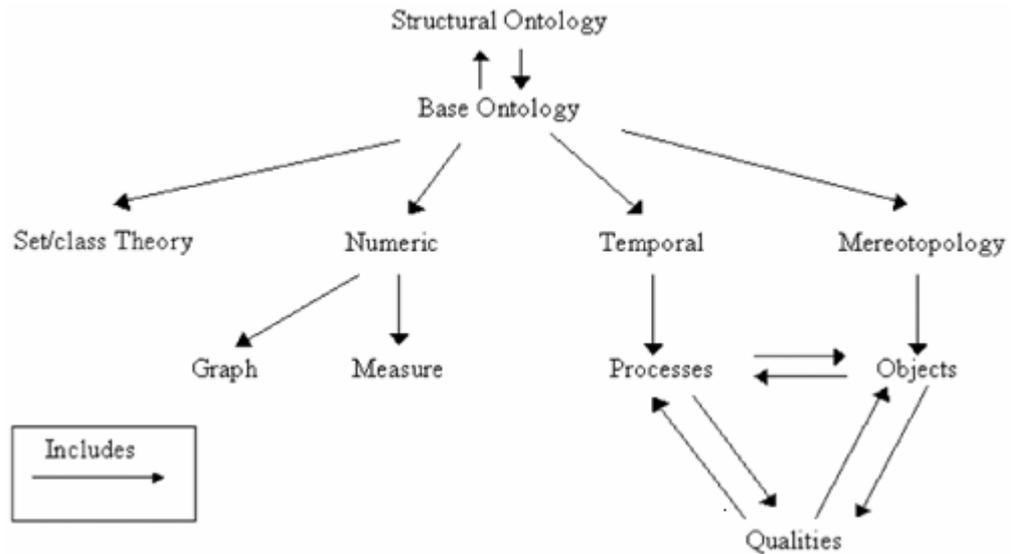


Fig. 17. Estructura modular de la ontología SUMO.

4.3.2.3 Ontologías lingüísticas

En este apartado describimos las denominadas **ontologías lingüísticas**, cuya finalidad consiste en la descripción de constructos semánticos en lugar de en el modelado de un dominio. Se aplican generalmente en el procesamiento del lenguaje natural y su principal característica es su vinculación con la semántica de las unidades gramaticales (palabras, grupos nominales, adjetivos, etc.).

Las ontologías lingüísticas más representativas son WordNet (y su versión multilingüe EuroWordNet), GUM (*Generalized Upper Model*), Mikrokosmos y SENSUS.

WordNet, EuroWordNet y Mikrokosmos ya han sido descritas en detalle en el apartado 3.4, por lo que en este apartado nos limitamos a describir brevemente las otras dos ontologías lingüísticas (GUM y Sensus).

4.3.2.3.1 Generalized Upper Model (GUM)

GUM es una ontología lingüística ligada a la semántica de los constituyentes gramaticales de la lengua. A diferencia de otras ontologías lingüísticas como WordNet, GUM no describe la semántica de las palabras, sino la que puede expresarse mediante unidades gramaticales superiores tales como grupos nominales, sintagmas preposicionales, etc. Contiene dos jerarquías, una de conceptos y otra de relaciones, que tienen su origen en la gramática funcional de Haliday (1985). La jerarquía de conceptos representa las entidades semánticas

básicas e incluye configuraciones para los procesos, así como los distintos tipos de objetos y cualidades. Una configuración es un conjunto de objetos que participan en alguna actividad o estado (por ejemplo *being&having*). La jerarquía de relaciones representa los *participantes* y las *circunstancias* implicadas en los *procesos*, y las *combinaciones lógicas* existentes entre ellos. Algunos ejemplos de participantes son *actor*, *mensaje* o *atributo*. *Compañía*, *comparación*, *causa*, *modo*, *tiempo*, *espacio*, etc. expresan circunstancias. En la Fig. 18 se muestra una figura que ilustra los primeros niveles de las jerarquías GUM.

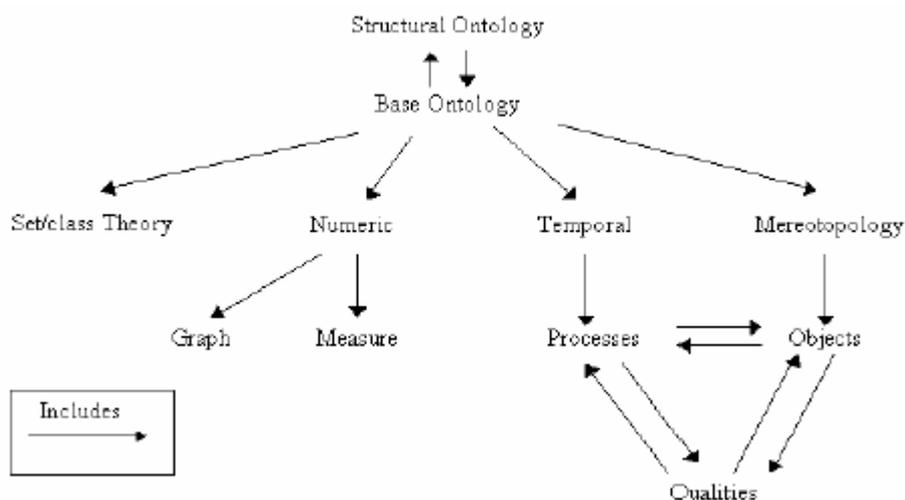


Fig. 18. Niveles superiores de la ontología GUM.

4.3.2.3.2 SENSUS

SENSUS es una ontología basada en el lenguaje natural que ofrece una estructura conceptual amplia para la traducción automática. SENSUS contiene más de 70.000 nodos que representan objetos, entidades, cualidades y relaciones. Esta ontología proporciona una base de conceptos estructurada de forma jerárquica que posee una ontología de nivel superior abstracta en la cual se representan las generalizaciones necesarias para el procesamiento del lenguaje durante la traducción automática. La región media de la ontología proporciona un marco para un modelo del mundo genérico y contiene elementos que representan el significado de un gran número de palabras en inglés. Las regiones inferiores (más específicas) de la ontología proporcionan puntos de anclaje para distintas lenguas. El contenido de SENSUS procede de la fusión de información de otras fuentes electrónicas, como el PENMAN Upper Model, ONTOS o WordNet.

4.3.2.4 *Ontologías de dominio*

Finalizamos nuestro breve repaso por las ontologías más representativas y el modo en el que definen las clases y las relaciones con la presentación sucinta de algunas de las ontologías de dominio más relevantes. Una ontología de dominio se define como un vocabulario reutilizable de los conceptos de un dominio y sus relaciones, las actividades que tienen lugar en dicho dominio, y las teorías y principios elementales que lo gobiernan. Algunas ontologías de dominio representativas son las siguientes:

En el ámbito del **comercio electrónico**, las ontologías existentes ofrecen clasificaciones de los productos y servicios disponibles. Algunos ejemplos son: UNSPSC (*United Nations Standard Products and Services Codes*), NAICS (*North American Industry Classification System*), SCTG (*Standard Classification of Transported Goods*), E-cl@ss (una clasificación para el intercambio de información sobre materiales y servicios entre los proveedores y sus clientes empleado por varias empresas alemanas) y RosettaNet (creada por un consorcio de compañías de componentes electrónicos, informática, fabricación de semiconductores y empresas de soluciones informáticas). La diversidad de clasificaciones indica que no existe consenso entre los proveedores de servicios de comercio electrónico, lo que dificulta la interoperabilidad entre distintas aplicaciones.

En el ámbito de la **medicina** son también varias las ontologías construidas con el fin de reutilizar y compartir datos sobre los pacientes mediante sistemas médicos de información. Algunos ejemplos de ontologías médicas son GALEN, UMLS y ON9. GALEN es una terminología clínica representada en lenguaje formal (GRAIL). UMLS (*Unified Medical Language System*) es una amplia base de datos diseñada para integrar los términos biomédicos recogidos en diversas fuentes. Por último, ON9 es un conjunto de ontologías médicas que incluye otros sistemas como el UMLS.

En el ámbito de la **ingeniería**, las ontologías diseñadas contienen modelos matemáticos que los ingenieros emplean para analizar el comportamiento de los sistemas físicos. Permieren compartir y reutilizar modelos de ingeniería entre instrumentos y usuarios. Algunos ejemplos son EngMath y Physis.

En el ámbito de la **empresa**, las ontologías se crean para definir y organizar el conocimiento relevante sobre actividades, procesos, organizaciones, estrategias, marketing, etc. Algunos ejemplos son la Enterprise Ontology y TOVE (Toronto Virtual Enterprise).

Las ontologías creadas en el ámbito de la **química** modelan la composición, estructura y propiedades de sustancias, procesos y fenómenos. Pueden ser utilizadas en campos como la educación, las ciencias medioambientales o los descubrimientos científicos. Algunos ejemplos de ontologías en el dominio de la química son las ontologías *Chemicals*, *Ions* y la *Environmental Pollutants Ontology*.

Ontologías de **gestión del conocimiento** (*knowledge management*): el objetivo de estas ontologías es fomentar el crecimiento, la comunicación y la conservación del conocimiento. Algunos ejemplos de este tipo de ontologías son las ontologías (KA)².

Valoración

En esta sección hemos presentado dos tipologías de las relaciones, una según la riqueza de la estructura interna, que distingue entre ontologías ligeras y pesadas, y otra en función de la temática de la conceptualización.

Como indicábamos al final del apartado anterior, nos interesa considerar las ontologías ligeras en las que la estructuración del conocimiento se realiza por medio de clases, atributos y relaciones, si bien no descartamos la posibilidad de incorporar a nuestra base de conocimiento sobre la cerámica mecanismos que permitan el razonamiento y la inferencia.

Por la temática de la conceptualización, la base de conocimiento desarrollada en el marco de este trabajo será una ontología del dominio de la industria cerámica. En nuestro repaso de las ontologías de nivel superior hemos podido comprobar cómo se estructuran los primeros niveles de las ontologías. Esto nos permitirá elegir una de estas ontologías de nivel superior en la que insertar nuestra ontología de dominio. Como veremos, en nuestro trabajo empírico nos hemos decantado por elaborar nuestra ontología del dominio de la industria cerámica tomando como punto de partida los niveles superiores de la ontología Mikrokosmos, que, siendo una ontología clasificada como de tipo lingüístico en la tipología presentada, se ajusta perfectamente a nuestra concepción de las clases de conceptos en el ámbito de la cerámica.

4.4 Herramientas para el desarrollo de ontologías

Como hemos indicado en el apartado 4.2.1 al hablar de las metodologías, la tarea de construir una ontología es muy compleja. Por este motivo, desde mediados de los noventa se han desarrollado herramientas que facilitan esta labor. Estas herramientas permiten llevar a cabo las principales actividades del proceso de desarrollo, como la conceptualización, la implementación, la comprobación de la consistencia y la documentación. Gómez Pérez *et al* (2004: 293) distinguen los siguientes grupos de herramientas :

- a) **Herramientas para el desarrollo de ontologías:** Permiten construir ontologías desde cero. Además de funciones de edición y navegación, estas herramientas ayudan en las fases de documentación, importación y exportación en diferentes formatos y lenguajes ontológicos y también en edición gráfica y gestión de bibliotecas de ontologías, entre otras. Algunos ejemplos de este tipo de herramientas son KAON, OntoEdit, *Protégé*, WebODE y WebOnto.
- b) **Herramientas para la evaluación de ontologías:** Permiten valorar el contenido de las ontologías construidas y las tecnologías asociadas a ellas.
- c) **Herramientas para la fusión y alineación de ontologías:** Permiten fusionar y alinear las distintas ontologías existentes para un mismo dominio. Algunos ejemplos de este tipo de herramientas son el *plug-in* PROMPT de *Protégé* o Chimaera en Ontolingua Server.
- d) **Herramientas de anotación basadas en ontologías:** Estas herramientas permiten insertar instancias de conceptos y relaciones en ontologías y mantener de forma (semi)automática el marcado de páginas web. Algunos ejemplos de este tipo de herramienta son COHSE y MnM.
- e) **Herramientas de interrogación y motores de inferencia:** Permiten consultar las ontologías de manera fácil y llevar a cabo inferencias a partir de la información contenida en ellas.

- f) **Herramientas de aprendizaje de las ontologías:** Permiten construir ontologías de forma semiautomática a partir de textos en lenguaje natural mediante técnicas de aprendizaje automático y análisis del lenguaje natural.

Las herramientas que más nos interesan en este trabajo son las que facilitan el desarrollo de ontologías y las *suites* que agrupan distintas herramientas. Podemos distinguir entre las que dependen del lenguaje ontológico y aquellas que son independientes del mismo. Las principales herramientas dependientes del lenguaje ontológico son las siguientes:

- a) **Ontolingua Server**, cuyo objetivo es desarrollar ontologías en un entorno colaborativo y proporcionar un repositorio de ontologías en lenguaje Ontolingua.
- b) **OntoSaurus**, cuyo objetivo es el desarrollo de ontologías LOOM por medio de un navegador web.
- c) **WebOnto**, desarrollado para consultar y desarrollar de manera colaborativa ontologías OCML.
- d) **OilEd**, que es un editor de ontologías desarrollado en principio para el lenguaje OIL, que posteriormente se ha adaptado a los lenguajes DAML+OIL y OWL.

La desventaja que presentan este tipo de herramientas es que su uso es totalmente dependiente del lenguaje subyacente, que debe conocerse en profundidad. Por ello, para un desarrollador de ontologías no experto en lenguajes, es recomendable recurrir a herramientas que poseen una estructura abierta que permita su ampliación y la integración con otras aplicaciones de forma independiente con respecto al lenguaje empleado. Entre ellas destacamos WebODE, OntoEdit, KAON y *Protégé*.

La mayoría de estas *suites* ha sido desarrollada por grupos de investigación universitarios, y esta es la razón por la que todas ellas excepto OntoEdit Professional se distribuyen de forma gratuita. Algunas, como OntoEdit y WebODE, tienen la ventaja añadida de que permiten el desarrollo de ontologías siguiendo metodologías concretas (On-To-Knowledge y METHONTOLOGY, respectivamente).

En cuanto a la arquitectura del software, algunas de ellas poseen una estructura cliente/servidor que dificulta su ampliación. Es el caso de Ontolingua,

OntoSaurus y WebOnto. Las herramientas más modernas tienen una estructura en tres niveles, distinguiendo entre los *backends* encargados del almacenamiento de las ontologías, los módulos lógicos que se encargan de la lógica de la aplicación y los *frontends*, encargados de las interfaces usuario de las aplicaciones. Este es el caso de OilEd, *Protégé* y OntoEdit.

En cuanto al modelo de conocimiento, la mayoría de las herramientas combinan los marcos con la lógica de primer orden, pero difieren en los componentes y la cantidad de información sobre ellos que pueden representar.

A continuación vamos a describir brevemente las cuatro herramientas principales para el desarrollo de ontologías independientes del lenguaje ontológico empleado (WebODE, OntoEdit, KAON y *Protégé*). Siguiendo a Gómez Pérez *et al* (2004), para cada una de ellas especificaremos los siguientes datos:

- a) Breve descripción de la herramienta y sus desarrolladores
- b) Arquitectura de la herramienta
- c) Modelo de conocimiento
- d) Edición de ontologías
- e) Interoperabilidad con otras aplicaciones

4.4.1 WebODE

WebODE es una herramienta desarrollada por el grupo de investigación sobre ontologías de la Universidad Politécnica de Madrid. Permite construir ontologías siguiendo la metodología METHONTOLOGY.

Arquitectura

WebODE se concibe como una plataforma de trabajo (*workbench*) extensible e integrada que permite llevar a cabo la mayoría de las actividades implicadas en el proceso de desarrollo de una ontología (conceptualización, razonamiento, intercambio) y además facilita la interoperabilidad con otros sistemas informáticos. Las ontologías en WebODE se almacenan en una base de datos relacional, que puede albergar ontologías de gran tamaño.

Modelo de conocimiento

El modelo de conocimiento con el que se conceptualizan las ontologías en WebODE es muy expresivo, y está basado en las representaciones intermedias presentadas en METHONTOLOGY que hemos visto en el apartado 4.2.1.1 (diccionario conceptual, tablas de relaciones e instancias, etc.).

Los componentes básicos incluidos son: conceptos y sus atributos locales (que pueden ser atributos de instancia y de clase), grupos de conceptos, que representan conjuntos de conceptos disjuntos, taxonomías de conceptos y particiones, relaciones binarias *ad hoc* y sus propiedades (simetría y transitividad), constantes, axiomas formales, reglas e instancias de conceptos y relaciones. Además permite añadir información bibliográfica, sinónimos y abreviaturas de los componentes.

Editor de ontologías

El editor de ontologías de WebODE es una aplicación web que integra diversos servicios: edición de la ontología, navegación, documentación, fusión y razonamiento, organizados en tres interfaces de usuario: un editor en formato HTML, llamado *OntoDesigner*, en la cual se editan todos los términos de la ontología excepto los axiomas y reglas; una interfaz de usuario gráfica, en la que se editan las taxonomías de conceptos y las relaciones gráficamente; y una tercera interfaz, denominada *WebODE Axiom Builder* (WAB), que permite editar axiomas formales y reglas.

Otros servicios para la creación de ontologías integrados en WebODE son el servicio de documentación, que permite generar ontologías en diferentes formatos para utilizarlos como documentación; el servicio de fusión supervisada de conceptos, atributos y relaciones de dos ontologías construidas para el mismo dominio (ODEMerge); un motor de inferencia que consiste en una implementación en Prolog de un subconjunto de las primitivas del protocolo OKBC; y, por último un servicio para la evaluación de taxonomías de conceptos (ODEClean), basado en el método OntoClean. Además de estos servicios, WebODE ofrece otros para la evaluación de ontologías en RDF(S), DAML+OIL y OWL, que detectan errores en las ontologías implementadas en estos lenguajes. Por último, WebODE permite crear grupos de usuarios para editar ontologías de manera colaborativa.

Interoperabilidad

Las ontologías creadas con WebODE pueden reutilizarse en otras aplicaciones basadas en ontologías, tanto de forma remota como desde un ordenador concreto. Además, WebODE permite exportar las ontologías en XML y

en otros lenguajes ontológicos como RDF(S), OIL, DAML+OIL, OWL, XCARIN y FLogic.

Las ontologías creadas en WebODE pueden convertirse en ontologías en *Protégé*, de forma que se pueden utilizar las capacidades adicionales ofrecidas por esta herramienta.

Por último, las ontologías creadas en WebODE pueden transformarse al lenguaje Java, lo que permite desarrollar otras aplicaciones en dicho lenguaje.

4.4.2 **OntoEdit**

OntoEdit es un entorno para el desarrollo de ontologías creado por el Instituto AIFB de la Universidad de Karlsruhe, actualmente distribuido por la empresa Ontoprise. Existen dos versiones, una gratuita y otra profesional, que incluye algunas funciones adicionales, como un motor de inferencia, instrumentos de consulta gráfica, módulos de importación y exportación y editores de reglas gráficos, entre otras.

Arquitectura

OntoEdit es una aplicación Java independiente que puede instalarse y utilizarse en un ordenador local. El editor de ontologías posee los módulos centrales de la aplicación, que se amplían mediante *plug-ins*. Las ontologías en OntoEdit se almacenan como archivos XML, aunque la versión profesional permite su almacenamiento como bases de datos relacionales.

Los *plug-ins* disponibles para ampliar OntoEdit varían en función de la versión (gratuita o profesional). Pueden distinguirse dos categorías: los filtros de importación y exportación de ontologías en diferentes lenguajes (RDF(S), OXML, DAML+OIL y FLogic) y los *plug-ins* generales, que permiten llevar a cabo la gestión de léxicos de dominio, la edición de instancias y la expresión de conceptos disjuntos. La versión profesional añade a estos editores gráficos de reglas, el motor de inferencia OntoBroker, herramientas de interrogación gráfica y comparación de ontologías.

Modelo de conocimiento

El modelo de conocimiento de OntoEdit se basa en marcos. Puede representar conceptos y sus atributos, relaciones binarias entre conceptos, axiomas formales e instancias. OntoEdit distingue entre conceptos abstractos (que no

pueden tener instancias) y concretos (que pueden tener instancias). Los conceptos se clasifican en taxonomías, que pueden aprovechar la herencia múltiple. Además, dos conceptos pueden declararse disjuntos de manera explícita.

Los atributos y las relaciones binarias se definen del mismo modo y se aplican al concepto para el que se definen de forma local. Por tanto, pueden crearse atributos y relaciones diferentes con el mismo nombre siempre que se asignen a conceptos distintos. Los tipos de valor que pueden tener los atributos son *String*, *Integer*, *Double* y *Boolean*, pero pueden ampliarse fácilmente. En cuanto a las relaciones, el rango puede ser cualquier concepto definido en la ontología.

En lo que concierne a los axiomas, es posible elegir el lenguaje formal en el que se representan, aunque por defecto es FLogic.

En cuanto a las posibilidades de importación de otras ontologías, OntoEdit no permite mostrar o acceder a los términos importados, sino que es necesario referirse a sus URIs.

Editor de ontologías

El editor de ontologías de OntoEdit permite navegar por la taxonomía de conceptos y editar las relaciones e instancias por medio de una estructura en árbol. Asimismo es posible utilizar las funciones copiar y pegar para facilitar la labor de edición de las taxonomías.

Una ventaja añadida de OntoEdit es que los términos de la ontología pueden ser descritos en varias lenguas.

Permite incluir referencias bibliográficas a toda la ontología e información adicional sobre los desarrolladores de la ontología o el dominio.

OntoEdit posee una interfaz de visualización mediante la que es posible navegar y editar la ontología de manera gráfica.

En la versión gratuita de OntoEdit no se incluye ningún razonador, mientras que la versión profesional posee dos *plug-ins* que integran el motor de inferencia Ontobroker y un editor de reglas gráfico.

Por último, la versión profesional de OntoEdit permite bloquear fragmentos de la taxonomía de conceptos.

Interoperabilidad

Existen dos modos de utilizar OntoEdit en aplicaciones basadas en ontologías. Por un lado, pueden exportarse las ontologías a distintos lenguajes

ontológicos y utilizarlas en las herramientas y sistemas de representación del conocimiento correspondientes a dichos lenguajes. OntoEdit permite importar ontologías a partir de estructuras de directorio y a partir de archivos en Excel. Por otro lado, OntoEdit es una aplicación Java, por lo que puede accederse a una ontología utilizando el API⁴⁴ Java.

4.4.3 KAON

La *suite* de herramientas KAON ha sido desarrollada por dos institutos de la Universidad de Karlsruhe y permite el acceso de forma gratuita.

Arquitectura

KAON posee una arquitectura flexible y ampliable. Se organiza en tres capas: *backends* para el almacenamiento de las ontologías, aplicaciones intermedias para los servicios ofrecidos por la herramienta y aplicaciones para el usuario o *frontends*. Las ontologías KAON pueden almacenarse de dos modos diferentes: como archivos RDF(S) y como modelos en bases de datos relacionales. Los principales servicios ofrecidos por KAON son el acceso a ontologías KAON, un servicio de control de la evolución de la ontología, un servicio para realizar un seguimiento de los cambios operados en la ontología y un servicio de detección de conflictos, entre otros.

Las aplicaciones para el usuario son el editor de ontologías OI-modeler y el portal KAON, que permite crear portales web basados en ontologías.

Modelo de conocimiento

El modelo de conocimiento que subyace a KAON está basado en una extensión del lenguaje RDF(S). En KAON podemos modelar conceptos, propiedades (para expresar relaciones y atributos) e instancias de conceptos y propiedades. El concepto superior de la taxonomía de conceptos es el concepto *Root*. Tanto los conceptos como las relaciones se organizan en taxonomías mediante la relación *subclass of* y *supproperty of*, respectivamente. Las extensiones que incorpora KAON con respecto a RDF(S) se refieren sobre todo a las propiedades. KAON permite definir su cardinalidad, distinguir entre atributos y relaciones dependiendo del rango (concepto o tipo de datos), definir las

⁴⁴ Una API (*Application Program Interface*) es un grupo de rutinas de un sistema operativo o una aplicación que definen cómo invocar cualquier servicio desde un programa.

relaciones como simétricas o transitivas y definir la relación inversa de una relación dada.

En el modelo KAON se pueden adjuntar etiquetas, documentación, sinónimos y raíces de palabras en lenguaje natural a los conceptos, propiedades e instancias.

En KAON no pueden expresarse axiomas formales, ni taxonomías de conceptos disjuntos o exhaustivas.

Las ontologías KAON pueden incluir términos de otras ontologías desarrolladas en KAON.

Editor de ontologías

OI-Modeler es el editor de ontologías de KAON. Posee una interfaz de usuario gráfica, en la que los conceptos se representan en forma de cuadros y las relaciones taxonómicas (*subclass of*) se representan por medio de líneas que enlazan los conceptos. Las propiedades, por su parte, se representan como cuadrados con las esquinas redondeadas.

Junto a la interfaz de usuario gráfica, se puede utilizar una interfaz de formulario para navegar y editar la ontología. A través de ella pueden editarse los super- y subconceptos, sus propiedades, sus etiquetas, la documentación, los sinónimos y las raíces en distintas lenguas y sus instancias.

Como hemos indicado anteriormente, el editor de ontologías KAON permite establecer una estrategia para controlar la evolución de la ontología. Por ejemplo, podemos especificar que al eliminar un concepto se eliminen también sus subclases o que se consideren subclases del concepto *Root*, o bien conectarlas a las superclases de la clase eliminada. Además, KAON contempla la edición colaborativa de ontologías.

Interoperabilidad

El acceso a las ontologías KAON puede realizarse mediante el API de Java. También puede accederse a la API de KAON utilizando servicios web por medio del KAON Web Service (KAON WS). Las ontologías pueden almacenarse como archivos o bases de datos, y pueden exportarse al lenguaje KAON. Por último, KAON permite importar ontologías generadas en lenguaje KAON, en RDF(S) y en el código RDF(S) generado por *Protégé*.

4.4.4 *Protégé*

Protégé es un conjunto de programas para la edición de ontologías desarrollado por la grupo SMI (Stanford Medical Informatics) de la Universidad de Stanford. Tiene más de 50.000 usuarios registrados y puede descargarse gratuitamente a través de su página web⁴⁵. La primera versión se desarrolló en 1987 con el objetivo de simplificar el proceso de adquisición del conocimiento en sistemas expertos. En la actualidad, *Protégé* está orientado hacia el desarrollo de ontologías y bases de conocimiento.

La versión más actual disponible de *Protégé* es la 3.2.1, aunque la versión 3.3 puede utilizarse en pruebas (junio de 2007). La plataforma *Protégé* proporciona dos modos diferentes de modelar ontologías: el editor *Protégé-frames* y el editor *Protégé-OWL*. El primero permite crear y almacenar ontologías basadas en marcos utilizando un modelo de conocimiento compatible con el protocolo OKBC. *Protégé-OWL* es una extensión que permite desarrollar ontologías utilizando el lenguaje ontológico estándar OWL, basado en lógica descriptiva. Dadas las diferencias notables existentes entre estos dos editores en cuanto al modelo de conocimiento empleado, hemos decidido tratarlas por separado. No obstante, su arquitectura es similar, por lo que la describiremos en primer lugar, para pasar a describir cada uno de los editores de forma independiente.

Arquitectura

Protégé es una aplicación basada en Java que se instala de forma independiente en un ordenador local. La aplicación central de *Protégé* es el editor de ontologías basado en marcos, si bien, gracias a su arquitectura abierta, puede ampliarse por medio de *plug-ins* que mejoran las posibilidades de importación o exportación o permiten configurar el editor de ontologías. *Protégé-OWL* se presenta como una extensión de *Protégé* marcos que se integra dentro de la *suite* como un *plug-in*.

Los distintos tipos de *plug-ins* que pueden agregarse a *Protégé* son *tab plug-ins*, *slot widgets* y *backends*. Los primeros añaden una pestaña al editor de ontologías desde la que se puede acceder a funciones adicionales del editor, como por ejemplo la visualización gráfica, la fusión y el tratamiento de versiones de las

⁴⁵ Datos extraídos de <http://protege.stanford.edu/> en junio de 2006.

ontologías, la gestión de recursos de conocimiento en línea, el acceso a ontologías OKBC, la creación y ejecución de restricciones y los motores de inferencia.

Los *slot widgets* permiten mostrar y editar los valores de los *slots* que difieren de los definidos originalmente en *Protégé*, como imágenes, vídeos y audio, gestión de fechas, medidas o cambio de valores de los *slots*.

Por último, los *backends* permiten exportar e importar ontologías en diferentes formatos (RDFSchemata, XML Schema, etc.).

Los *plug-ins* más habituales se encuentran integrados en la aplicación en Java, y todos ellos se encuentran accesibles para su descarga en la biblioteca de *plug-ins* y aplicaciones ubicada en la página web de *Protégé*.

4.4.4.1 *Protégé-frames*

Modelo de conocimiento

El modelo de conocimiento de *Protégé-frames* se basa en marcos y en lógica de primer orden. Los principales componentes para el modelado de ontologías son clases, *slots*⁴⁶, facetas e instancias o individuos, compatibles con el protocolo OKBC (*Open Knowledge Based Connectivity*).

Habida cuenta de que el editor de ontologías *Protégé-frames* es el elegido para el desarrollo de nuestra ontología cerámica, será descrito con mayor detenimiento en el capítulo 5. De momento no entraremos a definir cada uno de los componentes básicos. Remitimos al apartado 5.3.3 para completar la información sobre el funcionamiento del editor de ontologías.

Protégé permite importar otras ontologías desarrolladas con *Protégé* de dos modos distintos: incluyendo los conceptos como términos externos que no pueden ser editados, o bien fusionando ontologías, es decir, copiando la definición de la ontología que se quiere importar y redefiniéndola en *Protégé*.

Editor de ontologías

El editor de ontologías de *Protégé* se configura en torno a las pestañas (*Tabs*), cada una de las cuales permite gestionar un componente de la ontología. En la versión básica (sin *plug-ins* añadidos), *Protégé* presenta las pestañas

⁴⁶ La traducción literal de *slot* sería “ranura o muesca”, que no nos parece muy adecuada. Hemos mantenido la denominación en inglés a falta de un equivalente en español que sea lo suficientemente ilustrativo. Consideramos que *slot* es una denominación muy económica, y además es la que emplean generalmente los expertos en ingeniería del conocimiento consultados.

Classes, *Slots*, *Forms*, *Instances* y *Queries*, que se describen brevemente a continuación⁴⁷:

- a) *Classes*: En esta pestaña podemos gestionar la taxonomía de clases, que se presenta como una estructura arbórea. La clase que aparece siempre en la parte superior de la taxonomía es el concepto *:THING*, y de ella dependen todas las clases de la ontología. Las clases representan los conceptos en la ontología, determinados por sus propiedades (*slots*). En la jerarquía de clases en *Protégé* funcionan los mecanismos de herencia múltiple.
- b) *Slots*: Esta pestaña permite gestionar las propiedades asignadas a las clases, que pueden tener distintos tipos de valor (numérico, booleano, clases, instancias, cadenas de caracteres o símbolos). Las propiedades también pueden estructurarse de forma jerárquica, y permiten especificar atributos y relaciones *ad hoc* entre las clases de la ontología.
- c) *Forms*: Esta pestaña permite personalizar las vistas de pantalla empleadas para la introducción de instancias en una sola ventana. Los formularios definidos en esta pestaña se emplearán para introducir los individuos en la base de conocimiento.
- d) *Instances*: Mediante la selección de esta pestaña podemos introducir las instancias y sus valores específicos en la ontología. En las instancias se especifican los valores que se asignan a las propiedades de una clase concreta.
- e) *Queries*: La pestaña *Queries* se emplea para crear, ejecutar y guardar consultas que permiten seleccionar determinadas instancias de la ontología en función de los valores de uno o varios *slots*. Las consultas no forman parte integral de la base de conocimiento, pero posibilitan la identificación de los individuos contenidos en la ontología.

⁴⁷ De nuevo remitimos al apartado 5.3.3 para una descripción más detallada del uso de la herramienta.

Además de estas cinco pestañas básicas, podemos añadir otras que amplían la arquitectura de *Protégé* para llevar a cabo restricciones, representaciones gráficas o inferencias, entre otras utilidades.

El editor de ontologías permite generar distintos tipos de documentación sobre la ontología, tales como documentos en HTML o estadísticas sobre la ontología.

Existe un gran número de ontologías creadas con *Protégé* que pueden importarse para desarrollar ontologías de dominio a partir de ellas (como la UNSPSC, la ontología SUMO, etc.). Además, con *Protégé* es posible importar ontologías en diversos formatos a partir de bibliotecas de ontologías como DAML y Ontolingua Server, entre otras.

Por último, con *Protégé* también es posible crear entornos de trabajo colaborativo, que facilitan la edición de la ontología por varios usuarios a la vez.

En la Fig. 19 se presenta una captura de pantalla que ilustra cómo se presenta la información en *Protégé-frames*:

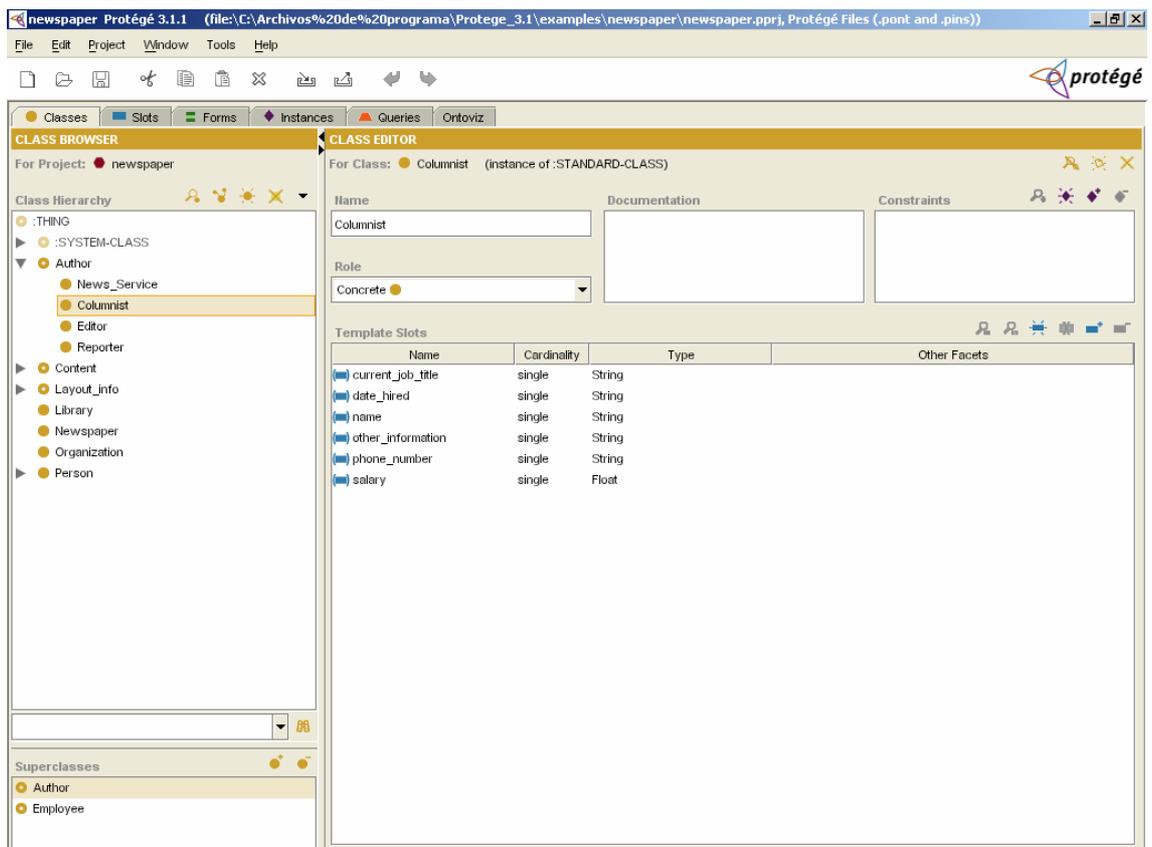


Fig. 19. Estructura del concepto *Columnist* en la ontología de muestra *Newspaper*.

Interoperabilidad

Las ontologías creadas por *Protégé* pueden funcionar conjuntamente con un gran número de aplicaciones basadas en ontologías. Se pueden exportar e importar ontologías en *Protégé* con los *backends* RDF(S), XML, XMLSchema y XMI. Una vez generado y guardado el archivo correspondiente, puede reutilizarse en cualquier aplicación que emplee estos lenguajes.

4.4.4.2 *Protégé-OWL*

Protégé-OWL es una extensión de la suite de programas *Protégé* que permite desarrollar ontologías en el lenguaje ontológico de marcado OWL. Está concebido como un *plug-in* de *Protégé*, que puede cargar y guardar ontologías en OWL y en RDF, así como editar y visualizar las clases, propiedades y reglas.

Dado que el lenguaje OWL emplea la lógica descriptiva como paradigma de representación del conocimiento, las características lógicas de las clases se definen como expresiones lógicas en lenguaje OWL. Asimismo, en las ontologías desarrolladas con *Protégé OWL* pueden utilizarse razonadores lógicos con el fin de crear clasificaciones automáticamente y llevar a cabo otras actividades de razonamiento. Por último, este editor permite editar individuos en OWL para el marcado se páginas para la web semántica.

Protégé OWL posee una estructura en tres capas: OWL-Lite, OWL-DL y OWL-Full, que difieren en cuanto a su expresividad, siendo OWL-Lite el menos expresivo y OWL-Full el módulo más expresivo. En esta descripción nos centramos en OWL-DL, puesto que es más expresivo que OWL-Lite y permite el razonamiento automático.

Modelo de conocimiento

El modelo de conocimiento de OWL-DL es el aspecto que más varía con respecto a *Protégé-frames*, puesto que en lugar de basarse en marcos y lógica de primer orden se basa en lógica descriptiva. Por ello, los componentes de las ontologías creadas con el *plug-in* OWL en *Protégé*, aunque aparentemente son similares, difieren en cuanto a la terminología empleada. A continuación describimos brevemente cuáles son los principales componentes de las ontologías creadas con *Protégé OWL*:

a) **Clases OWL**: Se conciben como conjuntos de individuos que se describen mediante descripciones matemáticas formales para definir de manera

precisa los requisitos necesarios y suficientes (condiciones) para que un individuo pertenezca a una clase. Las clases se organizan en taxonomías, de manera similar a como lo hacen en *Protégé-frames*. Una de las características distintivas de OWL-DL es que permite la generación automática de relaciones taxonómicas mediante un razonador.

Las clases pueden ser primitivas o definidas, es decir, se diferencia entre las clases para las que sólo se especifican condiciones necesarias, y aquellas para las que se definen condiciones necesarias y suficientes.

b) *Propiedades*: Las propiedades expresan relaciones binarias establecidas entre individuos. Las propiedades pueden tener propiedades inversas y pueden definirse mediante sus características matemáticas. Así, una propiedad puede ser funcional (cuando sólo tiene un valor), transitiva o simétrica. Las propiedades en OWL pueden ser de dos tipos: propiedades de objeto (*object properties*) y propiedades de tipos de datos (*datatype properties*). Las primeras establecen relaciones entre individuos, mientras que las segundas asignan otros tipos de valores a los individuos (números, cadenas de caracteres, etc.). *Protégé-OWL* también establece jerarquías de propiedades en las que las subpropiedades especializan la propiedad de la que dependen en la taxonomía.

Al igual que los *slots* definidos en *Protégé-Frames*, las propiedades tienen un dominio y un rango, es decir, relacionan un grupo de individuos (el dominio) con otro grupo de individuos o valores (el rango).

Uno de los rasgos distintivos de OWL es que las propiedades se emplean para crear restricciones que limitan los individuos que pueden pertenecer a una clase. Las restricciones expresan las condiciones necesarias y/o suficientes para que un individuo pertenezca a una clase, y se expresan mediante fórmulas en lógica descriptiva. Las restricciones en OWL pueden ser restricciones de la cantidad (que se dividen en existenciales y universales), restricciones de la cardinalidad y restricciones de los valores⁴⁸.

c) *Individuos*: Los individuos representan objetos en el dominio. Una diferencia fundamental entre *Protégé-frames* y *Protégé-OWL* es que en este último no se emplea la *Unique Name Assumption* (UNA). Esto quiere decir que

⁴⁸ La descripción exhaustiva de las restricciones excede los límites de este trabajo. Para una descripción más detallada remitimos a Horridge *et al* (2004).

dos nombres distintos pueden referirse al mismo individuo. Es necesario expresar de forma explícita si dos individuos son iguales o difieren entre sí.

Editor de ontologías

La presentación de la interfaz de usuario de *Protégé-OWL* es bastante similar a la de la versión con marcos. Sin embargo, el tipo de información que se define para cada clase difiere bastante entre las dos versiones de *Protégé*, y por ello también son distintas las pestañas que se presentan. Son las siguientes:

- a) **OWL Classes:** Permite establecer la jerarquía de clases y asignar propiedades a cada clase mediante expresiones en lógica descriptiva.
- b) **Properties:** Sirve para gestionar las propiedades de datos y las relaciones *ad hoc* que se asignan a una clase determinada.
- c) **Forms:** Al igual que en la versión basada en marcos, esta pestaña permite crear formularios para la introducción estructurada de los valores de los individuos.
- d) **Individuals:** Esta pestaña tiene la misma función definida para la pestaña *Instances* en la versión marcos, es decir, describir a los individuos asignando valores a las propiedades de las clases.
- e) **Metadata:** Esta pestaña es propia de los lenguajes ontológicos de marcado, y permite gestionar las etiquetas para anotar la web semántica.

Nótese que el editor *Protégé-OWL* no incluye por defecto una pestaña para la generación y ejecución de consultas, pero puede añadirse al editor mediante la configuración del *plug-in Queries Tab*.

En la Fig. 20 ofrecemos una captura de pantalla de la ontología *pizza.owl*, que se proporciona como ejemplo de ontología desarrollada utilizando *Protégé-OWL*:

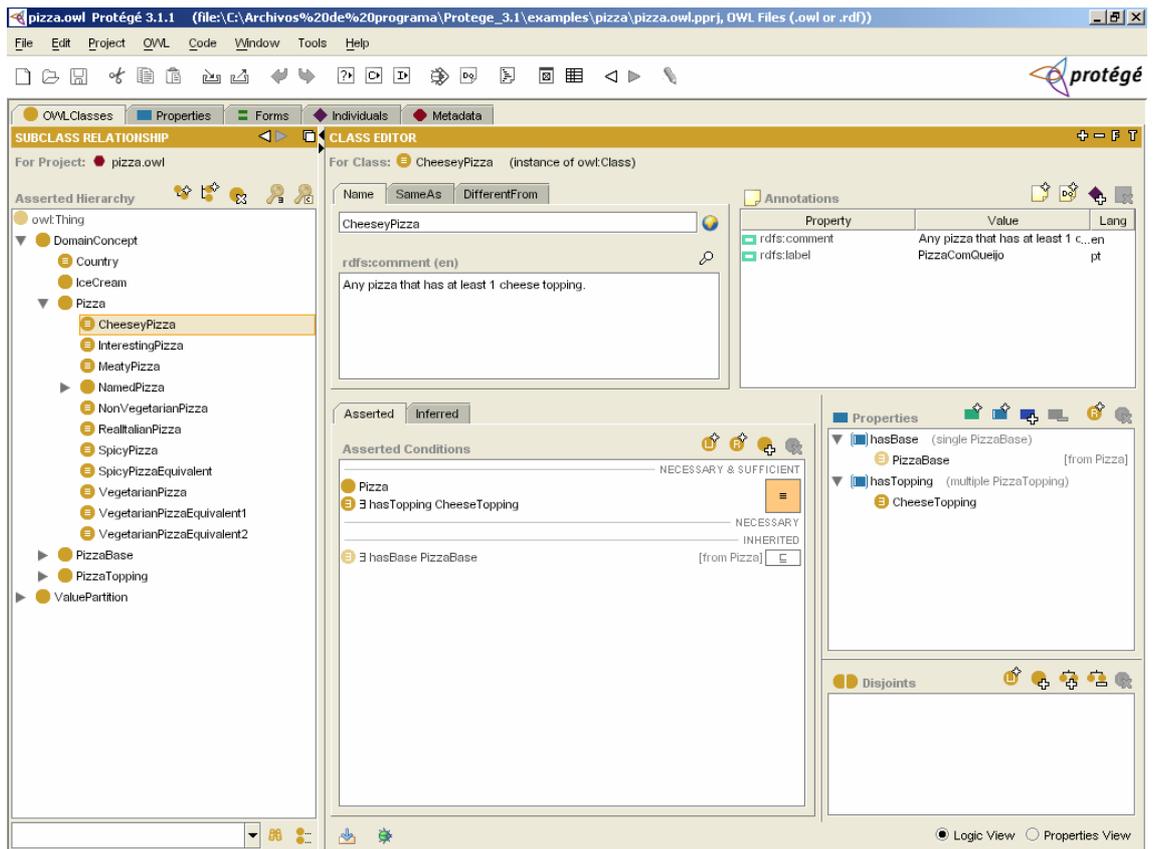


Fig. 20. Captura de pantalla de la clase *CheesyPizza* en la ontología *pizza.owl*.

En la parte izquierda de la pantalla podemos ver la jerarquía de clases, y en la parte derecha las condiciones y propiedades que definen la clase seleccionada (*CheesyPizza*). Además, vemos que en OWL es necesario especificar si la clase es igual o diferente de otras clases definidas en la ontología, y tenemos la posibilidad de anotar las clases, las propiedades y los individuos con distintos tipos de información o con metadatos, que permiten, por ejemplo, añadir etiquetas en otras lenguas o ampliar la información editorial sobre la ontología.

La Fig. 21 muestra la jerarquía de propiedades en el editor de *Protégé* OWL. En ella podemos ver que las propiedades en OWL se definen por su dominio y rango, y que pueden poseer también propiedades de anotación. Vemos como resulta muy sencillo definir las características matemáticas de una propiedad (funcional, simétrica y transitiva), así como indicar si una propiedad posee una inversa.

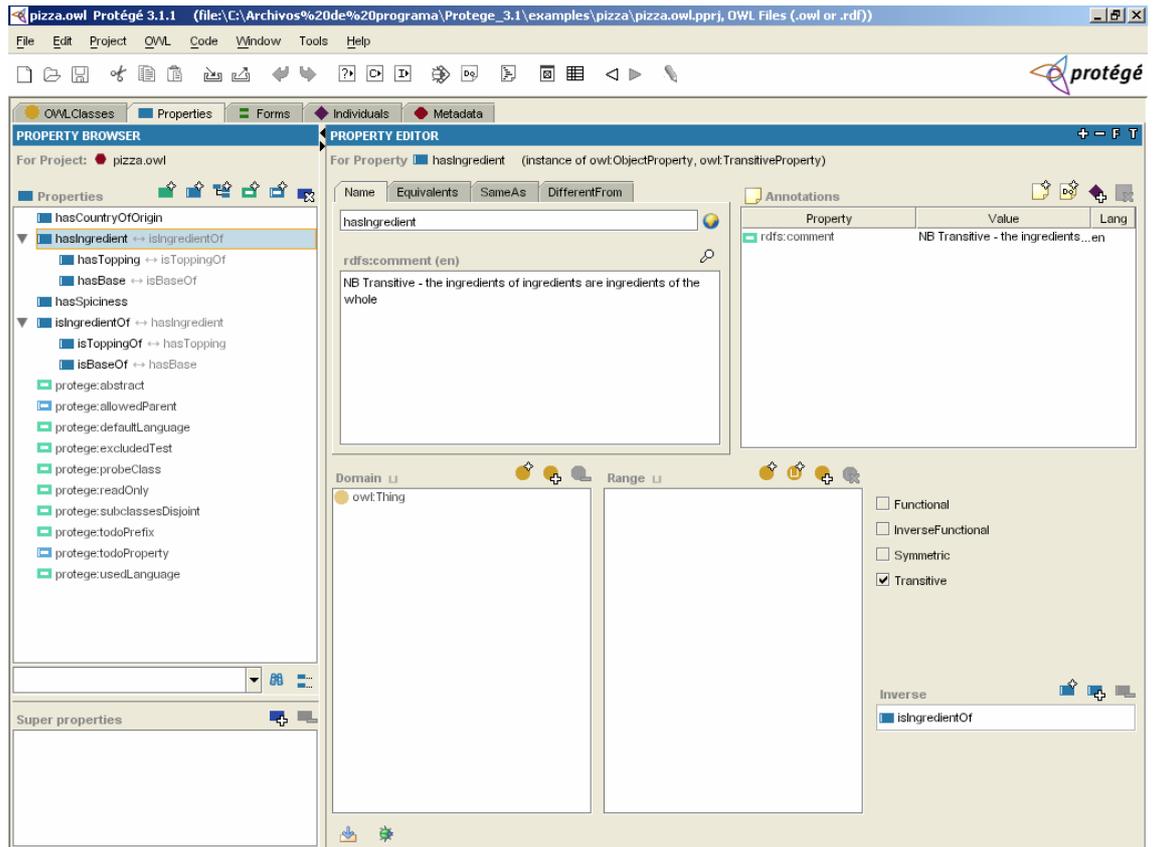


Fig. 21. Jerarquía de propiedades en el editor *Protégé*-OWL.

Interoperabilidad

En cuanto a la interoperabilidad de *Protégé* OWL, el rasgo más significativo es la posibilidad de emplear razonadores externos para comprobar la consistencia de la jerarquía y controlar la herencia múltiple mediante inferencias, así como para la generación automática de taxonomías de clases. El razonador más empleado es RACER, aunque pueden utilizarse otros que permitan el razonamiento en lógica descriptiva. Al igual que *Protégé-frames*, las ontologías en OWL pueden exportarse en distintos formatos, y, una vez exportadas, las ontologías pueden emplearse en cualquier aplicación basada en ontologías que soporte el lenguaje en el que se guarde.

Valoración

Una vez revisadas las principales herramientas para el desarrollo de ontologías, hemos de decantarnos por una de ellas para su empleo en nuestro trabajo empírico. Nuestra decisión está condicionada, en parte, por los requisitos que debe tener un editor de ontologías útil para la representación del conocimiento terminológico.

Feliu *et al* (2002: 42-43) identifican una serie de requisitos que, en su opinión, debe cumplir una herramienta para la gestión de ontologías aplicada al desarrollo de su ontología para el proyecto GENOMA-KB (ver apartado 3.3.4). En opinión de estos autores, un editor de ontologías debe permitir las siguientes acciones:

- a) Ampliar la jerarquía
- b) Asignar atributos predefinidos
- c) Añadir relaciones preestablecidas y completamente organizadas
- d) Bloquear la repetición de conceptos y/o relaciones y evitar la circularidad dentro del sistema.
- e) Tener en cuenta el sistema de herencia (monotónica/no monotónica; control de la herencia múltiple, entre otras).
- f) Tener un mecanismo para conectar la ontología con los posibles diccionarios existentes.
- g) Exportar el contenido de la ontología a diferentes formatos
- h) Permitir su fácil manejo

En definitiva, es necesario mantener la coherencia del sistema. La adición de nuevos atributos y relaciones debe estar controlada y protegida. Todos los atributos deben tener un espacio particular en el que se predefinan sus valores o se limite su rango. En cuanto a las relaciones, sus propiedades han de estar claramente definidas y sólo deben asignarse a los conceptos una vez que se haya superado un test de validación.

Dado que los objetivos perseguidos por el proyecto GENOMA-KB y los nuestros son bastante similares, el editor que seleccionemos debe cumplir estos requisitos. Los editores de ontologías presentados en este apartado se ajustan a nuestras necesidades, pero hemos decidido decantarnos por *Protégé-frames* porque, en nuestra opinión, tiene las siguientes ventajas:

- a) Se trata de una herramienta de distribución gratuita, lo que facilita su acceso para los desarrolladores de la ontología.
- b) El número de usuarios de *Protégé* es muy elevado, por lo que se han desarrollado un gran número de ontologías de dominio empleando esta herramienta. Las aportaciones de los usuarios, recogidas en la documentación y en la comunidad *Protégé*, servirán

de ayuda para resolver las dificultades que vayan surgiendo a medida que se desarrolla la base de conocimiento.

- c) *Protégé* permite el trabajo colaborativo. Si bien en este trabajo no vamos a explotar esta ventaja, sin duda será de gran ayuda cuando varios usuarios trabajen en la edición de la base de conocimiento.
- d) *Protégé-frames* es una herramienta independiente del lenguaje ontológico. Esto supone que el usuario del editor no necesita tener vastos conocimientos sobre lenguajes ontológicos. El desarrollo de ontologías empleando *Protégé* se realiza de una manera intuitiva, lo que no impide que puedan llevarse a cabo actividades complejas como la definición de restricciones o el uso de razonadores lógicos.
- e) *Protégé* se distribuye como una aplicación independiente que se instala en el ordenador. Esto es una ventaja frente a herramientas como WebOde, en las que las ontologías se desarrollan a través de una interfaz web.
- f) Su arquitectura abierta y modular determina que sus funcionalidades se amplíen con facilidad por medio de distintos *plug-ins*. Así, podremos emplear los mecanismos de representación gráfica y de definición de restricciones, entre otras.
- g) *Protégé* almacena las ontologías como bases de datos, lo que permite gestionar ontologías de gran tamaño.

A pesar de estas ventajas, *Protégé-frames* presenta algunos inconvenientes que será necesario superar en el desarrollo de nuestra base de conocimiento. Entre ellos destaca el hecho de que no es fácil formalizar las propiedades matemáticas de las relaciones *ad hoc* (transitividad, simetría, funcionalidad) mediante los *slots*. Sin embargo, como veremos en el capítulo 6, esta dificultad puede superarse definiendo nuevas propiedades para los *slots*. Otro inconveniente de *Protégé* es que, en principio, no permite reflejar variantes denominativas ni etiquetas en otros idiomas. Sería muy interesante para nuestro trabajo que pudiera hacer esto, puesto que trabajamos con varias lenguas. No obstante, esta dificultad también puede superarse mediante la definición de instancias para cada variante denominativa y para cada término equivalente en otras lenguas.

Podríamos habernos decantado también por la versión *Protégé-OWL*, que –como hemos visto– permite definir las relaciones por medio de sus propiedades

lógicas. Sin embargo, esta ventaja no compensa el hecho de que para definir las condiciones necesarias y suficientes de las clases es necesario manejar los conceptos de la lógica descriptiva con cierta soltura, así como las etiquetas RDF(S) y OWL.

En definitiva, todo esto justifica que la herramienta elegida para el desarrollo de nuestra base de conocimiento sobre los productos cerámicos acabados es *Protégé-frames*. A pesar de que en principio está pensada para la creación de ontologías tal y como se conciben en la ingeniería del conocimiento, creemos que se ajusta bastante bien a los requisitos necesarios para desarrollar una ontología terminológica con información sobre relaciones conceptuales.

4.5 Recapitulación y conclusiones

En este capítulo nos marcábamos como objetivo la presentación de los principios y herramientas de la ingeniería ontológica, con el fin de observar cómo se pueden reutilizar y adaptar sus herramientas para el proceso terminológico automatizado.

En la sección 4.1 hemos definido ontología como una conceptualización explícita y consensuada de los fenómenos del mundo, que puede ser reutilizada y compartida por distintos grupos de personas y aplicaciones informáticas. La revisión de las técnicas de modelado de ontologías nos permite concluir que, sea cual sea la técnica de modelado elegida, las ontologías pueden servir para dar cuenta de los conceptos, sus atributos y relaciones.

En la sección 4.2.1 hemos visto cuáles son los principios y métodos que condicionan y facilitan el desarrollo de ontologías. Hemos descrito en detalle la metodología METHONTOLOGY, ya que las tareas que propone se asemejan a las llevadas a cabo en el proceso de trabajo terminológico. En nuestro trabajo empírico retomaremos algunas de las tareas propuestas por esta metodología, con el fin de elaborar una tabla que recoja el catálogo de las relaciones conceptuales de nuestra ontología para su posterior formalización.

El apartado 4.3.2 se ha dedicado a la presentación somera de los distintos lenguajes ontológicos. Hemos visto cómo se puede distinguir entre lenguajes tradicionales y lenguajes de marcado. Todos ellos permiten representar conceptos y sus atributos y estructurarlos de manera jerárquica por medio de la relación

básica *SubclassOf*, que vertebra la ontología y, mediante mecanismos de herencia, facilita el desarrollo de ontologías de manera económica.

En la sección 4.3 hemos clasificado las ontologías atendiendo a dos criterios (riqueza de la estructura interna y temática de la conceptualización). Según estos criterios, la base de conocimiento que desarrollaremos como parte de nuestro trabajo empírico será una ontología de dominio ligera. Con el fin de facilitar su desarrollo e integración con otras ontologías, los niveles superiores de nuestra ontología se tomarán de la ontología lingüística Mikrokosmos, descrita en 3.4.2.

La sección 4.4 se ha dedicado a la descripción de las principales herramientas para el desarrollo de ontologías. En primer lugar hemos definido los distintos tipos de herramientas para el desarrollo de ontologías, para centrarnos posteriormente en la descripción de los editores de ontologías independientes de lenguaje ontológico empleado. Así, hemos descrito las herramientas WebODE, OntoEdit, KAON y *Protégé* en sus dos vertientes, la que emplea marcos y lógica de primer orden (*Protégé-frames*) y la que se basa en lógica descriptiva (*Protégé-OWL*).

La revisión de todas estas herramientas nos ha llevado a decantarnos por el editor de ontologías *Protégé*-marcos como editor para nuestra base de conocimiento sobre la cerámica en la parte empírica. Este editor cumple los requisitos básicos necesarios para desarrollar una ontología terminológica, y además posee otras ventajas relacionadas con su libre acceso y su facilidad de ampliación.

En el estudio empírico describiremos en detalle esta herramienta, y veremos cómo se ha utilizado para el desarrollo de nuestra ontología cerámica.

La conclusión general de este capítulo es que las herramientas desarrolladas en el marco de la ingeniería ontológica conforman una excelente plataforma de trabajo para el desarrollo de ontologías terminológicas. En el estudio empírico refrendaremos esta conclusión, y propondremos una serie de adaptaciones y mejoras que pueden hacer de los editores de ontologías uno más de los instrumentos empleados en el trabajo terminológico automatizado.

2ª PARTE: ESTUDIO EMPÍRICO

El principal objetivo de esta tesis doctoral es aplicar un editor de ontologías para la formalización de las relaciones conceptuales en terminología. En los fundamentos teóricos se han revisado los estudios sobre relaciones conceptuales llevados a cabo en la Terminología y la Lingüística (Capítulo 2), así como las propuestas de formalización de las mismas (Capítulo 3). El capítulo 4 se ha dedicado a presentar la ingeniería ontológica y una de sus principales herramientas, los editores de ontologías, con la intención de considerar las posibilidades de aplicación de esta herramienta en el trabajo terminológico automatizado.

En la segunda parte de esta tesis doctoral se presenta el trabajo empírico realizado con el fin de desarrollar los objetivos específicos de nuestro trabajo, que son: el desarrollo de un catálogo de relaciones completo, la formalización de las relaciones conceptuales y su implementación en el editor de ontologías *Protégé*. En el capítulo 5 se presentan las fases del trabajo empírico y los recursos empleados para llevarlo a cabo. El capítulo 6 está dedicado a la descripción del catálogo de relaciones y su formalización e implementación en el editor de ontologías *Protégé*, y en el capítulo 7 se muestra el proceso de elaboración de una base de datos conceptual sobre los productos cerámicos acabados, creada utilizando el editor de ontologías *Protégé*, en la que la información de las relaciones conceptuales se representa de acuerdo con el modelo de formalización propuesto. En este mismo capítulo se propone una serie de consultas sobre relaciones conceptuales que nos permitirán comprobar la utilidad de nuestra propuesta para la recuperación de información.

En el capítulo 8 se presentan las conclusiones y las nuevas vías de investigación que abre este trabajo.

**CAPÍTULO 5: DESCRIPCIÓN DEL TRABAJO
EMPÍRICO**

5 Descripción del trabajo empírico

En este capítulo se presenta el trabajo empírico realizado en el marco de esta tesis doctoral que, como se ha dicho en la Introducción, se inscribe dentro de los proyectos de investigación TXTCERAM⁴⁹ y ONTODIC⁵⁰, y cuyo principal objetivo es el desarrollo e implementación de una ontología terminológica en el ámbito de la cerámica industrial. Nuestra contribución a dichos proyectos es una propuesta de formalización e implementación de un catálogo de relaciones conceptuales con la finalidad de almacenar y recuperar mediante un editor de ontologías la información recogida sobre los conceptos y sus relaciones.

En el apartado 5.1 se describen las fases seguidas en la elaboración del trabajo.

El apartado 5.2 está dedicado a la presentación de los materiales textuales y terminológicos que nos han servido como fuente de información sobre las relaciones conceptuales. Se describen los textos empleados para la detección preliminar de relaciones, así como el corpus textual *TXTCERAM* y la base de datos terminológica *Cerámica*, con la que se ha completado el conocimiento sobre las relaciones entre conceptos que nos han permitido elaborar nuestra base de conocimiento.

El apartado 5.3 se describen las herramientas informáticas empleadas en el trabajo empírico: el programa de análisis de corpus textuales *WordSmith Tools*, el programa de gestión terminológica *Multiterm* y el editor de ontologías *Protégé*.

En el apartado 5.4 se ofrece una breve recapitulación de las fases del trabajo y la relación de los recursos presentados en este capítulo.

⁴⁹ El proyecto “TXTCERAM. Extracción semiautomática y análisis conceptual formal de términos de la cerámica a partir de un corpus electrónico. Su eficacia y utilidad en la mediación lingüística”, dirigido por la Dra. Amparo Alcina, está financiado por la Generalitat Valenciana (código del proyecto: GV05/260).

⁵⁰ El proyecto “ONTODIC. Metodología y tecnologías para la elaboración de diccionarios onomasiológicos basados en ontologías. Recursos terminológicos para la e-traducción”, dirigido por la Dra. Amparo Alcina, está financiado por el Ministerio de Ciencia y Tecnología (código del proyecto: TSI2006-01911).

5.1 Fases del trabajo empírico

El trabajo empírico se ha desarrollado en cinco fases. La primera consiste en la elaboración de un catálogo de relaciones completo, basado en la revisión teórica de las clasificaciones propuestas desde la teoría de la Terminología y la Lingüística y en la observación de las relaciones conceptuales en el ámbito de la industria cerámica. A continuación se ha llevado a cabo la formalización de las relaciones en función de las clases conceptuales implicadas en las mismas y de sus propiedades matemáticas. En tercer lugar, se ha afrontado la implementación de los conceptos y sus relaciones en el editor de ontologías *Protégé*. En la cuarta fase se han introducido los conceptos de la cerámica relacionados con el producto cerámico acabado. La quinta fase ha consistido en la resolución de dudas conceptuales por parte de una experta en el ámbito de la cerámica industrial. Por último, se han planteado una serie de consultas onomasiológicas sobre relaciones a las que podríamos dar respuesta a partir de la información contenida en la base de conocimiento. A continuación se resume brevemente cada uno de estos pasos:

a) Elaboración de un catálogo de relaciones conceptuales

En primer lugar, hemos revisado las clasificaciones de las relaciones conceptuales que se han presentado en los fundamentos teóricos. Esta revisión nos ha permitido constatar que, en general, las propuestas carecen del grado de formalización y sistematicidad necesario para implementar las relaciones en un editor de ontologías. Por otro lado, algunas relaciones conceptuales complejas, como la meronimia, no aparecen descompuestas en subtipos.

El estudio de las diferentes clasificaciones nos ha llevado a proponer un nuevo catálogo de relaciones exhaustivo, que pretende integrar las distintas propuestas, y ofrecer de este modo un catálogo completo y aplicable a todas las áreas de especialidad.

Se ha contrastado el catálogo propuesto mediante un análisis preliminar de las relaciones conceptuales en el ámbito de la cerámica, llevado a cabo sobre un texto introductorio sobre productos cerámicos acabados, lo cual nos ha permitido constatar la validez de nuestra propuesta.

El resultado de esta primera fase, que se presenta en el apartado 6.1, es un catálogo de relaciones conceptuales definidas en lenguaje natural, que se tomará como punto de partida para su formalización.

b) Formalización de las relaciones conceptuales

A partir del catálogo de relaciones fruto de la fase anterior, se ha abordado la descripción formal de cada relación por medio de sus propiedades matemáticas (simetría, transitividad, relaciones inversas y recursividad) y de las clases conceptuales (entidades, propiedades y actividades) entre las que puede establecerse cada una de ellas. El resultado de esta fase se presenta en el apartado 6.2.

c) Implementación de los conceptos y relaciones conceptuales en el editor de ontologías

Una vez determinado el catálogo de relaciones conceptuales y establecidas sus propiedades y las clases que participan en cada una de ellas, se han analizado sus posibilidades de implementación en el editor de ontologías *Protégé*. Para ello se ha contemplado el modo en el que los conceptos y sus relaciones pueden representarse en una base de datos conceptual elaborada empleando dicho editor. Es decir, en esta fase, presentada en el apartado 6.3, se especifica cómo se introducirán los conceptos, las relaciones (con sus propiedades) y las denominaciones en *Protégé*.

d) Introducción de los conceptos sobre producto cerámico acabado y sus conceptos relacionados

Con el fin de valorar la utilidad de nuestra propuesta de formalización e implementación del catálogo de relaciones conceptuales en el editor de ontologías *Protégé*, en esta fase se ha creado una base de datos conceptual, denominada *Ontoceram*. En ella se han almacenado los conceptos pertenecientes a la rama de la cerámica correspondiente a los tipos de producto acabado, y sus conceptos relacionados. Cada concepto se ha analizado a partir de las fuentes terminológicas y textuales descritas en el apartado 5.3. El resultado de esta fase es una base de conocimiento, en la que se expresan de manera explícita y formalizada las relaciones que los conceptos relativos al producto acabado establecen entre sí y con otros conceptos del ámbito de la cerámica, como los materiales que componen los productos, sus características y los procesos de fabricación.

e) Resolución de dudas sobre el ámbito de la cerámica con ayuda de un experto

La información sobre los conceptos recogida a partir del análisis de definiciones y contextos no siempre es completa ni suficientemente aclaratoria.

Durante el proceso de introducción de los datos han surgido algunas dudas relacionadas con la industria cerámica, como el lugar que debe ocupar determinado concepto en la jerarquía, es decir, cuáles son sus conceptos superordinados y subordinados, o la adecuación de las denominaciones detectadas durante el análisis, entre otras. Para resolverlas, hemos contado con la ayuda de un miembro del Instituto Técnico de la Cerámica (ITC) de la Universitat Jaume I⁵¹, cuyos conocimientos sobre la industria cerámica nos han servido para asegurar que los datos introducidos son adecuados desde la perspectiva del área de especialidad.

Las dudas que hemos planteado a la experta han sido del tipo ¿es lo mismo *esmalte* que *vidriado*? o ¿en qué se diferencian las *piezas especiales* de las *piezas complementarias*?

En otros casos no nos ha sido posible asegurar a ciencia cierta si un compuesto polilexemático debía considerarse o no un término lexicalizado, como por ejemplo *producto vitrocerámico*. Este tipo de consultas también han sido resueltas por la experta.

En cuanto a la posible ubicación de los conceptos en la jerarquía, hemos planteado a la experta preguntas como las siguientes: ¿una *pieza complementaria* es un tipo de *producto cerámico acabado*? o ¿la *pasta* es una *materia prima* o un *producto intermedio*?

La resolución de estas cuestiones por parte de la experta garantiza que la organización de la información contenida en la base de conocimiento es adecuada y que las denominaciones son las que realmente manejan los especialistas en la industria cerámica.

f) Generación de consultas

La última fase del estudio se ha dedicado a analizar las diferentes posibilidades de recuperación de la información a partir de los conceptos almacenados en la base de conocimiento. Se proponen una serie de consultas que podrían ejecutarse y que permitirían recuperar de manera onomasiológica la información sobre relaciones conceptuales. De este modo se muestran, por un

⁵¹ Agradecemos la colaboración en esta fase de D^a. María Jesús José Ferrandis, licenciada en Ciencias Químicas y responsable del Centro de Documentación del ITC.

lado, el tipo de consultas que se pueden plantear a la base de conocimiento y, por otro, las posibilidades de consulta que proporciona el editor de ontologías.

5.2 Recursos textuales y terminológicos empleados en el trabajo empírico

Para la realización del trabajo empírico hemos empleado varios recursos textuales y terminológicos que se describen en este apartado. Estos recursos son los siguientes:

1. Un texto introductorio sobre baldosas cerámicas, la *Guía de la baldosa cerámica* (ASCER et al, 2003).
2. Dos diccionarios especializados sobre términos cerámicos, los de Guillem y Guillem (1987) y McColm (1994)
3. El árbol de campo de los conceptos de la industria cerámica
4. La base de datos terminológica *Cerámica*
5. El corpus textual digitalizado *TXTCERAM*

Estos materiales textuales y terminológicos se seleccionaron y desarrollaron en el marco de los proyectos “Descripción terminológica de una rama profesional de la industria cerámica. Elaboración de un diccionario multilingüe”⁵² y “TXTCERAM. Extracción semiautomática y análisis conceptual formal de términos de la cerámica a partir de un corpus electrónico. Su eficacia y utilidad en la mediación lingüística”⁵³.

Para la selección de los textos especializados y diccionarios sobre cerámica se aplicaron un conjunto de criterios que permiten garantizar su fiabilidad. Cabré (1999: 280-281) identifica las siguientes condiciones que debe cumplir la documentación para el trabajo terminológico:

- Debe ser **pertinente**, es decir, representativa del campo en el que se trabaja y, a ser posible, estar redactada por un **autor de calidad**.

⁵² Proyecto financiado por Caixa de Castelló/Bancaja en el período 1998-2000 y por la Consellería de Educación y Ciencia de la Generalitat Valenciana entre 2001 y 2003, dirigido por la Dra. Pilar Civera (código del proyecto: GV00-143-9).

⁵³ Proyecto financiado por la Generalitat Valenciana, dirigido por la Dra. Amparo Alcina Caudet (código del proyecto: GV05/260).

- Ha de ser suficientemente **especializada** para que permita identificar bien los conceptos propios del área temática.
- Debe ser **completa**, y por tanto incluir todos los aspectos del ámbito de trabajo.
- Debe ser **actual** para que represente los términos que se emplean de manera efectiva en el momento presente.
- Se recomienda que sea **original** en la lengua en la que se desarrolla el trabajo.
- Ha de ser **explícita**, es decir, los documentos deben estar bien identificados y ser accesibles.

Para cumplir con estos criterios, los textos seleccionados son obra de autores especialistas en el ámbito de la cerámica, que contienen información fiable y adecuada desde el punto de vista del ámbito especializado. En su mayoría se trata de materiales publicados que gozan del respaldo de alguna institución, o bien cuyos autores poseen una trayectoria reconocida en el ámbito de la industria cerámica. Se procuró además que las obras seleccionadas fueran originales y no traducciones, y relativamente actualizadas, ya que la industria cerámica es un campo del saber en constante evolución, y las fuentes empleadas deben recoger, en la medida de lo posible, los últimos avances tecnológicos.

En los siguientes apartados se describen en detalle los recursos textuales especializados empleados, indicando en cada caso cómo hemos trabajado con ellos en la presente investigación. Las referencias bibliográficas de todos estos materiales se incluyen en el Anexo 1.

5.2.1 *Guía de la baldosa cerámica (2003)*

La redacción de esta guía, editada en 2003, es el fruto del trabajo conjunto de varias asociaciones en el marco del convenio de colaboración entre el Colegio Oficial de Arquitectos de la Comunidad Valenciana y la Consellería de Obras Públicas, Urbanismo y Transportes de la Generalitat Valenciana. Por tanto, su pertinencia queda garantizada por el respaldo oficial que posee la obra y por el hecho de que sus editores gozan de una reconocida autoridad en el campo de la industria cerámica.

La *Guía de la baldosa cerámica* recoge la descripción y clasificación de los distintos tipos de baldosas cerámicas con el fin de ofrecer a los diversos

participantes en el proceso de la construcción un instrumento para la selección de productos que, al mismo tiempo, garantice la actualización de sus conocimientos.

Este documento se ha utilizado en la fase de elaboración del catálogo de relaciones conceptuales para contrastar la validez del mismo en el ámbito de la cerámica. En él se han identificado los conceptos relacionados con las baldosas cerámicas, y nos ha servido para analizar la naturaleza de dichas relaciones.

5.2.2 *Diccionario cerámico científico-práctico (1987)*

Los autores de esta obra lexicográfica, publicada por la Sociedad Española de Cerámica y Vidrio, son Claudio Guillem, director técnico de la empresa Lladró, y M^a Carmen Guillem, profesora titular de universidad de Química Orgánica.

Según indican los propios autores en el prólogo, el diccionario recoge 2.396 términos en español relacionados con el ámbito de la cerámica y sus equivalentes en inglés. El diccionario fue editado en 1987, por lo que no se puede considerar un diccionario excesivamente actualizado. No obstante, el hecho de que sea el único diccionario de cerámica en español que contiene definiciones hace que sea una obra de referencia en este campo.

Este diccionario, así como el que se presenta en el apartado siguiente, se han empleado en el trabajo empírico para observar en sus definiciones posibles conceptos relacionados con el concepto que aparece en cada entrada, puesto que las definiciones de los términos son contextos ricos en conocimiento, que aportan información sobre relaciones conceptuales. Meyer (2001:279) define los contextos ricos en conocimiento (*knowledge-rich contexts*) como contextos que expresan información conceptual de un término, y que por tanto resultan útiles para la construcción de definiciones y para la adquisición de conocimientos sobre el dominio.

5.2.3 *Dictionary of ceramic science and engineering (1994)*

Esta obra lexicográfica, editada en 1994 en lengua inglesa, carece de una introducción en la que se explique su cobertura temática, si bien el título y la revisión del diccionario nos confirman que éste es específico del ámbito de la cerámica, y que en él se contienen tanto términos científicos como técnicos. El autor es un experto en la materia, que ha publicado varios libros y artículos sobre

ingeniería cerámica. Se trata de un diccionario editado por una editorial científica de prestigio y recomendado por los expertos consultados, lo que nos hace confiar en la veracidad y adecuación de la información contenida en él.

Es necesario en este punto justificar por qué hemos seleccionado un diccionario en inglés. En este trabajo nos interesa extraer información sobre los conceptos y sus relaciones conceptuales. En principio, los procesos de la industria cerámica están muy internacionalizados y normalizados, por lo que el grado de variación conceptual entre las dos lenguas es muy reducido. Dado que se trata de una de las obras de referencia recomendadas por los expertos en cerámica, y que es relativamente actual, hemos considerado su utilidad para la extracción de información sobre relaciones conceptuales a partir de las definiciones de los términos, que, como en el caso del diccionario de Guillem y Guillem (1987), nos proporcionan contextos ricos en conocimiento.

5.2.4 Base de datos terminológica *Cerámica*

La base de datos terminológica *Cerámica* fue elaborada en el marco del proyecto “Descripción terminológica de una rama profesional de la industria cerámica. Elaboración de un diccionario multilingüe”. Ha sido desarrollada utilizando el programa *Multiterm* descrito en el apartado 5.3.2 y está formada por 4.617 fichas terminológicas sobre el campo de la cerámica industrial. El contenido de dichas fichas se encuentra en español, aunque se incluyen equivalentes en inglés, francés, alemán e italiano para un gran número de términos.

La base de datos está estructurada en los siguientes campos de información:

- a) **Rama:** Este campo de atributo hace referencia a la rama del árbol de campo (descrito en el apartado 5.2.5) a la que pertenece el término.
- b) **Categoría gramatical:** En este campo se determina si el término es un nombre masculino (m.), un nombre femenino (f.), un adjetivo (adj.), un adverbio (adv.) o un verbo (v.).
- c) **Forma del término:** Este campo permite indicar si la forma del término es una forma abreviada, una sigla, o bien si se trata de la forma completa del mismo.
- d) **Contexto definitorio y fuente de la que proviene:** En el campo *Contexto definitorio* se recogen ejemplos extraídos de textos especializados reales

que aportan información conceptual sobre el término. El campo *Fuente* hace referencia al texto del que se ha extraído el ejemplo.

- e) **Contexto lingüístico y su fuente:** En el campo *Contexto lingüístico* se recogen ejemplos que ilustran el uso lingüístico del término. Al igual que en el caso anterior, este campo va acompañado de un campo *Fuente*, que indica el texto especializado del que proviene dicho ejemplo de uso del término.
- f) **Nota:** El campo nota se ha empleado en esta base de datos para que los terminólogos incluyan información útil sobre el desarrollo del trabajo terminológico, como por ejemplo si existen variantes denominativas que conviene seguir investigando.
- g) **Ilustración:** Este campo se emplea para incluir la referencia a tablas o figuras procedentes de textos especializados que sirven para aclarar el significado del término.

En nuestro trabajo, la base de datos *Cerámica* se ha empleado para la extracción de información conceptual sobre los términos a partir de los contextos definitorios y lingüísticos. Asimismo, dado que los términos aparecen clasificados por ramas de acuerdo con el árbol de campo descrito en el siguiente apartado, la base de datos nos ha servido para seleccionar los términos relacionados con el producto cerámico acabado, que son los que hemos introducido en la base de datos conceptual elaborada con el editor de ontologías *Protégé*. Los equivalentes en inglés nos han permitido buscar información conceptual en las definiciones del diccionario McColm (1994) descrito en el apartado 5.2.3.

5.2.5 Árbol de campo de la cerámica

El árbol de campo que subyace a la base de datos terminológica *Cerámica* fue diseñado en el marco del proyecto “Lenguaje de especialidad y terminología de la cerámica”. Refleja la estructuración básica del ámbito temático en el que se ha trabajado: la industria cerámica. Posee la siguiente estructura, descrita en Civera y Alcina (2000) y Civera (2002):

1. Producto

1.1 Producto acabado

1.1.1 Características del producto acabado

- 1.1.2 Defectos del producto acabado
- 1.1.3 Tipos de producto acabado
- 1.1.4 Partes del producto acabado
- 1.2 Producto intermedio
 - 1.2.1 Características del producto intermedio
 - 1.2.2 Defectos del producto intermedio
 - 1.2.3 Tipos de producto intermedio
- 1.3 Otros productos
- 2. Fabricación
 - 2.1 Procesos
 - 2.2 Maquinaria y accesorios
- 3. Materias primas y aditivos
 - 3.1 Materias primas
 - 3.2 Aditivos
- 4. Colocación
 - 4.1. Procesos de colocación
 - 4.2. Maquinaria y accesorios de colocación
 - 4.3. Materiales para colocación
 - 4.3.1 Materiales de base
 - 4.3.2 Materiales de adherencia
- 5. Calidad
 - 5.1. Procesos de control de calidad
 - 5.2. Maquinaria y accesorios de control de calidad
- 6. Unidades de medida
- 7. Organismos, Instituciones

En el presente trabajo, el árbol de campo nos ha servido para delimitar la rama en la que vamos a desarrollar nuestra base de conocimiento (1.1.3 Tipos de producto acabado). También nos ha aportado una visión global del área temática, que ha sido de gran ayuda a la hora de determinar las clases de conceptos implicadas en los procesos cerámicos.

5.2.6 Corpus textual *TXT CERAM*

Este corpus textual especializado en formato electrónico es el fruto del trabajo de selección y digitalización de textos llevado a cabo en el marco del

proyecto de investigación TXTCERAM, en estrecha colaboración con el proyecto de innovación educativa CREC para la Creación de Recursos Lingüísticos⁵⁴. El proceso de digitalización de los textos se describe en Soler, Alcina y Estellés (2006).

El corpus en español –que es el que se ha empleado en este trabajo– está constituido por un total de 34 referencias bibliográficas sobre el ámbito de la cerámica y consta de 2.444.791 palabras. En el Anexo 1 se recogen las referencias bibliográficas de los textos incluidos en el corpus en español.

En la Tabla 18 se resume la descripción del corpus *TXTCERAM*, teniendo en cuenta los parámetros propuestos por Bowker y Pearson (2002). Estos parámetros son el tamaño del corpus, el tipo de textos seleccionados (textos completos o extractos y las condiciones pragmáticas de producción de los mismos), el número de textos, el medio en el que aparecen, el tema, la autoría, la lengua y la cobertura temporal:

PARÁMETROS DEL CORPUS TEXTUAL <i>TXTCERAM</i>	
Tamaño del corpus	2.444.791 palabras
Tipo de textos	Textos completos en los que tanto el emisor como el receptor son especialistas. La situación comunicativa es formal y el tratamiento del tema es especializado.
Número de textos	34
Medio	Escrito
Tema	Cerámica industrial
Autoría	Los autores de los textos son especialistas en la materia.
Lengua	Español
Cobertura temporal	Textos publicados entre 1980 y 2004

Tabla 18. Descripción del corpus *TXTCERAM* según los parámetros propuestos por Bowker y Pearson (2002).

En el momento de la realización de esta investigación el corpus está constituido por texto crudo, sin etiquetado morfosintáctico ni semántico. En el marco del proyecto TXTCERAM está previsto llevar a cabo la anotación morfosintáctica y macroestructural de los textos, con el fin de extraer información conceptual a partir del corpus (Estellés y Alcina, 2006).

⁵⁴ Proyecto de innovación educativa financiado por la Unitat de Suport Educatiu de la Universitat Jaume I (2004-2006).

En este trabajo hemos empleado el corpus textual digitalizado *TXTCERAM* para la extracción de concordancias de los términos estudiados. Estas concordancias se han empleado para completar la información conceptual obtenida a partir de los diccionarios y de la base de datos terminológica descritas anteriormente. Las concordancias nos han ayudado a comprender el significado de las relaciones conceptuales, así como para completar la detección de relaciones conceptuales que no se habían identificado a partir de los diccionarios y la base de datos terminológica.

5.3 Descripción de las herramientas informáticas empleadas

En este apartado se describen las herramientas informáticas empleadas para llevar a cabo la búsqueda de información conceptual sobre relaciones en los recursos textuales y el almacenamiento de los datos extraídos. En el apartado 5.3.1 describimos la herramienta de análisis textual *WordSmith Tools*, que se ha utilizado para extraer concordancias de los conceptos. El apartado 5.3.2 está dedicado a la descripción del programa de gestión terminológica *Multiterm*, en el que está almacenada la base de datos *Cerámica*, descrita en el apartado 5.2.4. Por último, en el apartado 5.3.3 se presenta el editor de ontologías *Protégé*, que nos ha servido para desarrollar una base de conocimiento sobre los productos cerámicos acabados y sus conceptos relacionados. Puesto que las herramientas de análisis textual y de gestión terminológica son ampliamente conocidas por los expertos y profesionales en terminología, su descripción se hará de manera sucinta. En cambio, consideramos que el empleo de los editores de ontologías creados desde la ingeniería del conocimiento para la gestión de la información conceptual en terminología es más novedoso, por lo que la descripción de *Protégé* es más exhaustiva. En el Anexo 2 se ofrece un apartado específico que recoge las referencias bibliográficas y las fuentes de internet en las que puede obtenerse información sobre el uso y características de estas herramientas informáticas.

5.3.1 Herramienta de análisis de corpus: *WordSmith Tools*

Para el análisis del corpus *TXTCERAM* se ha utilizado la versión 3.0 de *WordSmith Tools*, programa desarrollado por el Prof. Mike Scott y disponible a través de Oxford University Press⁵⁵.

WordSmith Tools es un conjunto de programas desarrollado con el fin de explotar la información lingüística y conceptual contenida en un corpus textual en formato electrónico. En nuestra investigación hemos empleado la herramienta *Concord* para la obtención de concordancias, si bien *WordSmith Tools* ofrece otras posibilidades de análisis, como la elaboración de listas de palabras (estadísticas, por frecuencia y alfabéticas), la detección de palabras clave al comparar con un corpus de referencia y la posibilidad de alinear corpus paralelos, que no se han explotado en este trabajo.

Concord permite ver en contexto todas las apariciones del término de búsqueda definido por el analista en el corpus seleccionado previamente. La información se recupera en forma de índice de palabras en contexto (KWIC, *key words in context*), lo que nos permite observar, por un lado, el número absoluto de veces que aparece un término determinado en el corpus textual y, por otro, las palabras que aparecen en su entorno. Podemos restringir o ampliar el número de palabras que queremos que aparezcan a derecha e izquierda del término.

Concord ofrece como resultado una tabla en la que pueden observarse las apariciones del término de búsqueda y las palabras que lo acompañan. En el caso que nos ocupa, para extraer el máximo de información sobre relaciones conceptuales, consideraremos las oraciones completas en las que aparece un término determinado, en las que podemos ver la información sobre los conceptos relacionados con el término buscado.

En la Fig. 22 se ofrece un ejemplo de las concordancias del término *baldoa cerámica* en el corpus *TXTCERAM*.

⁵⁵ Para obtener una información exhaustiva sobre las características y el funcionamiento del programa consúltese la página personal de Mike Scott (SCOTT, M. [en línea]. *WordSmith Tools*. <http://www.lexically.net/wordsmith/> [Consulta: junio de 2007].

The screenshot shows a window titled 'Concord - [BALDOSAS CERÁMICA: 122 entries (sort: 5L,5L)]'. The interface includes a menu bar (File, View, Settings, Window, Help), a toolbar with various icons, and a main text area displaying search results. The results are organized into columns: 'N', 'Concordance', 'Set', 'Tag', and 'Word No.'. The text in the concordance column is highlighted in blue, showing the term 'baldosa cerámica' in its original context within various technical documents. The 'Word No.' column shows the corresponding line number in the source text.

Fig. 22. Concordancias del término *baldosa cerámica* obtenidas con el programa *Concord*.

Podemos ver cómo para cada aparición de *baldosa cerámica* en el corpus, *Concord* nos ofrece una línea de información en la que el término aparece en el centro del índice KWIC, rodeado por las palabras que configuran su contexto lingüístico.

Tomemos a modo de ejemplo el siguiente contexto del término *baldosa cerámica* recuperado por *WordSmith Tools* a partir del corpus *TXT CERAM*:

Las baldosas cerámicas se dividen en grupos según su método de fabricación (moldeo) y su absorción de agua.

Esta concordancia nos aporta información relevante sobre los conceptos relacionados con el concepto *baldosa cerámica*. A la vista del contexto podemos concluir que:

- a) El *moldeo* es un *método de fabricación* de las *baldosas cerámicas*. A partir de este contexto obtenemos información sobre dos relaciones distintas: por un lado, la relación que se establece entre *moldeo* y *método de fabricación* (relación

subordinado-superordinado) y por otro, la relación que existe entre *moldeo* y *baldosa cerámica* (relación proceso-paciente).

- b) La *absorción de agua* es una característica técnica de las *baldosas cerámicas* (relación característica-objeto).
- c) El moldeo y la absorción de agua son dos criterios que nos permitirán distinguir entre distintos tipos de baldosa.

En conclusión, el análisis de los términos en contexto facilita el acceso a la información sobre las relaciones que se establecen entre los conceptos de un ámbito temático. El gran número de contextos recuperados para cada término permite acceder a gran cantidad de información sobre los conceptos.

5.3.2 Herramienta de gestión de bases de datos terminológicas:

Multiterm

Multiterm es uno de los programas integrados en el paquete de herramientas de traducción asistida por ordenador (TAO) desarrollado por la empresa Trados, que se compone de una memoria de traducción (*Translator's Workbench*), un programa de alineación de textos paralelos (*WinAlign*) y otros programas accesorios como *Tag Editor* o *T-Window*. Se completa con un gestor de datos terminológicos, *Multiterm*, que es el que se ha utilizado para la elaboración de la base de datos terminológicos *Cerámica*.

Multiterm facilita la creación y la gestión de bases de datos terminológicas de manera flexible, ya que el usuario puede definir los campos de información sobre los términos que se ajustan a sus necesidades. Estos campos, de estructura jerárquica, pueden ser de tres tipos:

- a) **Campos de índice**, que recogen los términos en diversas lenguas y facilitan la búsqueda en la base de datos.
- b) **Campos de texto**, que permiten introducir información textual, como definiciones y contextos para cada término.
- c) **Campos de atributo**, en los que se recoge información adicional sobre el término y referencias a los documentos empleados como fuente. Los valores de los campos de atributo son predefinidos por el administrador de la base de datos para agilizar su introducción y asegurar que ésta posee una estructura homogénea.

El programa contempla además diversas opciones de búsqueda (simple, global o difusa) y permite aplicar filtros para utilizar únicamente una parte de la información contenida en la base de datos.

Una de las características más útiles de *Multiterm* es la función que permite definir distintos tipos de exportación para generar de manera sencilla listas de términos y glosarios preparados para la impresión en diversos formatos a partir de los registros contenidos en la base de datos.

Todas estas características hacen de *Multiterm* una herramienta apropiada para la presentación de los datos extraídos del corpus textual una vez procesado. En el apartado 5.2.4 se ha descrito en detalle la estructura de la base de datos terminológica *Cerámica*.

En la Fig. 23 se ofrece una imagen que muestra un fragmento de la ficha de *baldosa cerámica* contenida en la base de datos terminológica *Cerámica*.

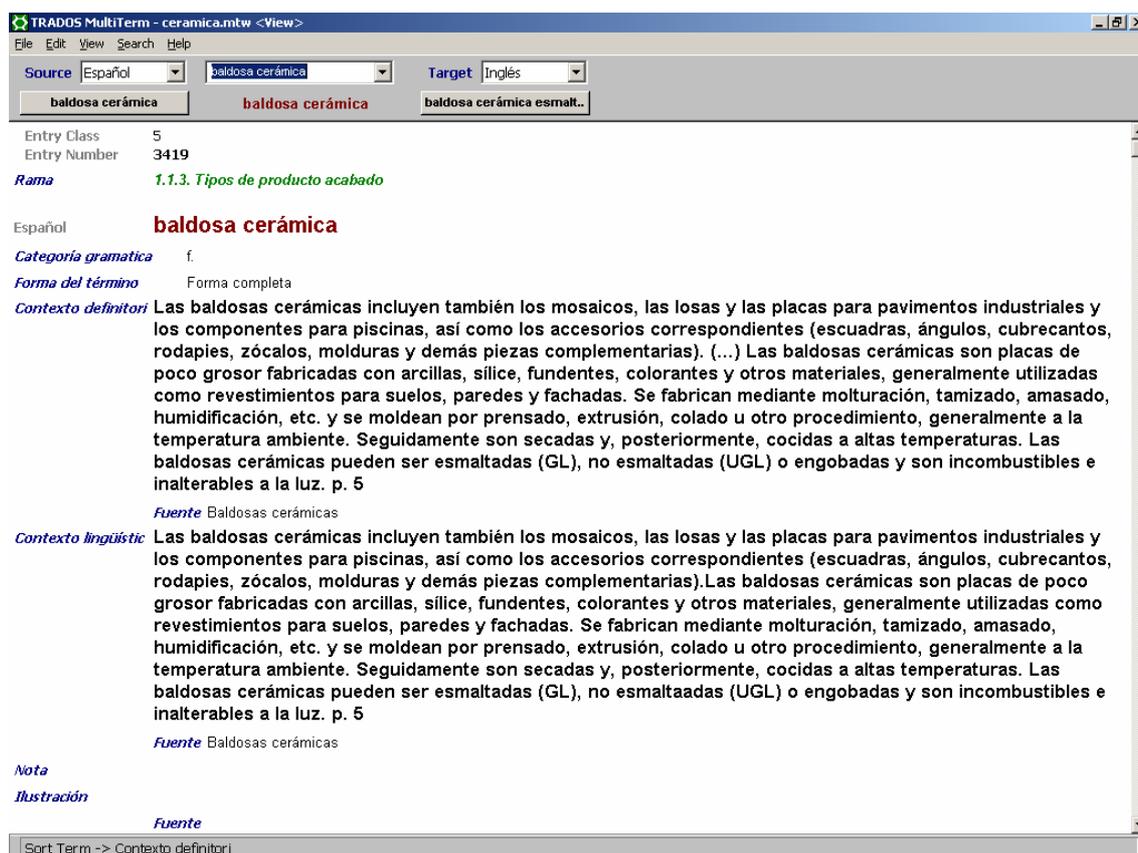


Fig. 23. Ficha terminológica del término *baldosa cerámica* en *Cerámica*.

En este extracto de la ficha terminológica podemos comprobar cómo se recoge la información sobre el término *baldosa cerámica*. Vemos que aparece información sobre la rama del árbol de campo a la que pertenece (1.1.3. Tipos de producto acabado), su categoría gramatical (femenino), el tipo de forma del

término de que se trata (forma completa) y sus contextos definatorios y lingüísticos. Los sinónimos de este término, si los hubiera, también aparecerían reflejados en esta ficha, junto con los equivalentes del término en las otras lenguas.

5.3.3 Editor de ontologías: *Protégé*

Como se ha visto en el capítulo 4, los editores de ontologías son programas desarrollados desde la ingeniería del conocimiento que facilitan la elaboración de ontologías desde cero. Una ontología es una representación estructurada del conocimiento en la que se almacena de forma explícita información consensuada sobre los fenómenos del mundo. Además de permitir la edición y la navegación por las ontologías, los editores ayudan en las fases de documentación, importación y exportación en diferentes formatos y lenguajes ontológicos, y en la edición gráfica de los datos, entre otras.

Protégé es una herramienta de edición de ontologías desarrollada por el Departamento de Informática Médica de la Universidad de Stanford. Es una plataforma gratuita, basada en Java y de código abierto, que se emplea para modelar ontologías y realizar consultas. Incorpora funciones para presentar la información contenida en ellas de forma gráfica y en formato HTML. Además, las ontologías creadas con *Protégé* pueden exportarse fácilmente a otros formatos y lenguajes ontológicos, como el lenguaje RDFS o el lenguaje OWL. En el apartado 4.4.4 se ha descrito la arquitectura y el modelo de conocimiento que subyacen a esta herramienta, así como su capacidad para interoperar con diversas aplicaciones.

En este apartado se describe la herramienta de edición de ontologías y su funcionamiento en la ingeniería del conocimiento⁵⁶. Para ello, en primer lugar se presentan los componentes que permiten modelar ontologías en *Protégé* y más adelante se describen las principales funciones del editor.

⁵⁶ Queremos agradecer la inestimable colaboración de D. Ismael Sanz, profesor del Departamento de Ciencia e Ingeniería de los Computadores de la Universidad Jaume I, que además de impartir varios cursillos sobre el manejo de *Protégé* al grupo de investigación Tecnolettra nos ha asesorado en todas las fases de elaboración de nuestra base de conocimiento.

Los **componentes** de las ontologías modeladas en *Protégé* son *Classes*, *Slots* e *Instances*⁵⁷. A continuación definimos cada uno de ellos, indicando cuáles son sus principales rasgos distintivos y cómo se organizan en el editor de ontologías. Asimismo adelantamos brevemente cómo se han utilizado en nuestro trabajo empírico estos componentes, si bien este aspecto se desarrollará con más detalle en los capítulos 6 y 7.

a) Classes

Las **classes** se definen como una representación abstracta de un concepto⁵⁸ en un dominio. Se estructuran en una jerarquía organizada de forma taxonómica, en la cual se mantiene la herencia, es decir, las propiedades se transmiten desde las *classes* situadas en un nivel superior a las situadas en un nivel inferior de la jerarquía. Cada *class* puede tener más de una *superclass*, es decir, puede heredar sus propiedades a partir de más de un padre, lo que se denomina herencia múltiple. Las *classes* pueden ser concretas o abstractas, dependiendo de si pueden poseer instancias directas o no.

En *Protégé* es posible definir **Metaclasses**, que se utilizan como plantillas para crear *classes* con propiedades predefinidas.

En la Fig. 24 se ilustra la pestaña *Classes* del editor de ontologías mediante un ejemplo de la ontología *Newspaper*, suministrada con *Protégé* a modo de ejemplo. Esta ontología contiene información sobre los conceptos que pueden manejarse en la redacción de un periódico, como la lista de los artículos publicados, las secciones de un periódico, los empleados o la publicidad que se incluye. Las posibles aplicaciones de esta ontología podrían ser la recuperación de

⁵⁷ Hemos decidido mantener la nomenclatura en inglés de los componentes, tal y como aparecen en el programa. De este modo evitamos la posible confusión conceptual al referirnos a las *clases* en *Protégé*. En nuestro trabajo empleamos *clase de conceptos* para referirnos a entidades, actividades y propiedades, mientras que el componente *Class* del editor es el que nos permite introducir conceptos. No obstante, cuando no exista posibilidad de confusión, emplearemos la denominación *clase* para referirnos al componente *Class* del editor.

⁵⁸ Nótese que la forma de entender los conceptos en la ingeniería ontológica no es necesariamente la misma que en terminología, en donde los conceptos hacen referencia a la representación mental de un objeto obtenida a través de la abstracción de sus características. En la ingeniería ontológica el significado de *concepto* se emplea de un modo más general.

la información sobre artículos publicados, honorarios de los empleados u organización de los mismos.

En la parte izquierda de la figura puede verse el navegador (*Class Browser*), que permite editar la taxonomía y ver cuáles son los superordinados de una *class* determinada. En este caso podemos ver que el concepto *Columnist* posee dos superordinados (*Author* y *Employee*), de los que hereda sus propiedades. Dichas propiedades se recogen en la parte derecha de la pantalla, en el denominado *Class Editor*. En esta parte del editor se introduce la siguiente información: el nombre de la clase, si es concreta o abstracta, la descripción en lenguaje natural en el apartado *Documentation*, las restricciones axiomáticas en el lenguaje PAL y los *slots* y sus valores para dicha clase. Algunos ejemplos de estos *slots* asignados a la clase *Columnist* son *current_job_title*, *data_hired* o *name*, que poseen valores concretos para cada *Instance* dependiente de *Columnist*.

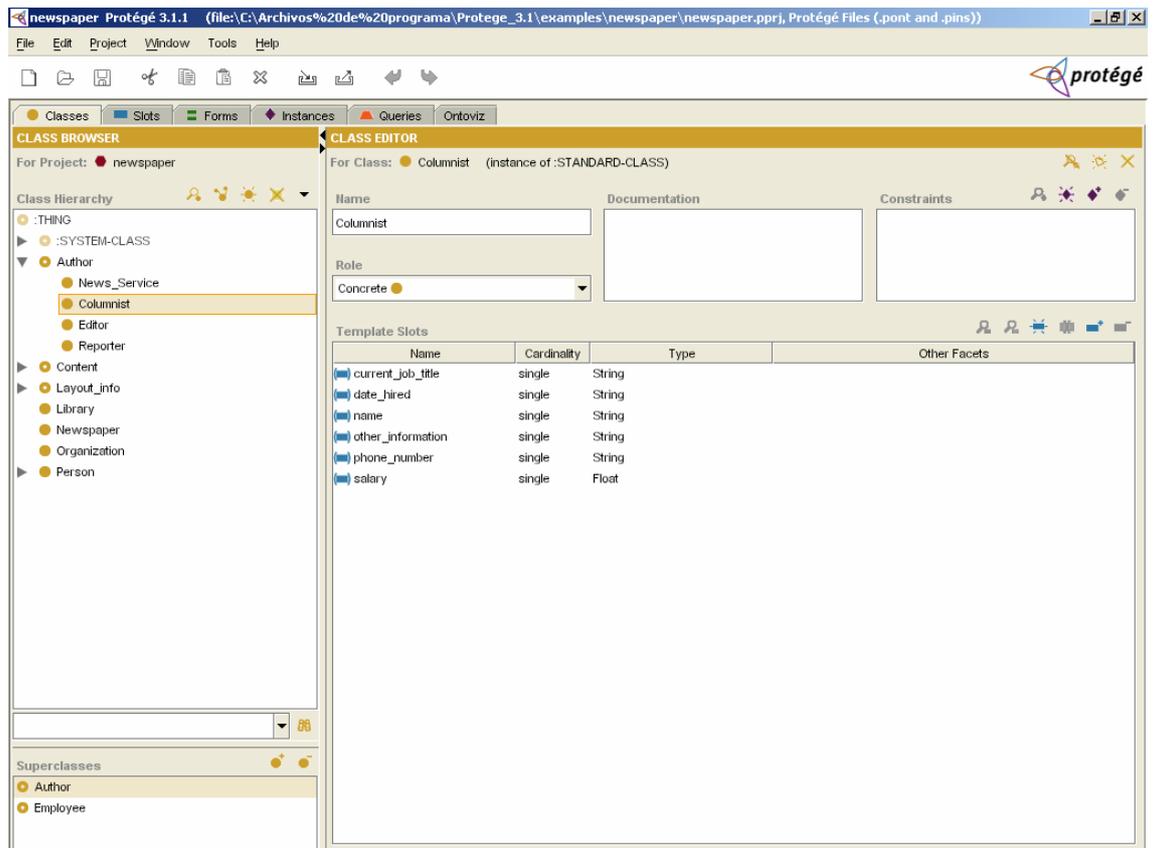


Fig. 24. La clase *Columnist* de la ontología *Newspaper*.

En el navegador de clases podemos crear nuevos subordinados para una clase, utilizar las funciones de búsqueda, copiar, pegar y arrastrar las clases en la estructura para recolocarlas en la taxonomía mediante el mecanismo *drag&drop*.

b) Slots

Los *slots* se definen en sentido amplio como propiedades asignadas a las *classes*. Aunque generalmente se considera que un *slot* está asociado a una clase determinada, se pueden manipular y definir de manera independiente mediante la pestaña *Slots* del editor y posteriormente asociarse a cualquiera de las *classes* que contiene la ontología. Cuando un *slot* se asocia a una clase determinada, sus valores pueden restringirse para ese concepto en concreto. Es lo que se conoce como *slot override*, que consiste en modificar los valores de un *slot* heredado o descrito de forma global en una clase.

Al igual que las *classes*, los *slots* pueden organizarse de forma jerárquica, lo que permite estructurar los diferentes tipos de atributos, diferenciando así el tipo de información sobre el concepto que proporcionan.

Para describir las características de los *slots* se emplean las facetas (*Facets*), que permiten introducir información sobre ellos, como por ejemplo el tipo de valor que pueden tener (*Class*, *Instance*, *Boolean*, etc.), el número de valores que pueden tener (*Cardinality*) o los valores por defecto del *slot*. En el capítulo 6 se revisarán con mayor detalle cuáles son las facetas estándar de los *slots*. A estas pueden añadirse otras que especifiquen aún más los valores del *slot* para adaptarlos a nuestras necesidades de representación.

En la Fig. 25 podemos ver cómo se disponen las facetas en la pestaña *Slots tab*. Tomamos como ejemplo el *slot containing_section* de la ontología *Newspaper*, que se utiliza para enumerar las secciones en las que se puede incluir un tipo de información determinado en un periódico:

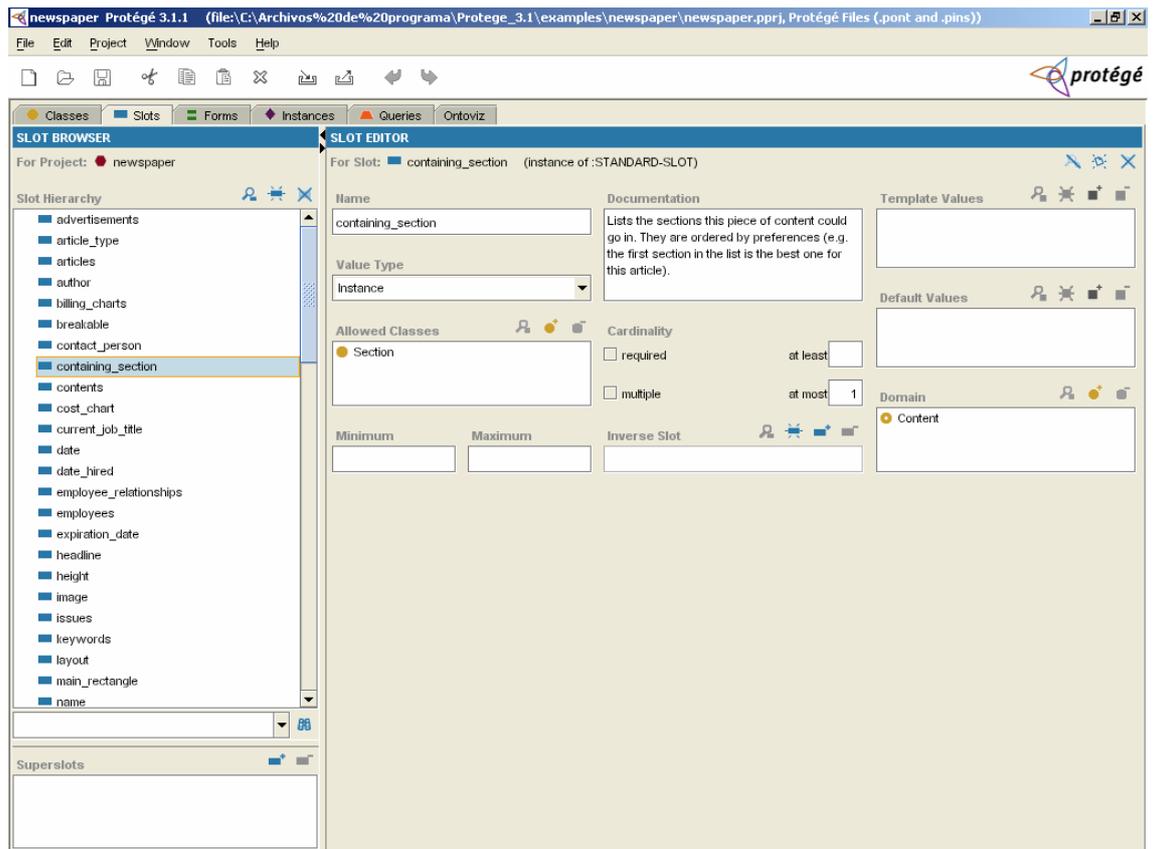


Fig. 25. Ejemplo del slot *containing_section* en la ontología *Newspaper*.

En la parte izquierda de la figura observamos el navegador de *slots*, y en la parte derecha las facetas del *slot* seleccionado y sus posibles valores. El *slot* *containing_section* permite establecer una relación entre el contenido y la sección del periódico en la que puede aparecer dicho contenido. Los tipos de valor posibles son *instances* que pertenezcan al concepto *Section*. Esta propiedad no es obligatoria (la casilla *required* no está marcada) y no puede ser múltiple (la casilla *multiple* aparece desmarcada y el número máximo de valores que puede tener es uno).

En la ontología *Newspaper* aparece un *slot* denominado *salary* definido para toda la ontología, utilizado para indicar el sueldo que percibe cada empleado del periódico, cuyo valor es una cifra (*Integer*). Supongamos que en la clase *Columnist* queremos restringir la cifra que indica el salario y decir que sólo puede estar entre 10.000 y 15.000 dólares. En ese caso, podríamos modificar el *slot* *salary* en la clase *Columnist* para limitar el valor que puede tener el *slot* para esa clase concreta.

El tipo de valores que pueden tener los *slots* permite especificar los atributos de las clases de la ontología y establecer relaciones entre ellas. Si los

slots pueden tener valores de tipo lógico, números (tanto enteros como con decimales), símbolos o una cadena de caracteres, entonces podemos representar los atributos escalares y literales de las clases. Si los valores que pueden tener los *slots* son *classes* o *instances*, entonces estaremos estableciendo relaciones entre los conceptos o instancias representados en la ontología. Definamos más concretamente a qué se refiere cada tipo de valor:

- a) Cualquiera (***Any***): El *slot* puede tener cualquier tipo de valor. Su valor específico se determinará para cada nueva clase o instancia incluidas en la ontología.
- b) Booleano (***Boolean***): El *slot* puede tener sólo uno de los dos valores lógicos booleanos: verdadero o falso.
- c) Cadena de caracteres (***String***): El valor del *slot* es una cadena de caracteres de texto en formato ASCII, que se escribe para cada clase o instancia concreta.
- d) Clase (***Class***): El valor del *slot* es otra clase de las contenidas en la ontología.
- e) Instancia (***Instance***): El valor del *slot* es una instancia de las definidas para alguna de las clases de la ontología.
- f) Número entero (***Integer***): El valor del *slot* es un número entero.
- g) Número decimal (***Float***): El valor del *slot* es un número con un decimal.
- h) Lista de valores predefinidos (***Symbol***): El valor del *slot* se escoge entre una lista de valores predefinidos y presentados en forma de lista.

En nuestro trabajo empírico utilizaremos los *slots* del editor de ontologías para establecer relaciones entre dos conceptos (representados en el componente *Class*) en los que el tipo de valor es otro concepto contenido en la base de conocimiento.

c) *Instances*

Los *Instances* en *Protégé* se definen como objetos concretos a los que se refiere la información sobre un dominio introducida en la ontología por medio de las *classes*. Modelan los datos de los individuos que se asignan a las clases concretas. Por ejemplo, para la clase *Columnist* podemos definir en la ontología

instancias concretas que se refieren a los empleados del periódico que trabajan como columnistas.

En *Protégé* los ejemplos concretos de una clase se gestionan mediante la pestaña *Instances*, cuya estructura del formulario para introducir los datos en el editor de instancias se define mediante la pestaña *Forms*.

En la Fig. 26 se muestran las instancias correspondientes a la clase *Columnist* en la ontología *Newspaper*. A la izquierda de la pantalla aparece la taxonomía de clases, indicando entre paréntesis el número de instancias que se asocian a cada clase. En la parte central aparece el navegador (*Instance browser*), en el que vemos las dos instancias dependientes de la clase *Columnist* (*Kelly* y *Kim*), es decir, los dos empleados que trabajan como columnistas en el periódico. En la parte derecha es posible delimitar los valores de estos ejemplos.

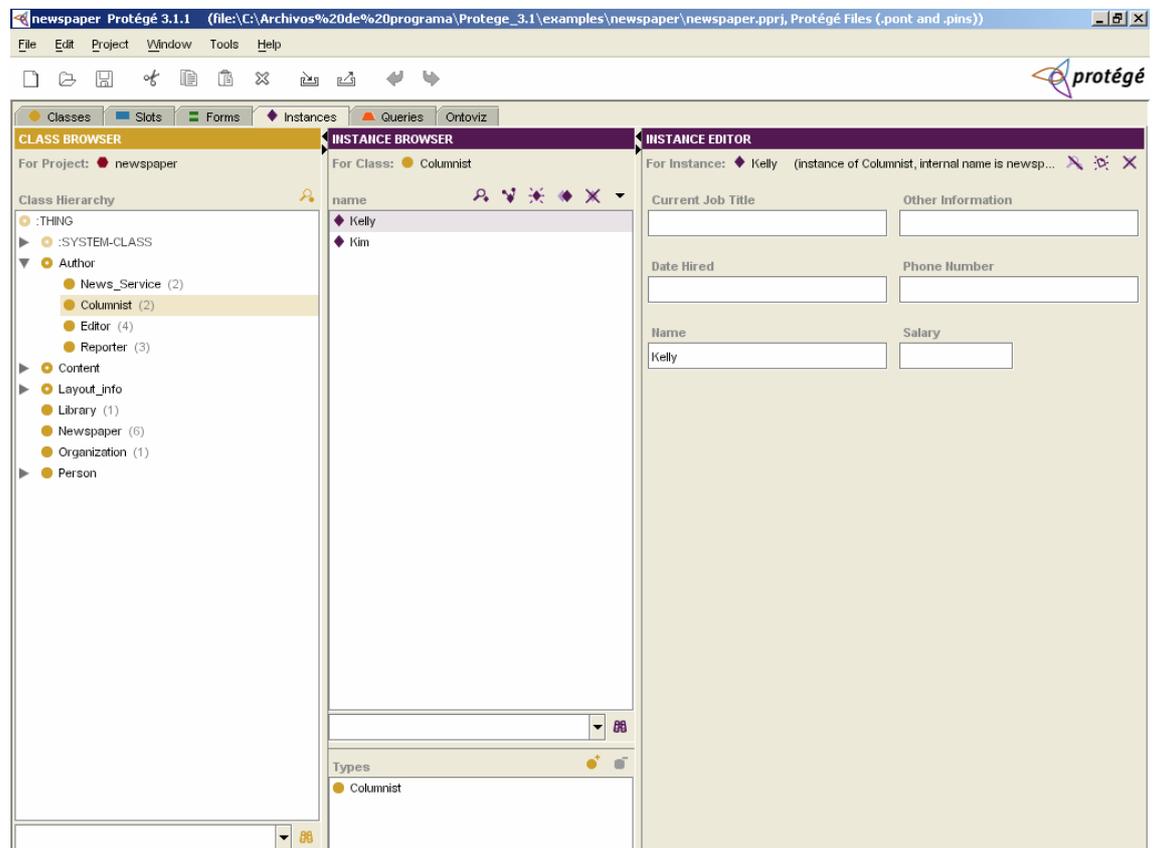


Fig. 26. *Instances* para la clase *Columnist* en la ontología *Newspaper*.

En nuestro trabajo empírico utilizaremos el componente *Instance* para reflejar las denominaciones empleadas para referirse a los conceptos representados por medio del componente *Class*, es decir, las instancias en nuestro trabajo representan las distintas posibilidades denominativas de los conceptos, y no objetos o ejemplos concretos.

Una vez presentados los componentes básicos que permiten crear y editar ontologías en *Protégé*, describimos brevemente algunas de las **funciones** que proporciona el editor, destacando las que nos servirán de ayuda en nuestra investigación. Además de las funciones de edición y almacenamiento de información conceptual ya descritas, *Protégé* incorpora otras que facilitan las siguientes tareas principales: consulta de las ontologías, visualización gráfica, importación y exportación de datos, gestión de proyectos, búsqueda y navegación. La estructura abierta de *Protégé* permite ampliar estas funciones mediante *plug-ins* desarrollados para adaptar y personalizar el entorno de trabajo a nuestras necesidades.

En este trabajo se han explorado principalmente las funciones de consulta. A continuación se presentan algunos ejemplos de cómo pueden utilizarse estas funciones.

En cuanto a la consulta de ontologías, además de las pestañas que nos permiten editar los principales componentes de la ontología (*Classes*, *Slots* e *Instances*), *Protégé* incorpora la pestaña *Queries*, mediante la que se definen consultas sobre las instancias de la ontología. Este *plug-in* se emplea para crear, ejecutar y guardar consultas con el fin de recuperar las instancias de la ontología que cumplen unos determinados criterios definidos en función de los valores de uno o varios *slots*.

Las consultas no forman parte integral de la base de conocimiento, pero permiten identificar las instancias contenidas en la ontología a través de sus propiedades representadas mediante *slots*. Esta pestaña consta de tres partes, que son el *Query Pane*, en el que se introduce y modifica la consulta, el *Search Results Pane*, que muestra los resultados de la búsqueda cuando se pulsa el botón *Find*, y por último la biblioteca de consultas (*Query Library*), en la que podemos almacenarlas y recuperarlas.

Veamos a continuación la consulta compleja *Articles whose authors are highly paid and articles are urgent with less than 50 pages* definida para la ontología *Newspaper*. Mediante esta consulta podemos recuperar los individuos de la clase *Article* cuyos autores reciben un salario elevado, que son urgentes y tienen menos de 50 páginas. La estructura de la pregunta en lenguaje natural sería la siguiente:



¿Qué artículos urgentes han sido escritos por autores muy bien remunerados y tienen menos de cincuenta páginas?

Como vemos en la Fig. 27, en la ontología *Newspaper* sólo hay un artículo que cumpla todos estos requisitos, *Cousteau Heads for Drydock*. Para definir esta consulta se han combinado tres consultas simples. La primera recupera aquellos artículos que posean un valor máximo de 50 páginas para el *slot page_number*; la segunda restringe los resultados de la anterior a los artículos que posean el valor *true* para el *slot urgent*; la tercera recupera los artículos en los que el valor del *slot author* contenga los resultados de otra consulta simple definida con anterioridad, la consulta *People with salary greater than 100,000*. Combinando todas estas consultas simples y pulsando el botón *Find*, obtenemos como resultado la instancia *Cousteau Heads for Drydock*, dependiente de la clase *Article*, que cumple todos estos criterios de búsqueda.

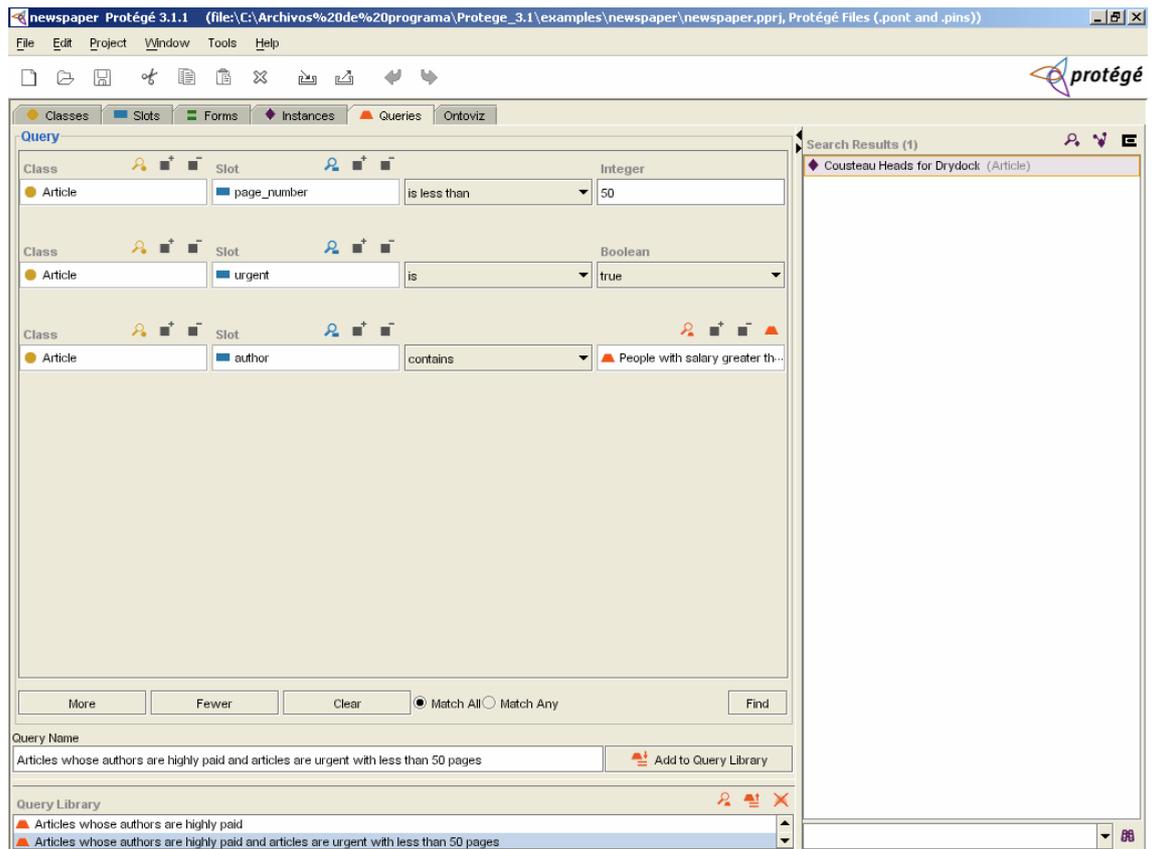


Fig. 27. Resultado de la consulta *Articles whose authors are highly paid and articles are urgent with less than 50 pages* en la ontología *Newspaper*.

La definición de consultas mediante la pestaña *Queries* es un instrumento válido para la recuperación de instancias contenidas en la ontología, si bien requiere que los valores de cada instancia estén definidos previamente.

En nuestro trabajo utilizaremos la pestaña *Queries* del editor de ontologías para recuperar en la base de conocimiento conceptos relativos a los productos cerámicos acabados a partir de los *slots* que indican relaciones.

Además de las consultas sobre instancias, *Protégé* posee un *plug-in* denominado *PAL Queries* que permite realizar consultas lógicas sobre *classes*, *slots* e *instances* contenidos en una ontología utilizando el lenguaje axiomático PAL⁵⁹ (*Protégé Axiom Language*). Con este *plug-in* es posible realizar consultas más complejas y potentes, pero tiene el inconveniente de que es necesario estar familiarizado con la sintaxis del lenguaje axiomático PAL.

En cuanto a la visualización de los contenidos de las ontologías creadas con *Protégé*, algunos de los *plug-ins* de representación gráfica que pueden incorporarse son *Jambalaya*, *OntoViz* y *TGViz*, que ofrecen la posibilidad de observar el contenido de la ontología utilizando diferentes métodos de representación. A modo de ejemplo, mostramos en la Fig. 28 la representación gráfica del concepto *Columnist* en la ontología *Newspaper* utilizando el *plug-in* *TGViz*.

Con este *plug-in* se puede visualizar una clase seleccionada (*Columnist*). Los *slots* se representan como líneas que vinculan la clase seleccionada con otras clases de la ontología, indicando cuál es la naturaleza de ese vínculo. Por ejemplo, *Columnist* está unido por la relación *sub-super* a la clase *Author*. La función de visualización gráfica puede resultar muy útil para observar las relaciones que se establecen entre las clases contenidas en una ontología.

⁵⁹ El Lenguaje Axiomático PAL facilita la definición de restricciones lógicas a los marcos contenidos en la ontología que se almacenan como parte de la misma.

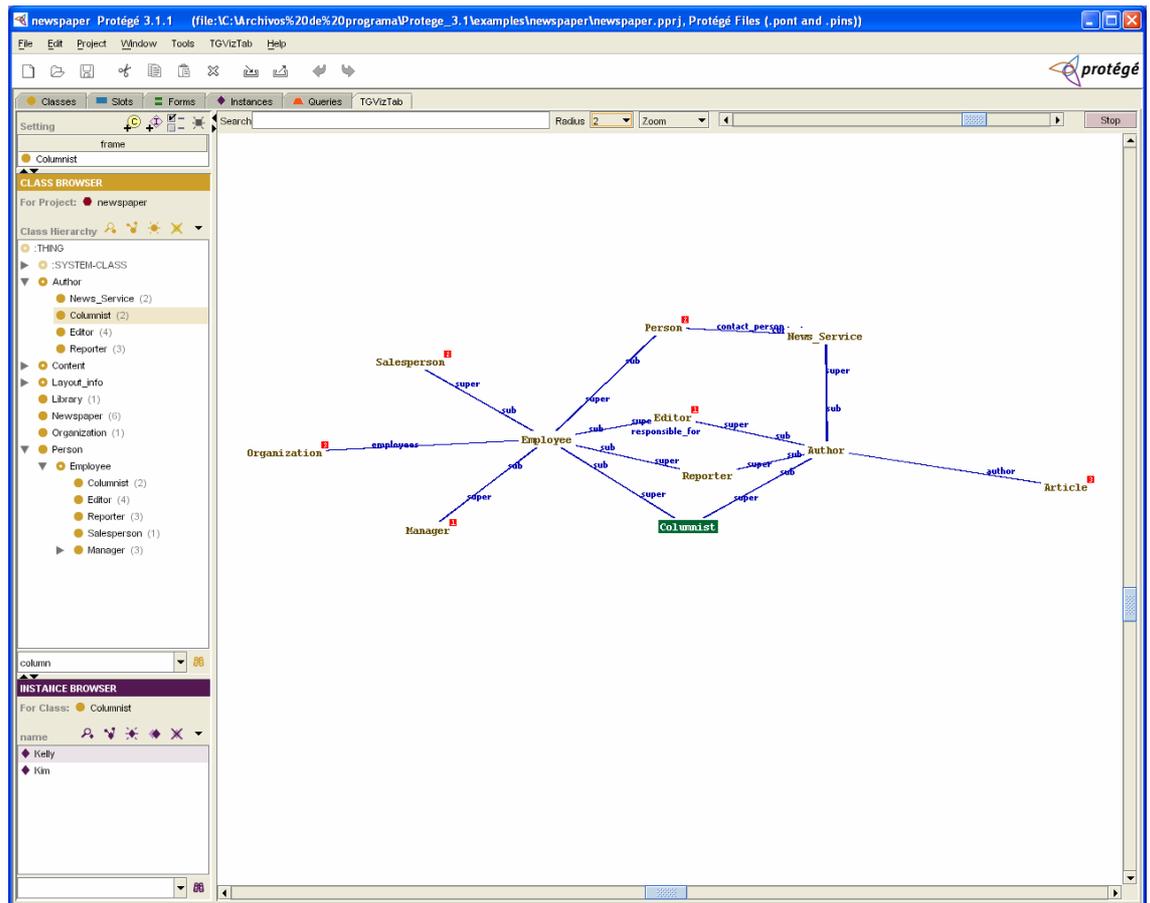


Fig. 28. Representación gráfica de la clase *Columnist* utilizando el *plug-in* TGVizTab.

5.4 Recapitulación

En este capítulo se han presentado de manera sucinta las fases en las que se ha llevado a cabo el trabajo empírico, así como los recursos textuales y las herramientas informáticas empleados. En los capítulos siguientes se explican en detalle los resultados obtenidos a partir de nuestro estudio, haciendo referencia a los recursos y herramientas aquí presentados.

**CAPÍTULO 6: FORMALIZACIÓN E
IMPLEMENTACIÓN DE LAS RELACIONES
CONCEPTUALES EN UN EDITOR DE ONTOLOGÍAS**

6 Formalización e implementación de las relaciones conceptuales en un editor de ontologías

En este capítulo se presenta el catálogo de relaciones conceptuales y nuestras propuestas de formalización e implementación de cada relación conceptual en el editor de ontologías *Protégé*.

En el apartado 6.1 se explica cómo hemos estructurado nuestro catálogo de relaciones a partir de la revisión de las clasificaciones presentadas en los fundamentos teóricos, y se definen las relaciones en lenguaje natural. A continuación se describe nuestra propuesta de formalización, teniendo en cuenta las clases conceptuales y las propiedades de cada relación (apartado 6.2). Asimismo, se presenta la metodología seguida para extraer las relaciones conceptuales a partir de un texto introductorio sobre baldosas cerámicas, y se explica cómo se ha almacenado la información obtenida antes de incorporarla a la base de datos conceptual.

En el apartado 6.3 se explica nuestra propuesta de implementación de las relaciones conceptuales en el editor de ontologías *Protégé*.

En el apartado 6.4 se presenta el catálogo completo de relaciones con su definición en lenguaje natural, su formalización y la propuesta de implementación en el editor *Protégé*. Toda esta información se completa con ejemplos extraídos a partir del análisis del corpus *TXT CERAM*.

Por último, en el apartado 6.5 se resumen los principales resultados de este capítulo y se ofrece un cuadro que resume las relaciones conceptuales del catálogo.

6.1 Catálogo de relaciones conceptuales

En este apartado se presenta el catálogo de las relaciones conceptuales que hemos considerado en la realización de este trabajo. Para determinarlo nos hemos basado en la revisión de las principales clasificaciones de las relaciones propuestas desde la terminología, descritas en detalle en el capítulo 2, que se han completado con aportaciones de la Lingüística y la Psicología.

En primer lugar, se explican las convenciones que hemos seguido para la denominación de las relaciones conceptuales (apartado 6.1.1), y a continuación se

describe cómo se ha estructurado el catálogo, indicando en qué autores nos hemos basado para hacerlo (apartado 6.1.2).

6.1.1 Convenciones para la denominación de las relaciones conceptuales

Uno de los rasgos que distinguen nuestro catálogo de relaciones es el modo en el que se denomina cada relación. Siguiendo el sistema propuesto en la clasificación de las relaciones complejas de Sager (1990, revisada en el apartado 2.2.1.2), cada relación se ha nombrado mediante una expresión que describe la naturaleza de los dos conceptos que vincula, separados mediante un guión. Así, por ejemplo, la relación que en muchas de las propuestas teóricas se denomina relación genérica o de hiponimia, establecida entre un superordinado y sus subordinados, se denomina en nuestra propuesta *relación superordinado-subordinado*. Además de ofrecer un nombre transparente, esta forma de denominar las relaciones deja constancia de la direccionalidad de las mismas, es decir, del sentido en el que se establecen (Oster, 2005: 135-137). De este modo destacamos la importancia que tiene la posición relativa de los elementos vinculados mediante la relación conceptual, lo que más adelante nos permitirá formalizar relaciones inversas.

La principal ventaja de este modo de denominar las relaciones radica en su transparencia. Creemos que una denominación como *relación componente-objeto funcional* resulta mucho más comprensible que hablar de un subtipo de relación meronímica.

Sin embargo, nuestra propuesta de denominación de las relaciones también presenta algunos inconvenientes. Por ejemplo, los nombres pueden resultar más extensos que si se emplea otro tipo de convención. Además, en ocasiones, la búsqueda de la sistematicidad en la forma de denominar las relaciones nos ha llevado a proponer nombres que tal vez resulten menos expresivos que el modo habitual en el que se denomina a la relación correspondiente en la teoría de la Terminología o en la Lingüística. Por ejemplo, al denominar las relaciones secuenciales temporales, hemos optado por la denominación *concepto-concepto simultáneo* para referirnos a la relación que se establece entre dos conceptos que concurren en el tiempo. La denominación *relación simultánea* (propuesta por

Arntz y Picht, 1989) es probablemente más ilustrativa. No obstante, nos hemos decantado por la primera para mantener la coherencia de la propuesta.

Los dos únicos casos en que no se mantienen estas convenciones son las relaciones asociativas y la relación de sinonimia. En el caso de la primera, esto se debe a que esta relación carece de una definición precisa, por lo que no es posible decidir *a priori* qué tipo de conceptos se sitúan a uno y otro lado de la relación. Por ello, en lugar de optar por una denominación vaga como *concepto-concepto*, que sería aplicable a cualquiera de las otras relaciones, hemos decidido mantener la denominación *relación asociativa*. En cuanto a la relación de sinonimia, dado que se trata de una relación entre denominaciones, hemos optado por la nomenclatura habitual, esto es, se ha denominado *relación de sinonimia*.

6.1.2 Estructuración del catálogo de relaciones conceptuales

Como se ha visto en el capítulo 2, la mayoría de los manuales de terminología y las normas internacionales (ISO 704, 2000; Arntz y Picht, 1989; Cabré, 1993, entre otros) establecen una distinción entre relaciones jerárquicas y no jerárquicas. Dentro de las primeras se incluyen las relaciones lógicas (que algunos autores denominan relaciones de abstracción o genéricas) y las relaciones meronímicas (que en ocasiones se llaman partitivas o parte-todo). Dentro de las relaciones no jerárquicas se suelen incluir las relaciones causales, las relaciones asociativas o pragmáticas y las relaciones secuenciales, entre otras.

En nuestro catálogo hemos optado por no mantener la distinción inicial entre relaciones jerárquicas y no jerárquicas, sino por establecer los siguientes cinco grupos: relaciones lógicas, relaciones meronímicas, relaciones secuenciales, relaciones argumentales y circunstanciales y otras relaciones.

Esta estructuración integra las propuestas de Sager (1990), Feliu (2004) y Dancette y L'Homme (2004), que se completan con la subdivisión de las relaciones meronímicas de Winston *et al* (1987).

El principal motivo que nos ha llevado a centrarnos en estas propuestas es que son exhaustivas en cuanto al número de relaciones consideradas. Si bien hemos visto otras clasificaciones, como la de Nuopponen (2005) o la de Weissenhofer (1995), que ofrecen un número de relaciones conceptuales más elevado, consideramos que la mayoría de ellas están incluidas las propuestas anteriores. Dado que nuestro objetivo es la formalización de las relaciones

utilizando un editor de ontologías, un número excesivo de relaciones podría complicar sobremedida la formalización y su implementación, al menos en las primeras fases de desarrollo del proyecto ONTODIC.

Por otro lado, las clasificaciones de Sager, Feliu y Dancette y L'Homme nos permiten identificar las clases conceptuales que participan en cada relación (entidades, actividades y atributos), que, como se verá en el apartado 6.2, serán uno de los aspectos que tendremos en cuenta al afrontar la formalización.

En los siguientes apartados se definen de forma general los cinco grupos de relaciones conceptuales considerados. Al final de cada apartado se ofrece la definición en lenguaje natural de las relaciones incluidas en cada grupo:

1. Relaciones lógicas

Las relaciones lógicas permiten establecer jerarquías o taxonomías entre conceptos en virtud de las características que comparten y que los diferencian. Dentro de las relaciones lógicas hemos considerado, siguiendo a Sager (1990), la distinción entre las que se establecen entre un superordinado y sus subordinados directos (relaciones verticales) y las relaciones que vinculan a los subordinados de un superordinado común (relaciones horizontales). A diferencia de Sager, no hemos tenido en cuenta las relaciones cuasigenéricas, es decir, aquellas en las que la atribución a una clase no es permanente, ya que consideramos que en un trabajo terminológico para un área temática concreta la adscripción de un concepto a una clase es constante.

- a) **Relación superordinado-subordinado:** Relación que se establece entre un concepto **a**, que posee un grado de abstracción superior y un concepto **b**, que se sitúa un nivel por debajo de él en una jerarquía.
- b) **Relación subordinado-subordinado:** Relación que se establece entre dos conceptos que se encuentran situados en el mismo nivel de abstracción por debajo de un superordinado común.

2. Relaciones meronímicas

Junto con las relaciones lógicas, las relaciones meronímicas poseen una gran relevancia en la estructuración del conocimiento especializado. Su naturaleza compleja hace que se hayan identificado varios subtipos, revisados en el apartado 2.4.2.

En nuestro catálogo hemos seguido la clasificación de relaciones meronímicas de Winston *et al* (1987), que identifica siete tipos de relación meronímica en función de cuatro rasgos: funcional (la posición espacial o temporal de las partes determina su función dentro del todo), homeomérica (la naturaleza de la parte es la misma que la del todo al que pertenece), separable (las partes y el todo pueden estar conectadas o no) y simultánea (la mayoría de las partes pertenecen al todo al mismo tiempo). Los siete tipos de relaciones meronímicas a que dan lugar estos rasgos son los siguientes: componente funcional-objeto, miembro-colección, porción-masa, material-objeto, etapa-proceso, característica-actividad y lugar-área.

En nuestro catálogo hemos añadido una octava relación, la relación parte-parte, que relaciona las partes que juntas constituyen un todo, con independencia de la relación meronímica que las vincule con él.

- a) **Relación componente funcional-objeto:** Relación que se establece entre los componentes funcionales de un objeto y el objeto del que forman parte.
- b) **Relación miembro-colección:** Relación que se establece entre partes que establecen una relación de proximidad espacial o conexión social con respecto al todo, sin necesidad de que cumplan una función específica ni que estén dispuestos estructuralmente de un modo específico.
- c) **Relación porción-masa:** Relación entre una parte y el todo donde la naturaleza de la porción es la misma que la de la masa.
- d) **Relación material-objeto:** Relación entre una parte intrínseca y el todo al que pertenece, del que no puede separarse por ser consustancial a él.
- e) **Relación etapa-proceso:** Relación que se establece entre cada una de las fases de una actividad y dicha actividad en conjunto.
- f) **Relación característica-actividad:** Relación que se establece entre las características de una actividad y el proceso complejo en el que se inscriben.

- g) **Relación lugar-área:** Relación que se establece entre un área espacial y las localizaciones concretas dentro de esa área.
- h) **Relación parte-parte:** Relación que se establece entre las partes que juntas componen un todo.

3. Relaciones secuenciales

Las relaciones secuenciales presentadas en nuestro catálogo están basadas en la propuesta de Feliu (2004). Son aquellas que se establecen por localización o sucesión en el espacio o en el tiempo de los elementos que vinculan. Feliu (2004) identifica dos tipos de secuencialidad: secuencialidad espacial y secuencialidad temporal. La primera, a su vez, puede ser de localización o de dirección, y la secuencialidad temporal puede ser de simultaneidad y de anterioridad-posterioridad.

- a) **Relación concepto-lugar:** Relación que se establece entre dos conceptos que se ubican juntos en el espacio.
- b) **Relación concepto-lugar al que se dirige:** Relación que se establece entre dos conceptos para indicar la dirección en la que se produce una actividad determinada.
- c) **Relación concepto-concepto simultáneo:** Relación que se establece entre dos actividades o entidades que se ubican en el mismo espacio de tiempo.
- d) **Relación concepto anterior-concepto posterior:** Relación que se establece entre dos conceptos que se sitúan en el tiempo de forma consecutiva.

4. Relaciones argumentales y circunstanciales

Esta agrupación de las relaciones argumentales y circunstanciales se basa en la propuesta de Dancette y L'Homme (2004) presentada en el apartado 3.2.5, que a su vez coincide en parte con las relaciones que Sager (1990) denomina *relaciones complejas*. Bajo este epígrafe agrupamos las relaciones conceptuales paradigmáticas que se establecen entre los predicados y sus argumentos (relaciones argumentales), así como aquellas que indican las circunstancias en las que se produce un predicado (relaciones circunstanciales).

Las relaciones argumentales y circunstanciales incluidas en nuestro catálogo son las siguientes: proceso-agente, proceso-producto, proceso-paciente, proceso-instrumento, proceso-estado, proceso-método, causa-efecto y objeto-uso.

- a) **Relación proceso-agente:** Relación que se establece entre un proceso y la entidad que lleva a cabo dicho proceso.
- b) **Relación proceso-producto:** Relación que se establece entre el proceso y el producto final de dicho proceso.
- c) **Relación proceso-paciente:** Relación que se establece entre un proceso y la entidad sobre la que se lleva a cabo.
- d) **Relación proceso-instrumento:** Relación entre un proceso y el instrumento empleado para llevarlo a cabo.
- e) **Relación proceso-estado:** Relación entre un proceso y el estado final que alcanza el objeto sobre el que se produce.
- f) **Relación proceso-método:** Relación que se establece entre un proceso y el método que se emplea para llevarlo a cabo.
- g) **Relación causa-efecto:** Relación que se establece entre una causa y el efecto que produce.
- h) **Relación objeto-uso:** Relación entre un objeto y la función o el uso para el que está destinado.

5. Otras relaciones

Bajo esta denominación hemos agrupado algunas relaciones que no encajan dentro de los cuatro grupos anteriores, pero que, no obstante, pueden establecerse entre los conceptos de la industria cerámica. Consideramos que su representación también resulta útil para la estructuración del conocimiento especializado de este campo del saber.

Incluimos en este grupo las relaciones siguientes: fenómeno-medida, objeto-característica, relación asociativa y la relación de sinonimia. Las dos primeras proceden de la clasificación de relaciones complejas propuesta por Sager

(1990). Por su parte, la relación de sinonimia no es estrictamente una relación entre conceptos distintos, sino una relación entre denominaciones. Por último, la relación asociativa permite reflejar aquellas relaciones conceptuales de naturaleza indeterminada.

- a) **Relación fenómeno-medida:** Relación que se establece entre un fenómeno o una característica y la unidad de medida empleada para expresar dicho fenómeno o característica.
- b) **Relación objeto-característica:** Relación que se establece entre una entidad y una característica que lo define.
- c) **Relación asociativa:** Relación que se establece por la correlación entre dos o más conceptos cuya naturaleza no se corresponde con ninguno de los tipos de relación identificados anteriormente.
- d) **Relación de sinonimia:** Relación que se establece por equivalencia entre dos o más denominaciones que se refieren al mismo concepto.

A continuación se presentan de forma esquemática estos cinco grupos de relaciones conceptuales, que se definirán una a una en la sección 6.4. Como veremos en el apartado 6.2, algunas de estas relaciones poseen una relación inversa que también se cumple:

- 1. *Relaciones lógicas*
 - a. Superordinado-subordinado
 - b. Subordinado-subordinado
- 2. *Relaciones meronímicas*
 - a. Componente funcional-objeto
 - b. Miembro-colección
 - c. Porción-Masa
 - d. Material-objeto
 - e. Etapa-proceso
 - f. Característica-actividad
 - g. Lugar-área
 - h. Parte-parte
- 3. *Relaciones secuenciales*
 - a. Concepto-lugar

- b. Concepto-lugar al que se dirige
- c. Concepto-concepto simultáneo
- d. Concepto anterior-concepto posterior

4. *Relaciones argumentales y circunstanciales*

- a. Proceso-agente
- b. Proceso-paciente
- c. Proceso-producto
- d. Proceso-estado
- e. Causa-efecto
- f. Proceso-instrumento
- g. Proceso-método
- h. Objeto-uso

5. *Otras relaciones*

- a. Fenómeno-medida
- b. Objeto-característica
- c. Relación de sinonimia
- d. Relación asociativa

6.2 Formalización de las relaciones conceptuales

Para poder representar formalmente las relaciones conceptuales, una vez definido el catálogo es necesario establecer algún tipo de formalismo que nos permita estructurar la información sobre las relaciones conceptuales. La formalización es, por tanto, un paso previo a la implementación de las relaciones conceptuales en el editor, ya que sólo si la información conceptual recogida sobre las mismas está estructurada formalmente podremos almacenar y recuperarla de forma onomasiológica por medios informáticos.

En esta sección presentamos, en primer lugar, la definición formal de relación conceptual y las propiedades de las relaciones conceptuales (6.2.1). A continuación explicamos nuestra concepción de las clases conceptuales a las que pueden pertenecer los conceptos vinculados mediante una determinada relación (6.2.2). En el apartado 6.2.3 se presenta el modo en el que se ha llevado a cabo la estructuración de la información sobre las relaciones a partir de los datos obtenidos durante el estudio de los conceptos sobre los productos cerámicos en un texto introductorio sobre baldosas cerámicas, completado con la información

extraída a partir de los diccionarios y la base de datos terminológica *Cerámica* y con las concordancias extraídas a partir del corpus *TXTCERAM* para la identificación de relaciones.

6.2.1 Definición formal y propiedades de las relaciones conceptuales

En primer lugar, presentamos el formalismo que nos sirve para describir las relaciones conceptuales. En este trabajo entendemos las relaciones conceptuales como vínculos semánticos establecidos entre dos o más conceptos especializados. Nuestra definición, basada en la de Otman (1996, presentada en el apartado 3.2.2) y en la reelaboración de su fórmula propuesta por Feliu (2004, presentada en 2.2.2.1), se expresa mediante el formalismo:

$$a \mathbf{R} b, n$$

donde *a* es un concepto perteneciente a una clase conceptual de las definidas por Sager y Kageura (1994), que –como veremos en el siguiente apartado– pueden ser entidades, propiedades o actividades. La variable *b* de la fórmula representa el/los concepto/s relacionado/s con *a*, y **R** representa la relación conceptual que se establece entre *a* y *b*. La variable *b* también puede ser una entidad, un atributo o una actividad. La constante *n* nos permite indicar que algunos conceptos pueden establecer una relación determinada con más de un concepto *b* simultáneamente.

El concepto *a* suele denominarse dominio de la relación, mientras que el o los conceptos *b* suelen denominarse rango de la relación.

En el apartado 2.3 se han definido las propiedades lógicas o matemáticas de las relaciones. A la hora de afrontar la formalización de éstas en nuestro catálogo, hemos tenido en cuenta algunas de estas propiedades que caracterizan cada relación. El hecho de considerar las propiedades de las relaciones en su formalización y posterior implementación nos permitirá, en fases posteriores de la investigación, llevar a cabo razonamientos e inferencias sobre las relaciones que nos permitan aprovechar fenómenos como el de la herencia.

Las propiedades más relevantes para nuestro trabajo son la transitividad (y sus implicaciones para el sistema de herencia), la simetría, la recursividad y la conversividad (relaciones inversas). A continuación recordamos brevemente qué entendemos por cada una de ellas, aportando ejemplos procedentes de nuestro estudio empírico sobre los conceptos cerámicos:

- a) **Transitividad:** Se dice que una relación es transitiva si se cumple que si $a R b$ y $b R c$, entonces $a R c$. Por ejemplo, la relación que se establece entre un superordinado y un subordinado es transitiva. Así, si consideramos que el concepto *baldosa cerámica* es un subordinado del concepto *producto cerámico acabado*, y a su vez *baldosa esmaltada* es un subordinado de *baldosa cerámica*, entonces el concepto *baldosa esmaltada* está subordinado al concepto *producto cerámico acabado*. La transitividad nos permite explotar los mecanismos de herencia en la creación de bases de datos conceptuales empleando un editor de ontologías.
- b) **Simetría:** Se dice que una relación es simétrica si y solo si, cuando se cumple $a R b$, entonces se cumple la misma relación en sentido contrario $b R a$. Por ejemplo, la relación entre las distintas partes que componen un todo es simétrica. Así, si los conceptos *cara vista* y *cara posterior* son ambos partes del mismo todo (*baldosa cerámica*), establecen entre sí la relación *parte-parte*, que se cumple en ambos sentidos.
- c) **Recursividad:** Una relación es recursiva si puede establecerse entre un concepto a y n conceptos b , siendo n mayor o igual que dos. Esta propiedad hace referencia al número de argumentos que puede aparecer a ambos lados de la relación. Por ejemplo, la relación *material-objeto* es una relación que puede presentar recursividad, es decir, un mismo material puede formar parte de varios objetos, como por ejemplo la *arcilla* se emplea tanto en las *baldosas cerámicas* como en el *baldosín catalán*.
- d) **Relaciones inversas:** Se dice que dos relaciones son inversas entre sí, si para $a R b$, se cumple una relación R' entre b y a que posee las mismas propiedades matemáticas ($b R' a$). Por ejemplo, la relación *proceso-producto* posee una relación inversa denominada *producto-proceso*, y por tanto, los conceptos *moldeo por extrusión* y *baldosa extrudida* se encuentran vinculados mediante ambas relaciones. Esta propiedad nos permite tener en cuenta la direccionalidad de las relaciones.

Estas propiedades matemáticas se aplican a todas las relaciones conceptuales, y nos permitirán definir las formalmente.

En el caso de las relaciones meronímicas, además de considerar las propiedades matemáticas que acabamos de ver, se han tenido en cuenta los rasgos especiales que distinguen los subtipos de meronimia, definidos en detalle en el apartado 2.4.2. A continuación retomamos brevemente cada una de las propiedades específicas de las relaciones meronímicas, aportando ejemplos procedentes de nuestro análisis de los conceptos cerámicos en caso de que se hayan encontrado:

- a) **Funcional:** La relación entre la parte y el todo es funcional cuando las partes están en una posición espacial o temporal específica que determina su función dentro del todo. La parte posee una función especial con respecto al todo, como por ejemplo las *aristas* con respecto a las *baldosas cerámicas*.
- b) **Homeomérica:** La relación entre la parte y el todo es homeomérica cuando la naturaleza de la parte es la misma que la del resto de las partes y la del todo al que pertenece. Por ejemplo, una porción de pizza sigue siendo pizza, mientras que un árbol no es lo mismo que un bosque. De momento no hemos encontrado ejemplos de relaciones homeoméricas en el ámbito de la cerámica.
- c) **Separable:** Esta propiedad hace referencia a que las partes y el todo pueden estar físicamente conectadas o no. Por ejemplo, el *esmalte* es separable con respecto a la *baldosa cerámica*, esto es, aunque es parte de la baldosa, el esmalte puede separarse de ella.
- d) **Simultánea:** Este rasgo de las relaciones meronímicas se añade en Chaffin *et al* (1988). Las partes simultáneas pertenecen a objetos que poseen la mayor parte de sus partes al mismo tiempo, como por ejemplo *arcilla*, que es uno de los materiales que componen las *baldosas cerámicas*. La *arcilla* aparece simultánea e inseparablemente ligada al todo del que forma parte, es decir, a la *baldosa cerámica*.

6.2.2 Clases de conceptos

Tal y como apuntábamos al presentar el formalismo que representa las relaciones conceptuales, los conceptos vinculados mediante una relación se adscriben a una clase. En el apartado 2.2.1.1 se han revisado los modelos conceptuales estático y dinámico propuestos por Sager y Kageura (1994). Estos autores proponen observar y definir las distintas clases de conceptos, ya que, en su opinión, la naturaleza de los conceptos permite determinar las estructuras conceptuales que se establecen en los campos especializados.

En nuestra investigación compartimos este punto de vista, y por ello hemos considerado que, para describir las relaciones de una manera formal, es necesario tener en cuenta qué clases de conceptos participan en ellas. El modelo conceptual propuesto por Sager y Kageura nos parece adecuado por su orientación terminológica, y porque en él se define de manera clara qué se entiende por cada una de las clases de conceptos que podemos encontrar en las áreas de especialidad.

Retomemos pues la tipología de las clases de conceptos propuesta por Sager y Kageura (1994), y veamos qué elementos de dicha tipología se han utilizado en nuestro trabajo empírico para la formalización de relaciones, junto con algún ejemplo procedente del estudio empírico. Según estos autores, existen cuatro clases de conceptos: entidades, relaciones, actividades y propiedades:

- a) **Entidad:** Clase de conceptos que se obtiene a partir de la abstracción de los elementos de la experiencia y de la reflexión, cuya existencia se considera independiente en el espacio y en el tiempo. Las entidades son elementos definibles por separado, necesarios para la identificación y clasificación de las unidades de la experiencia y el conocimiento. Ejemplos en el ámbito de la cerámica serían todos los productos cerámicos (*azulejo, baldosa*), las materias primas (*arcilla, gres*), la maquinaria y sus piezas (*prensa oleodinámica, horno monoestrato de rodillos*), así como los lugares en los que se desarrollan los procesos (*secadero*) y en los que se colocan los productos (*pared, suelo*).
- b) **Actividad:** Clase de conceptos obtenida mediante la abstracción de procesos, operaciones o acontecimientos realizados por o con

entidades. Su estructura es más compleja que la de las entidades, puesto que requieren la participación directa de estas para llevarse a cabo. En el ámbito de la cerámica, algunas actividades identificadas son los procesos de fabricación (*prensado en seco, cocción*) y los procesos de colocación (*embaldosado, colocación en capa fina*).

- c) **Propiedad:** Clase de conceptos derivada del análisis de los componentes y de las características de entidades, actividades y relaciones. Las propiedades siempre se consideran en un primer nivel de abstracción asociadas a otros conceptos, y sólo en un segundo nivel se constituyen como conceptos independientes. Permiten identificar y describir las diferencias entre entidades y actividades, así como reflejar rasgos o características de las mismas. Algunos ejemplos propios de nuestra área temática serían todas las características de los productos cerámicos (*resistencia a la heladicidad, porosidad, color*).
- d) **Relación:** Clase de conceptos obtenida a partir de la abstracción de relaciones físicas, temporales u otro tipo de relaciones ontológicas entre objetos y a partir de las relaciones lógicas establecidas entre entidades, relaciones y actividades. Las relaciones son la clase de conceptos que identifica los vínculos existentes o establecidos entre dos o más entidades, actividades o propiedades, o bien entre una combinación de ellos. Algunos ejemplos de conceptos de relación que hemos identificado para el ámbito de la cerámica son *fase* (señala una relación secuencial) o *composición* (señala una relación meronímica).

6.2.3 Análisis de las relaciones conceptuales en el ámbito de la cerámica: organización de la información conceptual extraída

Una vez configurado el catálogo de relaciones conceptuales a partir de la revisión de las propuestas teóricas, y decidido el modelo de formalización de las mismas, hemos desarrollado un primer estudio de las relaciones que se establecen entre los conceptos en el ámbito de la cerámica a partir de un texto especializado, con el fin de comprobar la validez de nuestra propuesta.

Este primer análisis se ha realizado a partir de un texto introductorio sobre baldosas cerámicas (la *Guía de la baldosa cerámica*, descrita en el apartado 5.2.1) y se ha completado con el estudio de las definiciones de los conceptos extraídos en los dos diccionarios especializados descritos en los apartados 5.2.2 y 5.2.3 (Guillem y Guillem, 1987 y McColm, 1994), y la información extraída a partir de los contextos lingüísticos y definatorios incluidos la base de datos terminológica *Cerámica*, descrita en el apartado 5.2.4.

En un primer momento se planteó la posibilidad de emplear como punto de partida para la extracción de información conceptual el corpus completo de textos sobre la industria cerámica *TXTCERAM* (descrito en el apartado 5.2.6), y explotarlo a partir de los términos que aparecen en la base de datos terminológicos *Cerámica* con ayuda de la herramienta de concordancias de *WordSmith Tools* (presentada en el apartado 5.3.2). Sin embargo, los primeros análisis nos hicieron darnos cuenta de que la información conceptual sobre relaciones en las concordancias de los términos que aparecían en el corpus se encontraba muy dispersa. En efecto, en ocasiones encontramos más de mil contextos para un término concreto. Algunos de ellos aportan información sobre relaciones conceptuales, pero es necesario revisarlos uno por uno para saber si contienen o no este tipo de información. Además, no resulta fácil discernir si la información aportada por un contexto es completa. Por ejemplo, para el concepto *baldosa cerámica* se recuperaron 903 concordancias en el corpus *TXTCERAM*. Si analizamos a título ilustrativo las 50 primeras encontradas, podemos hacer las siguientes observaciones:

- a) En la mayoría de los contextos aparecen conceptos relacionados con *baldosa cerámica*, pero la naturaleza de la relación que se establece entre ellos no suele hacerse explícita en el contexto.
- b) En algunos contextos se repite la información ya recuperada, y en otros (tres de cincuenta en nuestra muestra) no aparece ningún concepto relacionado.
- c) Los conceptos relacionados con *baldosa cerámica* recuperados hacen referencia a distintos tipos de baldosa cerámica (por ejemplo *baldosa cerámica extrudida*, *baldosa cerámica prensada en seco*, *gres porcelánico*), a las características técnicas y dimensionales de las baldosas (por ejemplo *dureza Mohs*, *resistencia a la abrasión*, *resistencia química*, *tamaño*,

forma) y al sistema y lugar de colocación de las baldosas (*sistema de colocación, colocación en capa gruesa, capa de agarre, alicatado*). Sin embargo, a la vista de estos cincuenta primeros contextos observados, no podemos determinar hasta qué punto la información recuperada es representativa del total de los conceptos que establecen algún tipo de relación con *baldosa cerámica*.

Para superar esta dispersión de la información sobre las relaciones en el corpus, podríamos haber utilizado como punto de partida para la identificación de conceptos relacionados en el corpus patrones de conocimiento (Meyer, 2001) o sondas de conocimiento (Ahmad y Fulford, 1992), que expresan la forma lingüística que posee cada tipo de relación. Dado que el número de relaciones conceptuales con las que trabajamos es bastante elevado (26 relaciones), y teniendo en cuenta que aún no disponemos de los patrones para las relaciones conceptuales consideradas, de momento la identificación de conceptos relacionados se ha hecho de forma manual. No obstante, en el marco de los proyectos TXTCERAM y ONTODIC está previsto considerar los patrones lingüísticos para la extracción de información conceptual⁶⁰.

Por todo ello, para comprobar la validez de nuestra propuesta en el ámbito de la cerámica hemos partido de un texto introductorio sobre baldosas cerámicas, en lugar de utilizar el corpus textual.

El estudio de las relaciones conceptuales establecidas por los conceptos identificados en la *Guía de la baldosa cerámica* se llevó a cabo del siguiente modo:

En primer lugar, mediante un proceso manual se extrajeron los conceptos y sus conceptos relacionados encontrados en la *Guía de la baldosa cerámica*. Una vez identificado cada concepto y sus conceptos relacionados, se comprobó si aparecía en la base de datos terminológicos *Cerámica* y en los dos diccionarios especializados (Guillem y Guillem, 1987 y McColm, 1994). En los diccionarios y la base terminológica se analizaron los contextos definitorios y lingüísticos de la base terminológica, lo que nos permitió extraer nuevos conceptos relacionados.

⁶⁰ En la actualidad, una de las colaboradoras en el proyecto TXTCERAM está estudiando los patrones de extracción de relaciones meronímicas en el corpus de la cerámica (Soler, 2005).

Con la información recogida en el texto, completada con la procedente de las fuentes lexicográficas y terminológicas, se ha elaborado una tabla en Excel⁶¹ en la que se organizó la información del siguiente modo. En la primera columna de la tabla se han situado los conceptos identificados en el texto. En la segunda columna se indica la clase a la que pertenece el concepto cuyas relaciones estamos estudiando (entidad, actividad o propiedad). Posteriormente se ha generado una columna para cada una de las relaciones del catálogo. Es decir, para cada concepto, en la primera columna situamos el concepto identificado; en la segunda indicamos si se trata de una entidad, una actividad o una propiedad, y en las columnas subsiguientes situamos los conceptos relacionados por debajo de la denominación de la relación conceptual que los vincula con el concepto objeto de estudio.

Este modo de organizar la información permite reflejar en forma de tabla las relaciones binarias que establece cada concepto. Así, por ejemplo, la relación existente entre la entidad *baldosa cerámica* y su concepto subordinado *baldosa cerámica extrudida* se refleja en la columna correspondiente a la relación superordinado-subordinado. La tabla nos permite recoger todos los conceptos relacionados y observar la naturaleza de las relaciones que establece el concepto, que se podría representar mediante el siguiente esquema:

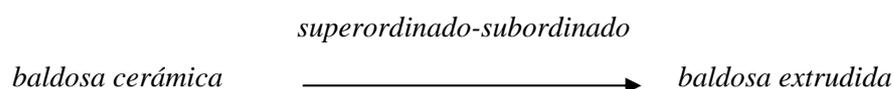


Fig. 29. Esquema de la relación conceptual entre *baldosa cerámica* y *baldosa extrudida*.

Para ejemplificar el modo en el que se ha elaborado la tabla, ofrecemos en la Tabla 19 un extracto de la misma que recoge los conceptos relacionados con el concepto *baldosa cerámica* extraídos a partir de nuestro estudio empírico. Como puede verse, el número de conceptos relacionados con el concepto *baldosa cerámica* identificados es elevado, y todos los conceptos relacionados se pueden adscribir a alguno de los tipos de relación propuestos en nuestro catálogo.

⁶¹ La tabla se ha elaborado utilizando el programa Microsoft Excel porque facilita la creación de tablas de grandes dimensiones.

<i>concepto/tipo de relación</i>	<i>clase de conceptos</i>	<i>superordinado-subordinado</i>	<i>subordinado-subordinado</i>	<i>componente funcional-objeto</i>	<i>miembro-colección</i>	<i>porción-masa</i>	<i>material-objeto</i>	<i>etapa-proceso</i>
<i>baldosa cerámica</i>	<i>Entidad</i>	<i>baldosa extrudida; baldosa prensada en seco; baldosa prensada; baldosa porosa; baldosa ligeramente porosa; baldosa no porosa; baldosa esmaltada; baldosa no esmaltada; baldosa engobada; azulejo; pavimento de gres; gres porcelánico; baldosín catalán; gres rústico; barro cocido mosaico; losa; placa para pavimentos industriales; componente para piscinas; accesorio</i>						

<i>concepto/tipo de relación</i>	<i>actividad-característica</i>	<i>área-lugar</i>	<i>parte-parte</i>	<i>concepto-lugar</i>	<i>concepto-lugar al que se dirige</i>	<i>concepto-concepto simultáneo</i>	<i>concepto anterior-concepto posterior</i>	<i>proceso-agente</i>	<i>proceso-paciente</i>
<i>baldosa cerámica</i>				<i>suelo; pared</i>		<i>material de agarre</i>			

<i>concepto/tipo de relación</i>	<i>producto-proceso</i>	<i>proceso-estado</i>	<i>causa-efecto</i>	<i>proceso-instrumento</i>	<i>proceso-método</i>	<i>objeto-uso</i>	<i>fenómeno-medida</i>	<i>objeto-característica</i>	<i>relación de sinonimia</i>	<i>relación asociativa</i>
<i>baldosa cerámica</i>	<i>moldeo; moldeo por extrusión; moldeo por prensado; colado; molienda; amasado; secado</i>					<i>revestimiento de paredes; revestimiento de fachadas; revestimiento de suelos</i>		<i>porosidad; absorción de agua; velocidad de succión; resistencia mecánica; impermeabilidad; incombustibilidad; inalterabilidad a la luz; resistencia a la abrasión; rectitud de lados; ortogonalidad; planitud de superficie; carga de rotura; resistencia química; resistencia al deslizamiento; resistencia a la helada; longitud; anchura; grosor; forma</i>	<i>baldosa; baldosín</i>	<i>pieza complementaria</i>

Tabla 19. Conceptos relacionados con el concepto *baldosa cerámica* recogidos en el estudio preliminar.

El modo de organizar la información sobre los conceptos y sus conceptos relacionados mediante esta tabla permite ver en cada fila un concepto junto con todos los conceptos relacionados con él. Además, si observamos la tabla columna a columna, podremos ver de un solo vistazo cuáles son las relaciones más frecuentes entre los conceptos extraídos.

Durante el análisis de los conceptos también hemos observado algunas relaciones inversas. Así, en el caso concreto de *baldosa cerámica*, no hemos identificado ejemplos de la relación *componente funcional-objeto* como tal, pero observamos que el concepto *baldosa cerámica* establece una relación *objeto-componente funcional* con las partes que la constituyen. A partir de este estudio hemos identificado algunos de los componentes funcionales de las baldosas, como *relieve*, *cara posterior* o *cara vista*, que también se recogen en la tabla de Excel en la fila correspondiente a estos conceptos. Lo mismo ocurre con los materiales con los que se pueden fabricar las baldosas, que se vinculan mediante la relación *material-objeto*. Si invertimos la direccionalidad de la relación, podemos decir que *baldosa cerámica* está vinculado con conceptos como *arcilla* o *sílice* mediante la relación *objeto-material*. Estas relaciones se hacen visibles en la tabla de Excel en la fila correspondiente a estos conceptos.

El estudio nos permite concluir que la mayoría de las relaciones puede presentar recursividad, es decir, un concepto puede estar vinculado con varios conceptos mediante el mismo tipo de relación. Por ejemplo, la relación *objeto-característica* que establece *baldosa cerámica* es recursiva, puesto que este objeto puede poseer (y de hecho posee) más de una característica (*resistencia a la helada*, *resistencia química*, etc.).

En la descripción detallada del catálogo de relaciones que se presenta en el apartado 6.4 se indican las clases conceptuales que pueden participar en cada una de las relaciones y cuáles son sus propiedades (transitividad, simetría, recursividad y relaciones inversas). En la propuesta de implementación de las relaciones conceptuales en el editor de ontologías *Protégé* que se presenta en el siguiente apartado, tendremos en cuenta todos estos elementos para la formalización de las relaciones, que nos permitirán estructurar en el editor de ontologías cada relación conceptual.

6.3 Implementación en el editor de ontologías

Después de haber establecido el catálogo de relaciones, y una vez descritas éstas formalmente mediante las clases conceptuales que participan en ellas y sus propiedades, en este apartado presentamos una propuesta de implementación de las relaciones conceptuales en el editor de ontologías *Protégé*. Para llegar a ella hemos tenido en cuenta las características y funciones del editor descritas en los apartados 4.4.4 y 5.3.3.

En el apartado 6.3.1 se expone el modo en que nos hemos servido de los componentes del editor de ontologías (*classes*, *slots* e *instances*) para estructurar los conceptos, denominaciones y relaciones establecidas en la base de conocimiento, a la que hemos llamado *Ontoceram*. Como se verá en el capítulo 7, en *Ontoceram* almacenaremos de manera estructurada los conceptos pertenecientes a la rama de los productos cerámicos acabados, las relaciones conceptuales que establecen dichos conceptos y los propios conceptos relacionados, así como las denominaciones que se emplean en el ámbito de la cerámica para referirse a ellos.

En el apartado 6.3.2 se describe el cuadro que utilizaremos en el apartado 6.4 para representar cómo se ha implementado cada una de las relaciones conceptuales dependiendo de su formalización.

6.3.1 Implementación de los conceptos, denominaciones y relaciones conceptuales en *Protégé*

Una vez definido el modelo de formalización de las relaciones conceptuales, el siguiente paso en nuestra investigación ha sido preguntarnos cómo introducir la información formalizada en la herramienta elegida para la elaboración de nuestra base de conocimiento, el editor de ontologías *Protégé*. Esta propuesta de introducción de los datos es lo que en informática se denomina *implementación*, es decir, una puesta en funcionamiento del modelo propuesto en una herramienta informática concreta.

Teniendo en cuenta las características y funciones del editor de ontologías *Protégé*, proponemos que se representen los distintos tipos de información conceptual en nuestra base de datos del modo que se describe en los siguientes subapartados:

6.3.1.1 Implementación de los conceptos

Los conceptos se implementan mediante el componente *Class* en el editor de ontologías *Protégé*. En la aplicación de las ontologías en la ingeniería del conocimiento, las clases se conciben como una representación abstracta de un concepto en un dominio. En terminología, los conceptos se definen como una representación mental de los objetos en un dominio de especialidad. Por ello creemos que las clases son un componente idóneo para representar los conceptos terminológicos.

Las clases se nombran en español y se escriben con mayúscula inicial. Por ejemplo, el concepto *Baldosa cerámica* se introduce en la jerarquía de clases con mayúscula inicial.

6.3.1.2 Implementación de las denominaciones

Las denominaciones que se emplean para referirse a los conceptos en el área de especialidad de la cerámica se implementan en el editor de ontologías *Protégé* por medio del componente *Instances*. En la ingeniería del conocimiento este componente se suele emplear para introducir los objetos a los que se refieren las clases concretas. La implementación de las denominaciones como instancias se utilizará en nuestra propuesta para reflejar las relaciones de sinonimia, y servirá para realizar consultas en la base de conocimiento, tal y como veremos en el capítulo 7.

Las denominaciones se escriben con minúscula, a no ser que en el corpus aparezcan siempre con mayúscula porque se trate de nombres de marcas o productos concretos, o porque se tomen prestados de otras lenguas como el alemán (como por ejemplo *Spaltplatten*). De este modo distinguimos los conceptos implementados como *Classes* de las denominaciones. Por ejemplo, el concepto *Baldosa cerámica* se denomina *baldosa cerámica*, *baldosa*, *baldosín* y *rasilla*. Todas estas denominaciones se recogen en *Ontoceram* mediante el componente *Instance* de la *Classes*.

6.3.1.3 Implementación de las relaciones conceptuales

Las relaciones conceptuales, definidas como vínculos semánticos que se establecen entre los conceptos de un dominio, pueden representarse en la base de conocimiento por distintos medios, dependiendo del tipo de relación de que se trate. A continuación presentamos nuestra propuesta de implementación de las diversas relaciones conceptuales en el editor de ontologías:

a) Implementación de la relación superordinado-subordinado y su inversa

Los conceptos, representados en nuestra base de conocimiento a través del componente *Class*, se organizan en *Protégé* como una taxonomía en la que se mantiene una relación de tipo IS-A, es decir, una relación hiponímica. Esta estructuración jerárquica de los conceptos permite explotar en la ontología el mecanismo de la herencia. La relación superordinado-subordinado y su inversa, la relación subordinado-superordinado, se hacen explícitas mediante la estructuración jerárquica de las clases en el propio editor, es decir, los superordinados y subordinados de cada concepto se determinan por el lugar que ocupan en la jerarquía. Así, por ejemplo, el concepto representado en *Protégé* por la clase *Baldosa cerámica* depende en la jerarquía de clases del concepto *Producto cerámico acabado*, que es el superordinado del que hereda sus propiedades.

b) Implementación del resto de las relaciones conceptuales

Proponemos que se implemente el resto de las relaciones conceptuales del catálogo en el editor de ontologías por medio de los *slots*, que es el componente de *Protégé* que permite asignar atributos a las clases. Estos atributos pueden poseer distintos valores, uno de los cuales pueden ser otras clases almacenadas en *Protégé*. Por tanto, se han creado *slots* de relación para establecer vínculos entre los conceptos representados por clases.

Los *slots* de relación se representan en español con minúscula inicial. Su nombre consta de dos elementos unidos por un guión, y coincide con el de las relaciones consideradas en el catálogo descrito en el apartado 6.1. El primer elemento del nombre del *slot* hace referencia al dominio de la relación, es decir, al concepto situado a la izquierda del formalismo $a R b$, mientras que el segundo elemento hace referencia al rango de la misma, es decir, al concepto o conceptos

con los que se relaciona el primero, y que se sitúa/n a la derecha en el formalismo. Un ejemplo de *slot* de relación es *objeto-característica*, que vincula objetos materiales (como los productos acabados) con sus características (como por ejemplo la *resistencia a la heladicidad*). Por tanto, el *slot* de relación *objeto-característica* puede emplearse para hacer explícita la relación que se establece entre un producto acabado (objeto) y la resistencia a la heladicidad, que es una de sus características.

Los *slots* de relación se implementan en la base de datos conceptual mediante la definición de sus distintas facetas. Las facetas definidas por defecto para los *slots* estándar (descritas en el apartado 5.3.3) nos permiten incorporar información sobre cada relación conceptual. Recordemos brevemente cuáles son estas facetas estándar. En cada una de ellas, indicaremos cómo se ha utilizado en nuestra implementación de las relaciones conceptuales:

- a) **Name:** Nombre que recibe el *slot*. En nuestra implementación esta faceta se emplea para asignar un nombre a la relación conceptual.
- b) **Documentation:** Descripción del *slot* en lenguaje natural. En nuestra implementación esta faceta se puede emplear para describir el significado de la relación.
- c) **Template Values:** Valores de plantilla del *slot*. Esta faceta no se ha utilizado en nuestra implementación de las relaciones, puesto que los valores en ella asignados no pueden modificarse posteriormente.
- d) **Default Values:** Valores por defecto del *slot*. Esta faceta es opcional, y en nuestra base de datos conceptual puede emplearse para señalar los valores por defecto que puede tener una relación, que más adelante pueden ser modificados para una clase o instancia concreta.
- e) **Value Type:** Tipo de valor que puede tener el *slot*. En los *slots* de relación, el tipo de valor siempre va a ser *Class*, puesto que ha de hacer referencia a otro concepto (representado mediante una clase) contenido en la base de datos conceptual.
- f) **Allowed Superclasses:** Indica las clases que pueden considerarse valores del *slot*. En los *slots* de relación, esta faceta indica el rango

de la relación, es decir, aquellas clases que pueden situarse a la derecha de la relación en el formalismo.

- g) **Cardinality:** Número de valores que puede tener el *slot* y su obligatoriedad. Pueden restringirse como obligatorio (*required*) y múltiple (*multiple*). Asimismo, podemos limitar el número mínimo (*at least*) y máximo (*at most*) de valores que puede tener el *slot*. Esta faceta se ha utilizado en nuestra implementación para reflejar la recursividad de las relaciones marcando la casilla múltiple en el caso de las relaciones recursivas.
- h) **Inverse slot:** Esta faceta indica si un *slot* tiene un *slot* inverso definido y, en su caso, el nombre del mismo. En nuestra propuesta de implementación se ha empleado para consignar las relaciones inversas, si las hubiere.
- i) **Minimum-maximum:** Valores máximos y mínimos. Esta faceta sólo es aplicable para los tipos de valor *Integer* o *Float*, que no son relevantes para el establecimiento de relaciones conceptuales. Por tanto, no se ha utilizado en nuestra implementación.
- j) **Domain:** Esta faceta indica el dominio del *slot*, es decir, las clases de la jerarquía a las que se puede asignar un *slot* determinado. En nuestra implementación se ha empleado para indicar qué conceptos pueden establecer una relación, a saber, qué clases conceptuales pueden situarse a la izquierda del formalismo que representa la relación.

Como vemos, las facetas estándar definidas para los *slots* en *Protégé* permiten dar cuenta de algunos de los aspectos de su formalización descritos en el apartado 6.1. Así, mediante dichas facetas podemos restringir las clases conceptuales vinculadas mediante una relación (mediante la faceta *Allowed Superclasses* y *Domain*), y también podemos indicar si la relación es recursiva (mediante la faceta *Cardinality*) y si posee una relación inversa (mediante la faceta *Inverse Slot*). Sin embargo, la implementación de las propiedades matemáticas de las relaciones conceptuales no queda suficientemente explícita con las facetas estándar. Para superar esta limitación, se ha definido un *metaslot* que permite incluir nuevas facetas en las que incorporar la información sobre dichas propiedades.

En concreto, hemos creado un subordinado de la clase :STANDARD SLOT al que hemos llamado :RELATION SLOT. En él hemos añadido a las facetas estándar una nueva, denominada *propiedades matemáticas*, mediante la que reflejamos si una relación es transitiva, simétrica o recursiva. Esta nueva faceta, cuyo tipo de valor es *symbol*, puede tener los siguientes valores: *transitiva*, *no_transitiva*, *simétrica*, *no_simétrica*, *recursiva*, *no_recursiva*. El valor asignado a este *slot* puede ser múltiple, para que sea posible reflejar las tres propiedades matemáticas de las relaciones simultáneamente. En la Fig. 30 se presenta una captura de pantalla en la que se indican las facetas del *slot propiedades matemáticas*:

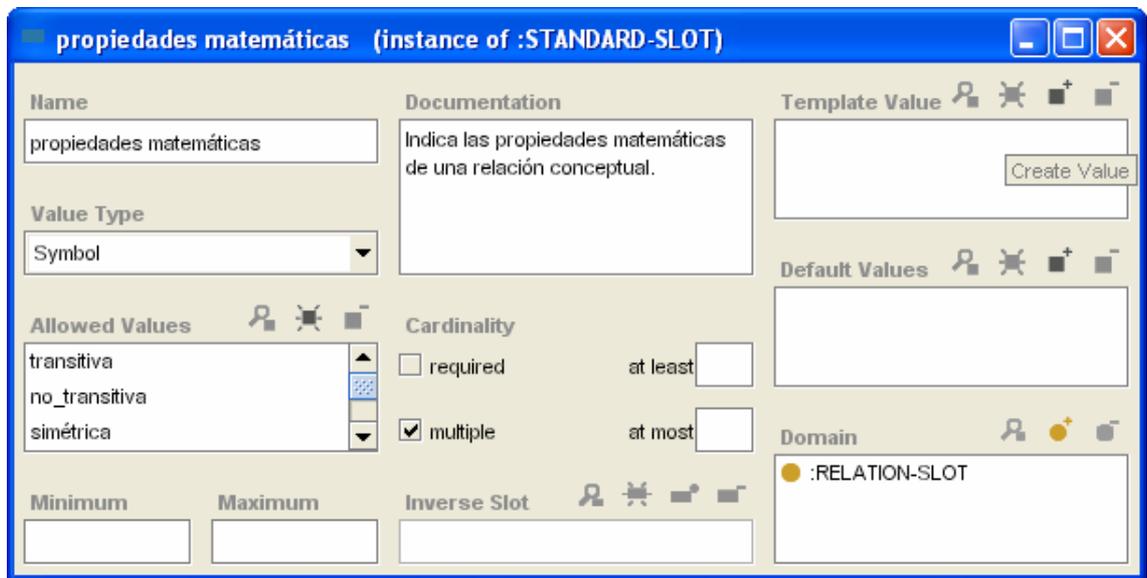


Fig. 30. Nueva faceta *propiedades matemáticas* añadida para todos los *slots* de relación.

Así, las relaciones conceptuales definidas mediante *slots* en *Ontoceram* se crearán dependiendo de :RELATION-SLOT, lo que nos permitirá hacer explícitas sus propiedades matemáticas. A modo de ejemplo, en la Fig. 31 se muestra una captura de pantalla de la implementación de la relación argumental *proceso-producto* en nuestra base de datos conceptual:

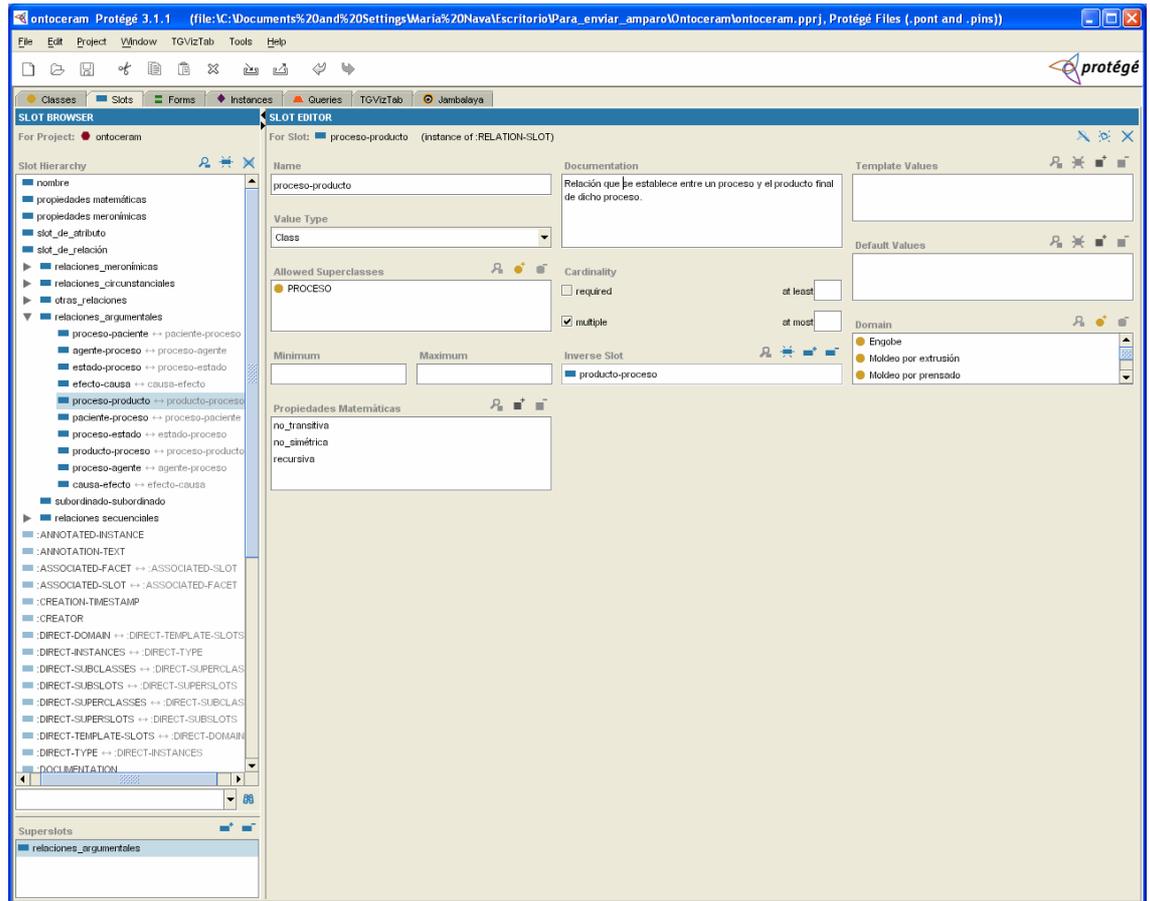


Fig. 31. Implementación de la relación *proceso-producto* en *Ontoceram*.

Como se ve en la figura, la relación *proceso-producto* es una relación definida a partir del *metaslot* `RELATION:SLOT` como subordinada de *relaciones_argumentales* en la jerarquía de *slots*. En cuanto a sus propiedades matemáticas, se trata de una relación no transitiva, no simétrica y recursiva. La recursividad de la relación también se refleja en la faceta *Cardinality*, donde se indica que puede ser una relación múltiple y no es obligatoria. En cuanto a su dominio y rango, es decir, las clases que pueden situarse a cada lado de la relación, en el apartado *Domain* aparecen los conceptos que pueden situarse a la izquierda de la misma, y en el apartado *Allowed Superclasses* se indica su rango, es decir, las clases que son susceptibles de situarse a la derecha de la relación, que en este caso podría ser cualquier OBJETO. Para cada concepto representado a través del componente *Class* es posible definir sus valores por defecto y, de este modo, introducir automáticamente el valor de la relación en las instancias. En ellas podremos modificar estos valores si es necesario. Por último, en la figura se aprecia que esta relación posee una relación inversa, la relación *producto-proceso*.

Un caso especial de las relaciones implementadas mediante *slots* de relación lo constituyen las **relaciones meronímicas**. Como veíamos en el apartado 6.2, los subtipos de esta relación se distinguen entre sí por cuatro rasgos adicionales a las propiedades matemáticas descritas para el resto de las relaciones. En el apartado 6.1 hemos distinguido siete tipos de relaciones meronímicas en función de cuatro propiedades definidas por Chaffin y Winston (1988): funcional, homeomérica, simultánea y separable. Para poder reflejar este hecho en nuestra base de datos conceptual es necesario añadir una nueva faceta a la relación que nos permita indicar estos rasgos.

Dado que son un caso especial de las relaciones conceptuales implementadas mediante *slots*, para dar cuenta de sus rasgos distintivos hemos creado una metaclassa dependiente de `:RELATION-SLOT`, denominada `:MERONYMIC-RELATION-SLOT`, en la que se añade la faceta *propiedades meronímicas*, con el fin de hacer explícitas las estas propiedades específicas.

En la Fig. 32 se muestra una captura de pantalla del slot *propiedades meronímicas*, que se añade a los *slots* de relación creados con la metaclassa `:MERONYMIC-RELATION-SLOT`.

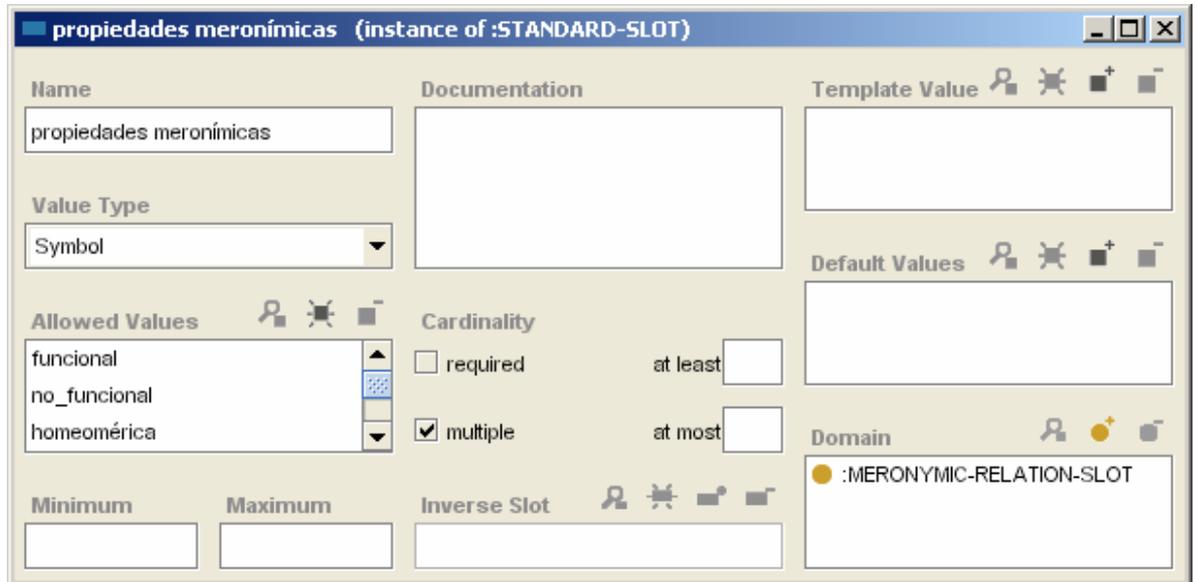


Fig. 32. Slot *propiedades meronímicas*, añadido a las relaciones creadas con la metaclassa `:MERONYMIC-RELATION-SLOT`.

Como podemos ver, este *slot* permite seleccionar los siguientes valores de las propiedades meronímicas: *funcional*, *no_funcional*, *homeomérica*, *no_homeomérica*, *simultánea*, *no_simultánea*, *separable* y *no_separable*.

En la Fig. 33 se muestra la relación meronímica componente-objeto, en la que podemos ver cómo se reflejan las propiedades meronímicas (funcional, no homeomérica, simultánea y separable), además de sus propiedades matemáticas (no transitiva, no simétrica y no reflexiva):

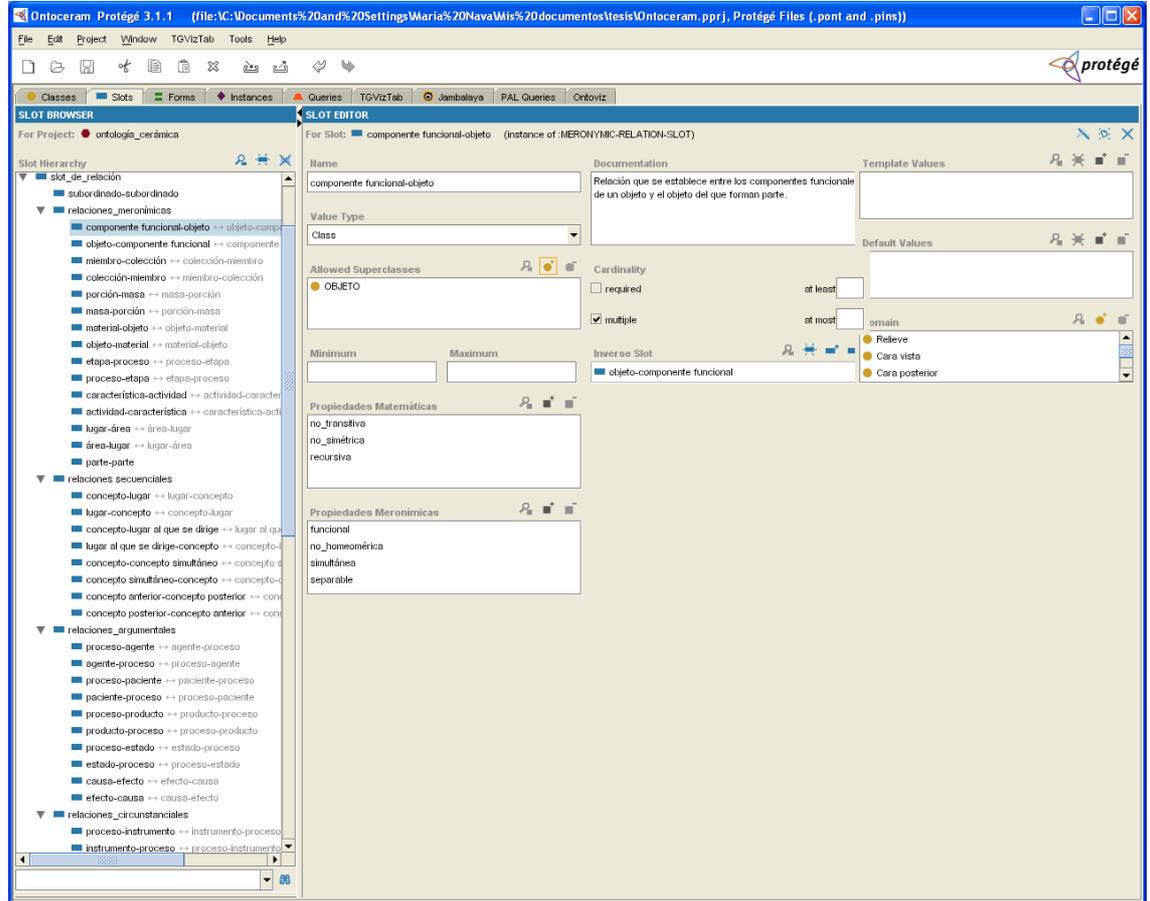


Fig. 33. Relación meronímica *componente funcional-objeto* en *Ontoceram*.

c) Implementación de la relación de sinonimia

Hasta ahora hemos descrito en este apartado cómo se implementan en *Ontoceram* las relaciones conceptuales. En el apartado 6.1, se incluye la relación de sinonimia al final del catálogo, dentro de *Otras relaciones*. Como apuntábamos entonces, la sinonimia no es *stricto sensu* una relación entre conceptos, sino entre denominaciones. En nuestra implementación, las denominaciones se implementan a través del componente *Instance*, y por ello será a través de ellas donde se haga explícito este tipo de relación.

En otras implementaciones en *Protégé* dentro del ámbito de la ingeniería del conocimiento, se ha propuesto la formalización de la relación de sinonimia mediante la creación de una metaclassa a la que se añade un *slot* que posibilita la

adición de sinónimos en una cadena de caracteres⁶². Sin embargo, en *Ontoceram* las distintas denominaciones que se emplean en los textos para referirse a un mismo concepto se hacen explícitas a través del componente *Instance*. Esta forma de introducir los sinónimos posee una ventaja adicional, y es que posibilitará las consultas en Protégé a partir de todas las posibles denominaciones de un concepto en la pestaña *Queries Tab*. En la Fig. 34 pueden observarse como ejemplo las denominaciones sinónimas que se emplean para referirse al concepto *Soporte*, que según nuestro análisis de los textos de ámbito de la cerámica son tres: *soporte*, *cuerpo* y *bizcocho*. En nuestra base de conocimiento hemos creado tres instancias para esta clase, que hacen referencia a estas tres denominaciones.

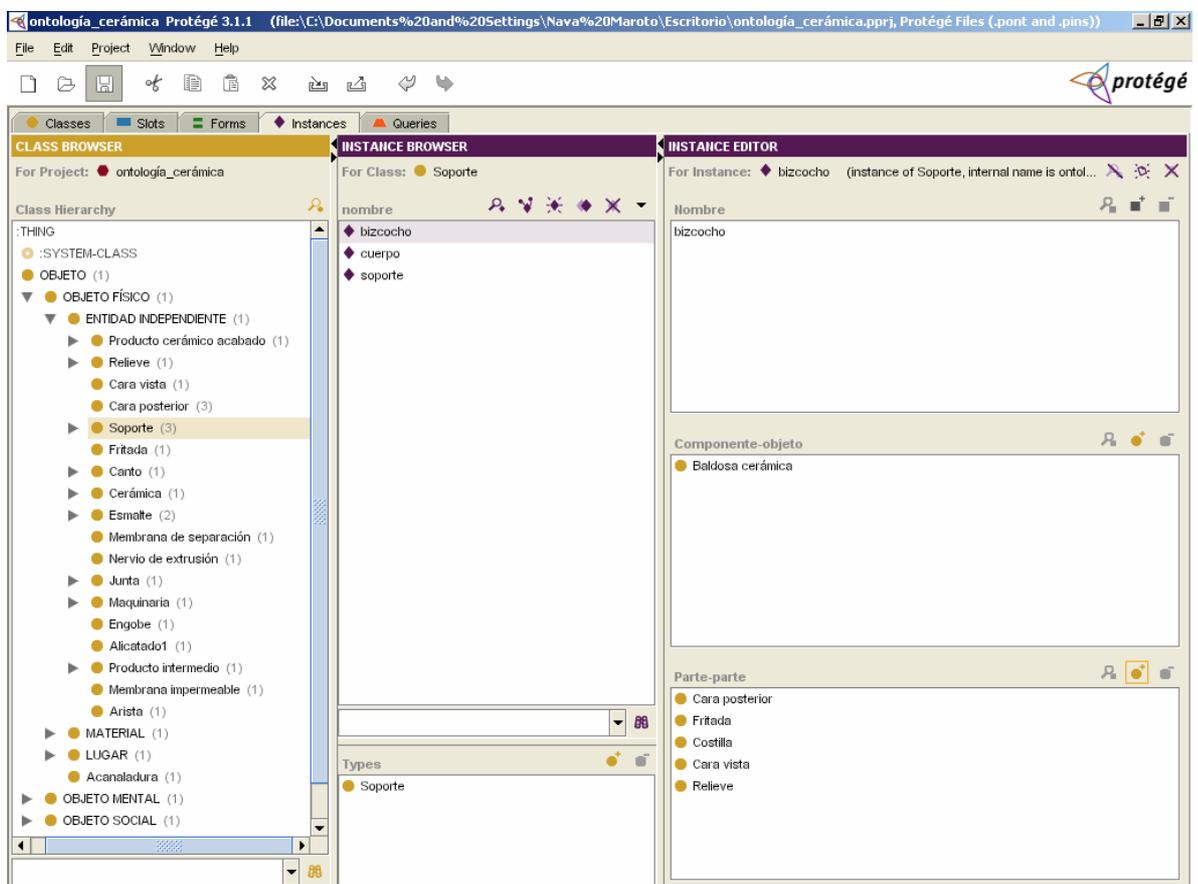


Fig. 34. Sinónimos empleados para referirse al concepto *Soporte* recogidos en *Ontoceram*.

⁶² Noy (2005) describe en detalle este tipo de implementación en Protégé en la página web de documentación de la herramienta. Para obtener una descripción más detallada, recomendamos leer el apartado *Creating Synonyms* de la página de Protégé, disponible en la dirección <http://protege.cim3.net/cgi-bin/wiki.pl?CreatingSynonyms> (Fecha de consulta: junio de 2007).

Este modo de incluir los sinónimos como distintas denominaciones a través del componente *Instance* podría adaptarse para la formalización de otras relaciones de equivalencia, como las que se establecen entre las distintas denominaciones para un mismo concepto en varias lenguas, si bien esta posibilidad no se ha explorado en la presente investigación.

6.3.2 Tabla representativa de la implementación de las relaciones

En este apartado presentamos el cuadro que se empleará en el apartado 6.4 para representar cómo se han implementado las relaciones conceptuales del catálogo en el editor de ontologías. Este cuadro reproduce de manera visual las facetas, definidas en el apartado 6.3.1.3, que se han definido para cada *slot* de relación, incluyendo las dos facetas introducidas para reflejar la propiedades matemáticas y meronímicas. En él se recogen las siguientes informaciones sobre los *slots* que representan cada relación:

- a) **Nombre de la relación conceptual**
- b) **Tipo de valor**
- c) **Cardinalidad de la relación**
- d) **Slot inverso**
- e) **Valores de plantilla**
- f) **Valores por defecto**
- g) **Propiedades matemáticas**
- h) **Propiedades meronímicas**
- i) **Valores permitidos (rango)**
- j) **Dominio**

La Tabla 20 presenta cómo se resume la implementación de cada relación conceptual en nuestra base de conocimiento *Ontoceram*.. Esta tabla se completará con los datos concretos de cada *slot* de relación en la descripción detallada del catálogo de relaciones que se propone en el apartado 6.4:

Nombre de la relación		
Tipo de Valor <input type="checkbox"/> Cualquiera <input type="checkbox"/> Booleano <input checked="" type="checkbox"/> Clase <input type="checkbox"/> Número con decimal <input type="checkbox"/> Instancia <input type="checkbox"/> Número entero <input type="checkbox"/> Cadena de caracteres <input type="checkbox"/> Símbolo	Cardinalidad <input type="checkbox"/> Obligatorio <input type="checkbox"/> Múltiple Slot inverso <input type="text"/>	Valores de Plantilla <input type="text"/>
Propiedades matemáticas <input type="checkbox"/> Transitiva <input type="checkbox"/> Simétrica <input type="checkbox"/> Recursiva	Valores permitidos (rango) <input type="text"/>	Valores por defecto <input type="text"/>
Propiedades meronímicas <input type="checkbox"/> Funcional <input type="checkbox"/> Homeomérica <input type="checkbox"/> Separable <input type="checkbox"/> Simultánea	Dominio <input type="text"/>	<input type="text"/>

Tabla 20. Cuadro resumen de la implementación de los slots de relación en *Ontoceram*.

6.4 Definición, formalización e implementación de las relaciones conceptuales del catálogo

En esta sección se presenta de forma detallada y conjunta toda la información sobre cada una de las relaciones conceptuales del catálogo que se ha descrito en las secciones anteriores. Para cada relación se recogen las siguientes informaciones:

- La descripción del significado de la relación en lenguaje natural
- Clases conceptuales entre las que puede establecerse
- Propiedades matemáticas de cada relación y propiedades específicas de las relaciones meronímicas
- Implementación en *Protégé*
- Ejemplos extraídos de nuestro análisis empírico

6.4.1 Relaciones lógicas

A continuación definimos los dos tipos de relación lógica considerados en nuestro catálogo, la relación superordinado-subordinado y la relación

subordinado-subordinado, y explicamos cómo se han formalizado e implementado en el presente trabajo.

6.4.1.1 Relación superordinado-subordinado

a) Definición

Relación que se establece entre un concepto *a* que posee un grado de abstracción superior y un concepto *b* que posee un grado de abstracción inferior, es decir, que se sitúa en un nivel inferior de la jerarquía. El concepto *b* posee las mismas características que el concepto *a*, más alguna característica adicional que aumenta su intensión, es decir, que lo hace más específico.

Debemos hacer hincapié en que en el presente trabajo se ha considerado únicamente la relación que se establece entre un superordinado y su/s subordinado/s directos. El sistema que vamos a emplear para la implementación de las relaciones conceptuales permite que se mantenga la herencia de características y relaciones entre conceptos vinculados mediante esta relación, lo que implica que los subordinados heredan todas las propiedades de su superordinado. De este modo, no será necesario especificar todas las características del subordinado de nuevo, sino que bastará con añadir la/s característica/s adicional/es o modificar los valores de las mismas en el subordinado, en caso de que varíen con respecto a los valores del superordinado.

b) Clases conceptuales

La relación puede establecerse entre entidades, propiedades y actividades. Los conceptos situados a ambos lados de la relación deben pertenecer a la misma clase.

c) Propiedades de la relación

La relación superordinado-subordinado es una relación transitiva, no simétrica, que puede presentar recursividad y que tiene una relación inversa, la relación subordinado-superordinado.

d) Implementación de la relación en *Protégé*

La relación entre un superordinado y sus subordinados y su inversa (la relación subordinado-superordinado) se han implementado aprovechando la organización jerárquica del componente *Class* en *Protégé*. Así, estas dos relaciones se hacen explícitas en virtud del lugar que ocupan los conceptos en la jerarquía.

e) Ejemplos

Relación entre entidades: *pieza cerámica – baldosa cerámica*

Relación entre propiedades: *coeficiente de fricción – coeficiente de fricción dinámico*

Relación entre actividades: *colocación – colocación en capa fina*

6.4.1.2 Relación subordinado-subordinado

a) Definición

Relación que se establece entre dos conceptos que se encuentran situados en el mismo nivel de abstracción por debajo de un superordinado común. Estos conceptos, que se denominan cohipónimos, comparten características con el superordinado común, pero tienen al menos un rasgo que los diferencia entre sí, y que los diferencia a su vez del superordinado.

b) Clases conceptuales

Este tipo de relación puede establecerse entre entidades, actividades y propiedades, siempre que ambos conceptos relacionados pertenezcan a la misma clase conceptual.

c) Propiedades de la relación

Es una relación transitiva, simétrica, puede presentar recursividad y la relación se cumple en ambos sentidos, es decir, es inversa de sí misma.

d) Implementación de la relación en *Protégé*

La relación que se establece entre dos subordinados de un superordinado común se implementa en *Protégé* mediante el *slot* de relación subordinado-subordinado, cuyas facetas se resumen en el siguiente cuadro:

Nombre de la relación		subordinado-	subordinado		
Tipo de valor	<input type="checkbox"/> Cualquiera <input type="checkbox"/> Booleano <input checked="" type="checkbox"/> Clase <input type="checkbox"/> Número con decimal <input type="checkbox"/> Instancia	Cardinalidad	<input type="checkbox"/> Obligatorio <input checked="" type="checkbox"/> Múltiple	Valores de plantilla	<input type="text"/>
	<input type="checkbox"/> Número entero <input type="checkbox"/> Cadena de caracteres <input type="checkbox"/> Símbolo		Slot inverso		<input type="text"/>
Propiedades lógicas	<input checked="" type="checkbox"/> Transitiva <input checked="" type="checkbox"/> Simétrica <input checked="" type="checkbox"/> Recursiva	Valores permitidos (rango)	PROPIEDAD ACTIVIDAD ENTIDAD	Dominio	PROPIEDAD ACTIVIDAD ENTIDAD
Propiedades meronímicas	<input type="checkbox"/> Funcional <input type="checkbox"/> Homeomérica <input type="checkbox"/> Separable <input type="checkbox"/> Simultánea				

Tabla 21. Implementación de la relación subordinado-subordinado.

e) Ejemplos

Relación entre entidades: *cubrecantos – escuadra*

Relación entre propiedades: *curvatura – grosor*

Relación entre actividades: *colocación con junta – colocación sin junta*

6.4.2 Relaciones meronímicas

A continuación se define cada uno de los subtipos de meronimia identificados, junto con sus correspondientes propiedades matemáticas y meronímicas, nuestra propuesta de implementación y ejemplos extraídos del corpus.

6.4.2.1 Relación componente funcional-objeto

a) Definición

La relación componente funcional-objeto se establece entre los componentes que desempeñan una función dentro del todo del que forman parte. Los componentes de un objeto integrado pueden ser canónicos (si son necesarios para la constitución del todo) o facultativos (si no son imprescindibles para que el todo continúe siéndolo).

b) Clases conceptuales

Generalmente, tanto las partes como el todo han de ser entidades.

c) Propiedades de la relación

Como vimos en el apartado 2.4.2, la relación entre un componente y el objeto en que se integra puede ser transitiva o no. La relación no es transitiva cuando la parte es un accesorio (*attachment*) del todo y tiene un dominio de función variable que puede ser descrito de manera independiente. La transitividad se mantiene cuando la parte es integral con respecto al todo y su dominio de función no es único, sino variable (puede tener varios todos posibles).

Es una relación no simétrica, que puede presentar recursividad (cuando el dominio de función es variable) y su relación inversa es la relación objeto-componente funcional.

En cuanto a las propiedades meronímicas, la relación se define como funcional, no homeomérica, separable y generalmente simultánea.

d) Implementación de la relación en *Protégé*

Nombre de la relación		componente	funcional-	objeto	
Tipo de valor	<input type="checkbox"/> Cualquiera <input type="checkbox"/> Booleano <input checked="" type="checkbox"/> Clase <input type="checkbox"/> Número con decimal <input type="checkbox"/> Instancia <input type="checkbox"/> Número entero <input type="checkbox"/> Cadena de caracteres <input type="checkbox"/> Símbolo	Cardinalidad	<input type="checkbox"/> Obligatorio <input checked="" type="checkbox"/> Múltiple	Valores de plantilla	<input type="text"/>
	Slot inverso		objeto-componente funcional	Valores por defecto	<input type="text"/>
Propiedades lógicas	<input checked="" type="checkbox"/> Transitiva <input type="checkbox"/> Simétrica <input checked="" type="checkbox"/> Recursiva	Valores permitidos (rango)	ENTIDAD	Dominio	ENTIDAD
Propiedades meronímicas	<input checked="" type="checkbox"/> Funcional <input type="checkbox"/> Homeomérica <input checked="" type="checkbox"/> Separable <input checked="" type="checkbox"/> Simultánea				

Tabla 22. Implementación de la relación componente funcional-objeto.

e) Ejemplos

Relación entre entidades: *arista – baldosa*

olambrilla – baldosín catalán

6.4.2.2 Relación miembro-colección

a) Definición

Relación que se establece entre partes que mantienen una relación de proximidad espacial o conexión social con respecto al todo, sin necesidad de que cumplan una función específica ni de que estén dispuestas estructuralmente de un modo determinado. Es su similitud con el resto de los miembros la que determina su pertenencia a la colección.

b) Clases conceptuales

Generalmente esta relación se establece entre entidades.

c) Propiedades de la relación

Se trata de una relación no transitiva, no simétrica y que puede presentar recursividad. Su relación inversa es la relación colección-miembro.

En cuanto a sus propiedades meronímicas, la relación miembro-colección posee los rasgos siguientes: separable, generalmente simultánea, no funcional y no homeomérica.

d) Implementación de la relación en Protégé

Nombre de la relación		miembro- colección			
Tipo de valor	<input type="checkbox"/> Cualquiera <input type="checkbox"/> Booleano <input checked="" type="checkbox"/> Clase <input type="checkbox"/> Número con decimal <input type="checkbox"/> Instancia <input type="checkbox"/> Número entero <input type="checkbox"/> Cadena de caracteres <input type="checkbox"/> Símbolo	Cardinalidad	<input type="checkbox"/> Obligatorio <input checked="" type="checkbox"/> Múltiple	Valores de plantilla	<input type="text"/>
		Slot inverso	colección-miembro	Valores por defecto	<input type="text"/>
Propiedades Lógicas	<input type="checkbox"/> Transitiva <input type="checkbox"/> Simétrica <input checked="" type="checkbox"/> Recursiva	Valores permitidos (rango)	ENTIDAD	Dominio	ENTIDAD
Propiedades meronímicas	<input type="checkbox"/> Funcional <input type="checkbox"/> Homeomérica <input checked="" type="checkbox"/> Separable <input checked="" type="checkbox"/> Simultánea				

Tabla 23. Implementación de la relación miembro-colección.

e) Ejemplos

Relación entre entidades: *tesela - mosaico*

6.4.2.3 Relación porción-masa

a) Definición

Relación entre una parte y el todo en la que la naturaleza de la porción es la misma que la de la masa. Es precisamente este rasgo, la condición de ser homeomérica, el que la distingue de otros tipos de meronimia.

b) Clases conceptuales

Los conceptos vinculados mediante la relación porción-masa suelen ser entidades.

c) Propiedades de la relación

Se trata de una relación transitiva, no simétrica, que puede presentar recursividad. Su relación inversa es la relación masa-porción.

En cuanto a las propiedades meronímicas de la relación, es homeomérica, no funcional, separable y simultánea.

d) Implementación de la relación en *Protégé*

Nombre de la relación		porción-masa			
Tipo de valor	<input type="checkbox"/> Cualquiera <input type="checkbox"/> Booleano <input checked="" type="checkbox"/> Clase <input type="checkbox"/> Número con decimal <input type="checkbox"/> Instancia <input type="checkbox"/> Número entero <input type="checkbox"/> Cadena de caracteres <input type="checkbox"/> Símbolo	Cardinalidad	<input type="checkbox"/> Obligatorio <input checked="" type="checkbox"/> Múltiple	Valores de plantilla	<input type="text"/>
		Slot inverso	masa- porción	Valores por defecto	<input type="text"/>
Propiedades lógicas	<input checked="" type="checkbox"/> Transitiva <input type="checkbox"/> Simétrica <input checked="" type="checkbox"/> Recursiva	Valores permitidos (rango)	ENTIDAD	Dominio	ENTIDAD
Propiedades meronímicas	<input type="checkbox"/> Funcional <input checked="" type="checkbox"/> Homeomérica <input checked="" type="checkbox"/> Separable <input checked="" type="checkbox"/> Simultánea				

Tabla 24. Implementación de la relación porción-masa.

e) Ejemplos

De momento no hemos encontrado ningún ejemplo de la relación porción-masa en nuestro estudio empírico.

6.4.2.4 Relación material-objeto

a) Definición

Relación entre una parte intrínseca y el todo al que pertenece, del que no puede separarse por ser consustancial al mismo. Su rasgo distintivo es que no es separable con respecto al todo.

b) Clases conceptuales

Esta relación se establece generalmente entre entidades.

c) Propiedades de la relación

Es una relación transitiva, no simétrica, y puede presentar recursividad. Su relación inversa es la relación objeto-material.

En cuanto a las propiedades meronímicas, el material es simultáneo con respecto al todo, pero no es separable y no tiene por qué ser ni funcional, ni homeomérico.

d) Implementación de la relación en *Protégé*

Nombre de la relación		material-objeto			
Tipo de valor	<input type="checkbox"/> Cualquiera <input type="checkbox"/> Booleano <input checked="" type="checkbox"/> Clase <input type="checkbox"/> Número con decimal <input type="checkbox"/> Instancia <input type="checkbox"/> Número entero <input type="checkbox"/> Cadena de caracteres <input type="checkbox"/> Símbolo	Cardinalidad	<input type="checkbox"/> Obligatorio <input checked="" type="checkbox"/> Múltiple	Valores de plantilla	<input type="text"/>
		Slot inverso	objeto-material	Valores por defecto	<input type="text"/>
Propiedades lógicas	<input checked="" type="checkbox"/> Transitiva <input type="checkbox"/> Simétrica <input checked="" type="checkbox"/> Recursiva	Valores permitidos (rango)	ENTIDAD	Dominio	ENTIDAD
Propiedades meronímicas	<input type="checkbox"/> Funcional <input type="checkbox"/> Homeomérica <input type="checkbox"/> Separable <input checked="" type="checkbox"/> Simultánea				

Tabla 25. Implementación de la relación material-objeto.

e) Ejemplos

Relación entre entidades: *sustancia viscosa – adhesivo*
arcilla – baldosa cerámica
arena natural – base de arena

6.4.2.5 Relación etapa-proceso

a) Definición

Relación que se establece entre las fases de una actividad y dicha actividad entendida como un todo. El principal rasgo que caracteriza a las etapas con respecto al proceso del que forman parte es la simultaneidad.

b) Clases conceptuales

Las clases conceptuales relacionadas mediante este tipo de meronimia suelen ser actividades.

c) Propiedades de la relación

Es una relación transitiva, no simétrica, que puede presentar recursividad. Su relación inversa es la relación proceso-etapa.

En cuanto a sus propiedades meronímicas, la relación etapa-proceso es simultánea, no necesariamente funcional (aunque puede serlo), no homeomérica y no separable del proceso.

d) Implementación de la relación en *Protégé*

Nombre de la relación		etapa- proceso			
Tipo de valor	<input type="checkbox"/> Cualquiera <input type="checkbox"/> Booleano <input checked="" type="checkbox"/> Clase <input type="checkbox"/> Número con decimal <input type="checkbox"/> Instancia <input type="checkbox"/> Número entero <input type="checkbox"/> Cadena de caracteres <input type="checkbox"/> Símbolo	Cardinalidad	<input type="checkbox"/> Obligatorio <input checked="" type="checkbox"/> Múltiple	Valores de plantilla	
		Slot inverso	proceso-etapa	Valores por defecto	
Propiedades lógicas	<input checked="" type="checkbox"/> Transitiva <input type="checkbox"/> Simétrica <input checked="" type="checkbox"/> Recursiva	Valores permitidos (rango)	ACTIVIDAD	Dominio	ACTIVIDAD
Propiedades meronímicas	<input type="checkbox"/> Funcional <input type="checkbox"/> Homeomérica <input type="checkbox"/> Separable <input checked="" type="checkbox"/> Simultánea				

Tabla 26. Implementación de la relación etapa-proceso.

e) Ejemplos

Relación ente actividades: *etapa de gran fuego – cocción*
desolidarización – colocación
ensayo – control de calidad

6.4.2.6 Relación característica-actividad

a) Definición

Relación que se establece entre las características de una actividad y el proceso complejo en el que se inscriben.

b) Clases conceptuales

Las clases conceptuales relacionadas son propiedades (dominio) y actividades (rango).

c) Propiedades de la relación

La relación entre las características de una actividad y dicha actividad no es necesariamente transitiva, no es simétrica y puede presentar recursividad. La relación inversa entre el proceso y las partes características también se cumple.

En cuanto a las propiedades meronímicas, la relación característica-actividad es funcional y simultánea, pero no es homeomérica ni separable.

d) Implementación de la relación en *Protégé*

Nombre de la relación		característica-	Actividad		
Tipo de valor	<input type="checkbox"/> Cualquiera <input type="checkbox"/> Booleano <input checked="" type="checkbox"/> Clase <input type="checkbox"/> Número con decimal <input type="checkbox"/> Instancia <input type="checkbox"/> Número entero <input type="checkbox"/> Cadena de caracteres <input type="checkbox"/> Símbolo	Cardinalidad	<input type="checkbox"/> Obligatorio <input checked="" type="checkbox"/> Múltiple	Valores de plantilla	<input type="text"/>
		Slot inverso	<input type="text" value="actividad-característica"/>	Valores por defecto	<input type="text"/>
Propiedades lógicas	<input type="checkbox"/> Transitiva <input type="checkbox"/> Simétrica <input checked="" type="checkbox"/> Recursiva	Valores permitidos (rango)	<input type="text" value="ACTIVIDAD"/>	Dominio	<input type="text" value="PROPIEDAD"/>
Propiedades meronímicas	<input checked="" type="checkbox"/> Funcional <input type="checkbox"/> Homeomérica <input type="checkbox"/> Separable <input checked="" type="checkbox"/> Simultánea				

Tabla 27. Implementación de la relación característica-actividad.

e) Ejemplos

Debido a que en nuestro trabajo empírico nos hemos fijado sobre todo en las relaciones establecidas entre los productos acabados, de momento no hemos encontrado ninguna relación entre una actividad y sus características.

6.4.2.7 Relación lugar-área

a) Definición

Relación que se establece entre un área espacial y las localizaciones concretas situadas dentro de esa área. El rasgo que distingue a este tipo de meronimia es que el lugar es homeomérico con respecto al área, es decir, posee la misma naturaleza.

b) Clases conceptuales

Las clases conceptuales relacionadas mediante la relación lugar-área suelen ser entidades.

c) Propiedades de la relación

Se trata de una relación transitiva, no simétrica y que normalmente no presenta recursividad. La relación inversa (relación área-lugar) también se cumple en todos los casos.

En cuanto a las propiedades meronímicas, la relación lugar-área es homeomérica, no funcional, no separable y no necesariamente simultánea.

d) Implementación de la relación en Protégé

Nombre de la relación		lugar-área			
Tipo de valor	<input type="checkbox"/> Cualquiera <input type="checkbox"/> Booleano <input checked="" type="checkbox"/> Clase <input type="checkbox"/> Número con decimal <input type="checkbox"/> Instancia <input type="checkbox"/> Número entero <input type="checkbox"/> Cadena de caracteres <input type="checkbox"/> Símbolo	Cardinalidad	<input type="checkbox"/> Obligatorio <input type="checkbox"/> Múltiple	Valores de plantilla	<input type="text"/>
		Slot inverso	área-lugar Defecto	Valores por	<input type="text"/>
Propiedades lógicas	<input checked="" type="checkbox"/> Transitiva <input type="checkbox"/> Simétrica <input type="checkbox"/> Recursiva	Valores permitidos (rango)	ENTIDAD	Dominio	ENTIDAD
Propiedades meronímicas	<input type="checkbox"/> Funcional <input checked="" type="checkbox"/> Homeomérica <input type="checkbox"/> Separable <input type="checkbox"/> Simultánea				

Tabla 28. Implementación de la relación lugar-área.

e) Ejemplos

De momento no hemos encontrado ningún ejemplo de la relación meronímica entre lugares y áreas.

6.4.2.8 Relación parte-parte

a) Definición

Relación horizontal que se establece entre las partes que juntas constituyen un todo. En este caso no consideramos operativo distinguir entre los distintos tipos de meronimia, por lo que proponemos una sola relación para expresar las relaciones entre componentes, miembros, porciones, materiales, etapas y características.

b) Clases conceptuales

Pueden ser entidades, actividades o propiedades, dependiendo del subtipo de meronimia de que se trate.

c) Propiedades de la relación

Es una relación transitiva, simétrica y que puede presentar recursividad. Esta relación es inversa de sí misma.

Dado que no hemos separado entre los distintos tipos de meronimia, no consideramos los rasgos meronímicos específicos de esta relación.

d) Implementación de la relación en *Protégé*

Nombre de la relación		parte-parte			
Tipo de valor	<input type="checkbox"/> Cualquiera <input type="checkbox"/> Booleano <input checked="" type="checkbox"/> Clase <input type="checkbox"/> Número con decimal <input type="checkbox"/> Instancia <input type="checkbox"/> Número entero <input type="checkbox"/> Cadena de caracteres <input type="checkbox"/> Símbolo	Cardinalidad	<input type="checkbox"/> Obligatorio <input checked="" type="checkbox"/> Múltiple	Valores de plantilla	<input type="text"/>
		Slot inverso	<input type="text"/>	Valores por defecto	<input type="text"/>
Propiedades lógicas	<input checked="" type="checkbox"/> Transitiva <input checked="" type="checkbox"/> Simétrica <input checked="" type="checkbox"/> Recursiva	Valores permitidos (rango)	ACTIVIDAD ENTIDAD PROPIEDAD	Dominio	ACTIVIDAD ENTIDAD PROPIEDAD
Propiedades meronímicas	<input type="checkbox"/> Funcional <input type="checkbox"/> Homeomérica <input type="checkbox"/> Separable <input type="checkbox"/> Simultánea				

Tabla 29. Implementación de la relación parte-parte.

e) Ejemplos

Relación entre entidades: *cara vista – arista*
esmalte – soporte

Relación entre actividades: *encolado – dosificación*
engobe – vidriado

6.4.3 Relaciones secuenciales

Las relaciones secuenciales son aquellas que se establecen por localización o sucesión en el espacio o en el tiempo de los elementos que vincula. Este tipo de relación presenta dos subtipos: secuencialidad espacial y secuencialidad temporal. La primera a su vez puede ser de localización o de dirección y la secuencialidad temporal puede ser de simultaneidad y de anterioridad-posterioridad. A continuación describimos en detalle cada uno de estos subtipos.

6.4.3.1 Relación concepto-lugar

a) Definición

Relación que se establece entre dos conceptos que se ubican juntos en el espacio.

b) Clases conceptuales

Las clases conceptuales pueden ser entidades y actividades.

c) Propiedades de la relación

Es una relación simétrica, puede ser transitiva o no, dependiendo de los casos. Esta relación puede presentar recursividad. Su relación inversa es la relación lugar-concepto.

d) Implementación de la relación en *Protégé*

Nombre de la relación		concepto- lugar			
Tipo de valor	<input type="checkbox"/> Cualquiera <input type="checkbox"/> Booleano <input checked="" type="checkbox"/> Clase <input type="checkbox"/> Número con decimal <input type="checkbox"/> Instancia <input type="checkbox"/> Número entero <input type="checkbox"/> Cadena de caracteres <input type="checkbox"/> Símbolo	Cardinalidad	<input type="checkbox"/> Obligatorio <input checked="" type="checkbox"/> Múltiple	Valores de plantilla	<input type="text"/>
		Slot inverso	<input type="text" value="lugar-concepto"/>	Valores por defecto	<input type="text"/>
Propiedades lógicas	<input checked="" type="checkbox"/> Transitiva <input checked="" type="checkbox"/> Simétrica <input checked="" type="checkbox"/> Recursiva	Valores permitidos (rango)	<input type="text" value="ACTIVIDAD ENTIDAD"/>	Dominio	<input type="text" value="ACTIVIDAD ENTIDAD"/>
Propiedades meronímicas	<input type="checkbox"/> Funcional <input type="checkbox"/> Homeomérica <input type="checkbox"/> Separable <input type="checkbox"/> Simultánea				

Tabla 30. Implementación de la relación concepto-lugar.

e) Ejemplos

Relación entre entidades: *adhesivo – adherendo*
baldosa cerámica – suelo

6.4.3.2 Relación concepto-lugar al que se dirige

a) Definición

Relación que se establece entre dos conceptos para indicar la dirección en la que se produce una actividad determinada.

b) Clases conceptuales

Las clases conceptuales entre las que puede establecerse esta relación son entidades y actividades.

c) Propiedades de la relación

Esta relación puede ser transitiva, simétrica, y puede presentar recursividad. Su relación inversa es la relación lugar al que se dirige-concepto.

d) Implementación de la relación en *Protégé*

Nombre de la relación	concepto-	lugar al que se dirige
Tipo de valor <input type="checkbox"/> Cualquiera <input type="checkbox"/> Booleano <input checked="" type="checkbox"/> Clase <input type="checkbox"/> Número con decimal <input type="checkbox"/> Instancia <input type="checkbox"/> Número entero <input type="checkbox"/> Cadena de caracteres <input type="checkbox"/> Símbolo	Cardinalidad <input type="checkbox"/> Obligatorio <input checked="" type="checkbox"/> Múltiple	Valores de plantilla <div style="border: 1px solid black; width: 100px; height: 40px;"></div>
Propiedades lógicas <input checked="" type="checkbox"/> Transitiva <input checked="" type="checkbox"/> Simétrica <input checked="" type="checkbox"/> Recursiva	Slot inverso lugar al que se dirige-concepto	Valores por defecto <div style="border: 1px solid black; width: 100px; height: 80px;"></div>
Propiedades meronómicas <input type="checkbox"/> Funcional <input type="checkbox"/> Homeomérica <input type="checkbox"/> Separable <input type="checkbox"/> Simultánea	Valores permitidos (rango) ACTIVIDAD ENTIDAD	Dominio ACTIVIDAD ENTIDAD

Tabla 31. Implementación de la relación concepto-lugar al que se dirige.

e) Ejemplos

De momento no han aparecido ejemplos en nuestro análisis.

6.4.3.3 Relación concepto-concepto simultáneo

a) Definición

Relación que se establece entre dos actividades o entidades que suceden en el mismo espacio de tiempo.

b) Clases conceptuales

Las clases conceptuales relacionadas suelen ser actividades, aunque también pueden ser entidades.

c) Propiedades de la relación

Es una relación simétrica y transitiva, que puede presentar recursividad. Su relación inversa es la relación concepto simultáneo-concepto. Es una relación inversa de sí misma.

d) Implementación de la relación en *Protégé*

Nombre de la relación		concepto-	concepto	simultáneo	
Tipo de valor	<input type="checkbox"/> Cualquiera <input type="checkbox"/> Booleano <input checked="" type="checkbox"/> Clase <input type="checkbox"/> Número con decimal <input type="checkbox"/> Instancia <input type="checkbox"/> Número entero <input type="checkbox"/> Cadena de caracteres <input type="checkbox"/> Símbolo	Cardinalidad	<input type="checkbox"/> Obligatorio <input checked="" type="checkbox"/> Múltiple	Valores de plantilla	
		Slot inverso	concepto Simultáneo- concepto	Valores por defecto	
Propiedades lógicas	<input checked="" type="checkbox"/> Transitiva <input checked="" type="checkbox"/> Simétrica <input checked="" type="checkbox"/> Recursiva	Valores permitidos (rango)	ACTIVIDAD ENTIDAD	Dominio	ACTIVIDAD ENTIDAD
Propiedades meronímicas	<input type="checkbox"/> Funcional <input type="checkbox"/> Homeomérica <input type="checkbox"/> Separable <input type="checkbox"/> Simultánea				

Tabla 32. Implementación de la relación concepto-concepto simultáneo.

e) Ejemplos

De momento no se han encontrado ejemplos de este tipo de relación en el corpus.

6.4.3.4 Relación concepto anterior-concepto posterior

a) Definición

Relación que se establece entre dos conceptos que se suceden en el tiempo de forma consecutiva.

b) Clases conceptuales

Las clases conceptuales entre las que se puede establecer esta relación son generalmente actividades, aunque también podría darse entre entidades.

c) Propiedades de la relación

Es una relación no simétrica, transitiva, y que puede presentar recursividad. Su inversa es la relación concepto posterior-concepto anterior.

d) Implementación de la relación en *Protégé*

Nombre de la relación		concepto anterior-	concepto posterior
Tipo de valor	<input type="checkbox"/> Cualquiera <input type="checkbox"/> Booleano <input checked="" type="checkbox"/> Clase <input type="checkbox"/> Número con decimal <input type="checkbox"/> Instancia <input type="checkbox"/> Número entero <input type="checkbox"/> Cadena de caracteres <input type="checkbox"/> Símbolo	Cardinalidad	<input type="checkbox"/> Obligatorio <input checked="" type="checkbox"/> Múltiple
		Slot inverso	Valores de plantilla <input type="text"/>
			Valores por defecto <input type="text"/>
Propiedades lógicas	<input checked="" type="checkbox"/> Transitiva <input type="checkbox"/> Simétrica <input checked="" type="checkbox"/> Recursiva	Valores permitidos (rango)	Dominio ACTIVIDAD ENTIDAD
Propiedades meronómicas	<input type="checkbox"/> Funcional <input type="checkbox"/> Homeomérica <input type="checkbox"/> Separable <input type="checkbox"/> Simultánea		ACTIVIDAD ENTIDAD

Tabla 33. Implementación de la relación concepto anterior-concepto posterior.

e) Ejemplos

De momento no se han encontrado ejemplos de este tipo de relación en el análisis.

6.4.4 Relaciones argumentales y circunstanciales

En este apartado se presentan las relaciones argumentales y circunstanciales, que, como se ha visto en el apartado 6.1.2, son relaciones conceptuales paradigmáticas. Las relaciones argumentales y circunstanciales son: proceso-agente, proceso-paciente, proceso-producto, proceso-estado, causa-efecto, proceso-instrumento, proceso-método y objeto-uso.

6.4.4.1 Relación proceso-agente

a) Definición

Relación que se establece entre un proceso y las entidades que llevan a cabo dicho proceso.

b) Clases conceptuales

Las clases conceptuales implicadas pueden ser actividades (procesos) y entidades (agente), y generalmente los conceptos relacionados no pertenecen a la misma clase conceptual.

c) Propiedades de la relación

Esta relación es no transitiva, no simétrica, y puede presentar recursividad. Su relación inversa es la relación agente-proceso.

d) Implementación de la relación en *Protégé*

Nombre de la relación		proceso-	agente		
Tipo de valor	<input type="checkbox"/> Cualquiera <input type="checkbox"/> Booleano <input checked="" type="checkbox"/> Clase <input type="checkbox"/> Número con decimal <input type="checkbox"/> Instancia <input type="checkbox"/> Número entero <input type="checkbox"/> Cadena de caracteres <input type="checkbox"/> Símbolo	Cardinalidad	<input type="checkbox"/> Obligatorio <input checked="" type="checkbox"/> Múltiple	Valores de plantilla	<input type="text"/>
	Propiedades lógicas		<input type="checkbox"/> Transitiva <input type="checkbox"/> Simétrica <input checked="" type="checkbox"/> Recursiva	Slot inverso	agente-proceso
Propiedades meronímicas		<input type="checkbox"/> Funcional <input type="checkbox"/> Homeomérica <input type="checkbox"/> Separable <input type="checkbox"/> Simultánea	Valores permitidos (rango)		ENTIDAD
					ACTIVIDAD

Tabla 34. Implementación de la relación proceso-agente.

e) Ejemplos

De momento no hemos encontrado ejemplos de este tipo de relación.

6.4.4.2 Relación proceso-producto

a) Definición

Relación que se establece entre el proceso y el producto final de dicho proceso.

b) Clases conceptuales

Las clases conceptuales implicadas en esta relación son actividades (proceso) y entidades (producto), que para establecer esta relación no pertenecen a la misma clase.

c) Propiedades de la relación

Se trata de una relación no necesariamente transitiva, no simétrica, y puede presentar recursividad. Su relación inversa es la relación producto-proceso.

d) Implementación en *Protégé*

Nombre de la relación		proceso-	producto		
Tipo de valor	<input type="checkbox"/> Cualquiera <input type="checkbox"/> Booleano <input checked="" type="checkbox"/> Clase <input type="checkbox"/> Número con decimal <input type="checkbox"/> Instancia <input type="checkbox"/> Número entero <input type="checkbox"/> Cadena de caracteres <input type="checkbox"/> Símbolo	Cardinalidad	<input type="checkbox"/> Obligatorio <input checked="" type="checkbox"/> Múltiple	Valores de plantilla	<input type="text"/>
Propiedades lógicas	<input type="checkbox"/> Transitiva <input type="checkbox"/> Simétrica <input checked="" type="checkbox"/> Recursiva	Slot inverso	producto-proceso	Valores por defecto	<input type="text"/>
Propiedades meronómicas	<input type="checkbox"/> Funcional <input type="checkbox"/> Homeomérica <input type="checkbox"/> Separable <input type="checkbox"/> Simultánea	Valores permitidos (rango)	ENTIDAD	Dominio	ACTIVIDAD

Tabla 35. Implementación de la relación proceso-producto.

e) Ejemplos

moldeo – baldosa cerámica

engobe – baldosa engobada

doble encolado – macizado

6.4.4.3 Relación proceso-paciente

a) Definición

Relación que se establece entre un proceso y la entidad sobre la que se lleva a cabo dicho proceso.

b) Clases conceptuales

La relación se establece entre una actividad (proceso) y una entidad (paciente).

c) Propiedades de la relación

Es una relación no transitiva, no simétrica, y que puede presentar recursividad. Su relación inversa es la relación paciente-proceso.

d) Implementación de la relación en Protégé

Nombre de la relación		proceso-	paciente		
Tipo de valor	<input type="checkbox"/> Cualquiera <input type="checkbox"/> Booleano <input checked="" type="checkbox"/> Clase <input type="checkbox"/> Número con decimal <input type="checkbox"/> Instancia	Cardinalidad	<input type="checkbox"/> Obligatorio <input checked="" type="checkbox"/> Múltiple	Valores de plantilla	<input type="text"/>
	<input type="checkbox"/> Número entero <input type="checkbox"/> Cadena de caracteres <input type="checkbox"/> Símbolo		Slot inverso		paciente- proceso
Propiedades lógicas	<input type="checkbox"/> Transitiva <input type="checkbox"/> Simétrica <input checked="" type="checkbox"/> Recursiva	Valores permitidos (rango)	ENTIDAD	Dominio	ACTIVIDAD
Propiedades meronímicas	<input type="checkbox"/> Funcional <input type="checkbox"/> Homeomérica <input type="checkbox"/> Separable <input type="checkbox"/> Simultánea				

Tabla 36. Implementación de la relación proceso-paciente.

e) Ejemplos

alicatado – azulejo

colocación – baldosa cerámica

enlucido – superficie

6.4.4.4 Relación proceso-instrumento

a) Definición

Relación entre un proceso y el instrumento empleado para llevarlo a cabo.

b) Clases conceptuales

Las clases conceptuales relacionadas pueden ser actividades (proceso) y entidades (instrumento), y generalmente los conceptos relacionados no pertenecen a la misma clase.

c) Propiedades de la relación

Relación no transitiva, no simétrica, y que puede presentar recursividad. Su relación inversa es la relación instrumento-proceso.

d) Implementación de la relación en *Protégé*

Nombre de la relación		proceso-	instrumento		
Tipo de valor	<input type="checkbox"/> Cualquiera <input type="checkbox"/> Booleano <input checked="" type="checkbox"/> Clase <input type="checkbox"/> Número con decimal <input type="checkbox"/> Instancia <input type="checkbox"/> Número entero <input type="checkbox"/> Cadena de caracteres <input type="checkbox"/> Símbolo	Cardinalidad	<input type="checkbox"/> Obligatorio <input checked="" type="checkbox"/> Múltiple	Valores de plantilla	<input type="text"/>
Propiedades lógicas	<input type="checkbox"/> Transitiva <input type="checkbox"/> Simétrica <input checked="" type="checkbox"/> Recursiva	Slot inverso	Instrumento-proceso	Valores por defecto	<input type="text"/>
Propiedades meronómicas	<input type="checkbox"/> Funcional <input type="checkbox"/> Homeomérica <input type="checkbox"/> Separable <input type="checkbox"/> Simultánea	Valores permitidos (rango)	ENTIDAD	Dominio	ACTIVIDAD

Tabla 37. Implementación de la relación proceso-instrumento.

e) Ejemplos

colocación en capa gruesa – cemento aditivado

control de resistencia al cuarteo – autoclave

colocación con junta – mortero

6.4.4.5 Relación proceso-estado

a) Definición

Relación entre un proceso y el estado final que alcanza el paciente de dicho proceso.

b) Clases conceptuales

Las clases conceptuales entre las que se establece esta relación son una actividad (proceso) y una propiedad o entidad (estado).

c) Propiedades de la relación

Se trata de una relación no transitiva, no simétrica y que no suele presentar recursividad. Su relación inversa es la relación estado-proceso.

d) Implementación de la relación en Protégé

Nombre de la relación		proceso-	estado		
Tipo de valor	<input type="checkbox"/> Cualquiera <input type="checkbox"/> Booleano <input checked="" type="checkbox"/> Clase <input type="checkbox"/> Número con decimal <input type="checkbox"/> Instancia <input type="checkbox"/> Número entero <input type="checkbox"/> Cadena de caracteres <input type="checkbox"/> Símbolo	Cardinalidad	<input type="checkbox"/> Obligatorio <input checked="" type="checkbox"/> Múltiple	Valores de plantilla	
		Slot inverso	estado-proceso	Valores por defecto	
Propiedades lógicas	<input type="checkbox"/> Transitiva <input type="checkbox"/> Simétrica <input type="checkbox"/> Recursiva	Valores permitidos (rango)	ENTIDAD PROPIEDAD ACTIVIDAD	Dominio	ENTIDAD PROPIEDAD ACTIVIDAD
Propiedades meronímicas	<input type="checkbox"/> Funcional <input type="checkbox"/> Homeomérica <input type="checkbox"/> Separable <input type="checkbox"/> Simultánea				

Tabla 38. Implementación de la relación proceso-estado.

e) Ejemplos

De momento no se han encontrado ejemplos de esta relación en el estudio empírico.

6.4.4.6 Relación proceso-método

a) Definición

Relación que se establece entre un proceso y el método que se emplea para llevarlo a cabo.

b) Clases conceptuales

Las clases conceptuales relacionadas suelen ser actividades.

c) Propiedades de la relación

Se trata de una relación no transitiva, no simétrica, y que puede presentar recursividad. Su relación inversa es la relación método-proceso.

d) Implementación en *Protégé*

Nombre de la relación		proceso-	método		
Tipo de valor	<input type="checkbox"/> Cualquiera <input type="checkbox"/> Booleano <input checked="" type="checkbox"/> Clase <input type="checkbox"/> Número con decimal <input type="checkbox"/> Instancia <input type="checkbox"/> Número entero <input type="checkbox"/> Cadena de caracteres <input type="checkbox"/> Símbolo	Cardinalidad	<input type="checkbox"/> Obligatorio <input checked="" type="checkbox"/> Múltiple	Valores de plantilla	
		Slot inverso	método-proceso	Valores por defecto	
Propiedades lógicas	<input type="checkbox"/> Transitiva <input type="checkbox"/> Simétrica <input checked="" type="checkbox"/> Recursiva	Valores permitidos (rango)	ACTIVIDAD	Dominio	ACTIVIDAD
Propiedades meronómicas	<input type="checkbox"/> Funcional <input type="checkbox"/> Homeomérica <input type="checkbox"/> Separable <input type="checkbox"/> Simultánea				

Tabla 39. Implementación de la relación proceso-método.

e) Ejemplos

De momento no se han encontrado ejemplos de esta relación en el corpus.

6.4.4.7 Relación causa-efecto

a) Definición

Relación que se establece entre una causa y el efecto que produce.

b) Clases conceptuales

Las clases conceptuales entre las que puede establecerse este tipo de relación son entidades, actividades y propiedades, y que no tienen que pertenecer necesariamente a la misma clase a ambos lados de la relación.

c) Propiedades de la relación

La relación causa-efecto es no simétrica y no necesariamente transitiva. Puede presentar recursividad, y su relación inversa es la relación efecto-causa.

d) Implementación de la relación en Protégé

Nombre de la relación		causa-efecto			
Tipo de valor	<input type="checkbox"/> Cualquiera <input type="checkbox"/> Booleano <input checked="" type="checkbox"/> Clase <input type="checkbox"/> Número con decimal <input type="checkbox"/> Instancia <input type="checkbox"/> Número entero <input type="checkbox"/> Cadena de caracteres <input type="checkbox"/> Símbolo	Cardinalidad	<input type="checkbox"/> Obligatorio <input checked="" type="checkbox"/> Múltiple	Valores de plantilla	<input type="text"/>
		Slot inverso	efecto-causa	Valores por defecto	<input type="text"/>
Propiedades lógicas	<input type="checkbox"/> Transitiva <input type="checkbox"/> Simétrica <input checked="" type="checkbox"/> Recursiva	Valores permitidos (rango)	ENTIDAD ACTIVIDAD PROPIEDAD	Dominio	ENTIDAD ACTIVIDAD PROPIEDAD
Propiedades meronímicas	<input type="checkbox"/> Funcional <input type="checkbox"/> Homeomérica <input type="checkbox"/> Separable <input type="checkbox"/> Simultánea				

Tabla 40. Implementación de la relación causa-efecto.

e) Ejemplos

fricción – abrasión

porosidad abierta – absorción de agua

calafateado – sellado hermético

6.4.4.8 Relación objeto-uso

a) Definición

Relación entre un objeto y la función o el uso para el que está destinado.

b) Clases conceptuales

Las clases conceptuales entre las cuales puede establecerse esta relación son entidades, actividades y propiedades, y que no necesariamente tienen que ser idénticas a ambos lados de la relación.

c) Propiedades de la relación

Esta relación es no simétrica y no transitiva. Puede presentar recursividad, y su relación inversa es la relación uso-objeto.

d) Implementación de la relación en *Protégé*

Nombre de la relación		objeto-uso			
Tipo de valor	<input type="checkbox"/> Cualquiera	Cardinalidad	<input type="checkbox"/> Obligatorio	Valores de plantilla	<input type="checkbox"/>
	<input type="checkbox"/> Booleano		<input checked="" type="checkbox"/> Múltiple		<input type="checkbox"/>
	<input checked="" type="checkbox"/> Clase	Slot inverso	uso-objeto	Valores por defecto	<input type="checkbox"/>
	<input type="checkbox"/> Número con decimal				<input type="checkbox"/>
	<input type="checkbox"/> Instancia				<input type="checkbox"/>
	<input type="checkbox"/> Número entero				<input type="checkbox"/>
	<input type="checkbox"/> Cadena de caracteres				<input type="checkbox"/>
	<input type="checkbox"/> Símbolo				<input type="checkbox"/>
	<input type="checkbox"/>				<input type="checkbox"/>
Propiedades lógicas	<input type="checkbox"/> Transitiva	Valores permitidos (rango)	ACTIVIDAD PROPIEDAD	Dominio	ENTIDAD
	<input type="checkbox"/> Simétrica				
	<input checked="" type="checkbox"/> Recursiva				
Propiedades meronómicas	<input type="checkbox"/> Funcional				
	<input type="checkbox"/> Homeomérica				
	<input type="checkbox"/> Separable				
	<input type="checkbox"/> Simultánea				

Tabla 41. Implementación de la relación objeto-uso.

e) Ejemplos

baldosa – revestimiento de suelos

capa de desolidarización – aislamiento del soporte

barniz – uso decorativo

6.4.5 Otras relaciones

Presentamos a continuación las relaciones fenómeno-medida, objeto-característica, la relación de sinonimia y la relación asociativa.

6.4.5.1 Relación fenómeno-medida

a) Definición

Relación que se establece entre un fenómeno o una característica y la unidad de medida empleada para expresar dicho fenómeno.

b) Clases conceptuales

Los conceptos relacionados pueden ser actividades, entidades y propiedades que generalmente no pertenecen a la misma clase conceptual.

c) Propiedades de la relación

Es una relación no transitiva, no simétrica y que puede presentar recursividad. Su relación inversa es la relación medida-fenómeno.

d) Implementación de la relación en Protégé

Nombre de la relación		fenómeno-	medida		
Tipo de valor	<input type="checkbox"/> Cualquiera <input type="checkbox"/> Booleano <input checked="" type="checkbox"/> Clase <input type="checkbox"/> Número con decimal <input type="checkbox"/> Instancia <input type="checkbox"/> Número entero <input type="checkbox"/> Cadena de caracteres <input type="checkbox"/> Símbolo	Cardinalidad	<input type="checkbox"/> Obligatorio <input checked="" type="checkbox"/> Múltiple	Valores de plantilla	<input type="text"/>
Propiedades lógicas	<input type="checkbox"/> Transitiva <input type="checkbox"/> Simétrica <input checked="" type="checkbox"/> Recursiva	Slot inverso	<input type="text" value="medida-fenómeno"/>	Valores por defecto	<input type="text"/>
Propiedades meronímicas	<input type="checkbox"/> Funcional <input type="checkbox"/> Homeomérica <input type="checkbox"/> Separable <input type="checkbox"/> Simultánea	Valores permitidos (rango)	<input type="text" value="ENTIDAD"/>	Dominio	<input type="text" value="ENTIDAD ACTIVIDAD PROPIEDAD"/>

Tabla 42. Implementación de la relación fenómeno-medida.

e) Ejemplos

dureza Mohs – escala de Mohs

carga de rotura – newtons

6.4.5.2 Relación objeto-característica

a) Definición

Relación entre una entidad y una característica que define a dicha entidad.

b) Clases conceptuales

Generalmente, las clases relacionadas son entidades y propiedades, y esta relación suele establecerse entre conceptos pertenecientes a clases distintas.

c) Propiedades de la relación

Esta relación no tiene por qué ser transitiva (aunque podría serlo), no es simétrica y su relación inversa es la relación característica-objeto.

d) Implementación en *Protégé*

Nombre de la relación		objeto-	característica		
Tipo de valor	<input type="checkbox"/> Cualquiera <input type="checkbox"/> Booleano <input checked="" type="checkbox"/> Clase <input type="checkbox"/> Número con decimal <input type="checkbox"/> Instancia <input type="checkbox"/> Número entero <input type="checkbox"/> Cadena de caracteres <input type="checkbox"/> Símbolo	Cardinalidad	<input type="checkbox"/> Obligatorio <input checked="" type="checkbox"/> Múltiple	Valores de plantilla	<input type="text"/>
		Slot inverso	característica-objeto	Valores por defecto	<input type="text"/>
Propiedades lógicas	<input type="checkbox"/> Transitiva <input type="checkbox"/> Simétrica <input checked="" type="checkbox"/> Recursiva	Valores permitidos (rango)	PROPIEDAD	Dominio	ENTIDAD
Propiedades meronómicas	<input type="checkbox"/> Funcional <input type="checkbox"/> Homeomérica <input type="checkbox"/> Separable <input type="checkbox"/> Simultánea				

Tabla 43. Implementación de la relación objeto-característica.

e) Ejemplos

baldosa cerámica – absorción de agua

azulejo – resistencia a la helada

6.4.5.3 Relación asociativa

a) Definición

Relación que se establece por la correlación entre dos o más conceptos, cuya naturaleza no se corresponde con ninguno de los tipos de relación identificados anteriormente.

b) Clases conceptuales

Este tipo de relación puede establecerse entre entidades, propiedades y actividades, y que no necesariamente tienen que pertenecer a la misma clase para estar relacionadas.

c) Propiedades de la relación

Es una relación no necesariamente simétrica ni transitiva. Puede presentar recursividad, y su relación inversa sería ella misma.

d) Implementación en Protégé

Nombre de la relación		Relación asociativa			
Tipo de valor	<input type="checkbox"/> Cualquiera <input type="checkbox"/> Booleano <input checked="" type="checkbox"/> Clase <input type="checkbox"/> Número con decimal <input type="checkbox"/> Instancia <input type="checkbox"/> Número entero <input type="checkbox"/> Cadena de caracteres <input type="checkbox"/> Símbolo	Cardinalidad	<input type="checkbox"/> Obligatorio <input checked="" type="checkbox"/> Múltiple	Valores de plantilla	<input type="text"/>
		Slot inverso	<input type="text"/>	Valores por defecto	<input type="text"/>
Propiedades lógicas	<input type="checkbox"/> Transitiva <input type="checkbox"/> Simétrica <input checked="" type="checkbox"/> Recursiva	Valores permitidos (rango)	ENTIDAD ACTIVIDAD PROPIEDAD	Dominio	ENTIDAD ACTIVIDAD PROPIEDAD
Propiedades meronímicas	<input type="checkbox"/> Funcional <input type="checkbox"/> Homeomérica <input type="checkbox"/> Separable <input type="checkbox"/> Simultánea				

Tabla 44. Implementación de la relación asociativa.

e) Ejemplos

adherendo – adhesivo

ensayo – propiedad de los materiales

embaldosado directo – capa de desolidarización

6.4.5.4 Relación de sinonimia

a) Definición

Relación que establece la equivalencia conceptual entre dos o más denominaciones referidas al mismo concepto.

b) Clases conceptuales

Los conceptos relacionados pueden ser entidades, propiedades o actividades.

c) Propiedades de la relación

Es una relación simétrica y transitiva, que además puede presentar recursividad. Esta relación es inversa de sí misma.

d) Implementación en Protégé

Su implementación en *Protégé* no se realiza mediante *slots*, sino que se explicita mediante el componente *Instance*, que permite reflejar cada una de las denominaciones consideradas sinónimas para un mismo concepto.

e) Ejemplos

cuerpo – soporte

capa aislante – capa de aislamiento

canto – arista

6.5 Recapitulación y cuadro resumen del catálogo de relaciones conceptuales

En este capítulo hemos presentado nuestro catálogo de relaciones conceptuales y hemos propuesto un modelo de formalización e implementación de las mismas en el editor de ontologías *Protégé*. Veamos cuáles son las principales aportaciones de este capítulo:

En primer lugar, nuestro catálogo contempla las relaciones más comunes que se establecen entre los conceptos de los ámbitos especializados. En efecto, aunque el ámbito de estudio en el que se ha desarrollado nuestro trabajo empírico es la industria cerámica, el catálogo de relaciones propuesto no se limita a las relaciones específicas de este ámbito, sino que tiene vocación de exhaustividad y por tanto es válido para cualquier área de especialidad. Prueba de ello es el hecho de que el catálogo refleja la complejidad de las relaciones meronímicas que, salvo excepciones, no suele tenerse en cuenta en el trabajo terminológico práctico, a pesar de su conveniencia de cara a la estructuración conceptual. Aun cuando consideramos que el catálogo es bastante completo, nuestra propuesta de formalización posibilita su ampliación en caso necesario para dar cuenta de las relaciones específicas del área temática, cuyo acomodo en el catálogo podría hacerse fácilmente describiendo nuevas relaciones específicas con los mismos parámetros.

En segundo lugar, la estructuración de las relaciones conceptuales del catálogo en cinco grupos (relaciones lógicas, relaciones meronímicas, relaciones secuenciales, relaciones argumentales y circunstanciales y otras relaciones) permite presentarlas de forma estructurada en función de su temática. Creemos que de este modo es posible superar las clasificaciones en las que las relaciones no

jerárquicas se suelen ofrecer como listas más o menos extensas de relaciones de diversa naturaleza agrupadas con escaso orden aparente.

En tercer lugar, nuestra propuesta de formalización de las relaciones conceptuales permite representarlas de un modo estructurado, teniendo en cuenta aspectos tales como las clases de conceptos que intervienen en la relación y sus propiedades matemáticas y rasgos distintivos (en el caso de las relaciones meronímicas). El hecho de utilizar el formalismo $a R b, n$ permite definir las relaciones conceptuales con vistas a su implementación en un editor de ontologías.

Asimismo, creemos que la denominación de las relaciones en función de los argumentos que aparecen a uno y otro lado de la relación hace que el catálogo sea más transparente para el usuario potencial de una base de conocimiento.

Otro aspecto innovador de nuestra propuesta de formalización e implementación de las relaciones conceptuales reside en que hemos tenido en cuenta la direccionalidad de las mismas, para lo que se han propuesto relaciones inversas que permiten relacionar los conceptos en los dos sentidos posibles.

Por último, en el catálogo se ha incluido la relación de sinonimia como una relación que, si bien no es estrictamente conceptual, conviene tener en consideración en las obras terminográficas para dar cuenta de la riqueza expresiva de las terminologías. Además, como veremos en el siguiente capítulo, el hecho de contemplar los sinónimos como instancias de las clases nos permitirá recuperar la información a partir de cualquiera de las posibilidades denominativas para un mismo concepto.

En definitiva, la implementación formal de las relaciones conceptuales en una base de conocimiento creada con un editor de ontologías posibilita que la información que dichas relaciones vehiculan pueda ser utilizada en aplicaciones informáticas.

Para terminar, presentamos una tabla que resume las relaciones conceptuales de nuestro catálogo. En ella se indican los siguientes aspectos de cada relación:

- a) nombre de la relación conceptual,
- b) descripción de la relación en lenguaje natural,

- c) dominio de la relación (clases de conceptos a partir de los cuales puede establecerse),
- d) rango de la relación (clases de conceptos con los que puede establecerse)
- e) transitividad
- f) simetría
- g) recursividad
- h) relación inversa
- i) ejemplos

TABLA RESUMEN DE LAS RELACIONES CONCEPTUALES

Relación	Definición	Dominio	Rango	Transitividad	Simetría	Recursividad	Relación inversa	Ejemplos
Superordinado-subordinado	Relación que se establece entre un concepto a que posee un grado de abstracción superior y un concepto b que se sitúa un nivel por debajo de él en una jerarquía.	Entidades Actividades Propiedades	Entidades Actividades Propiedades	Sí	No	Sí	Subordinado-superordinado	acabado superficial – engobe; coeficiente de fricción – coeficiente de fricción dinámico; calafateado – calafateado elastomérico
Subordinado-subordinado	Relación que se establece entre dos conceptos que se encuentran situados en el mismo nivel de abstracción por debajo de un superordinado común.	Entidades Actividades Propiedades	Entidades Actividades Propiedades	Sí	Sí	Sí	Ella misma	cubrecantos – escuadra; curvatura – grosor; colocación con junta – colocación sin junta
Componente funcional-objeto	Relación que se establece entre los componentes funcionales de un objeto y el	Entidades	Entidades	Sí	No	Sí	Objeto-componente funcional	arista – baldosa; olambrilla – baldosín catalán; esmalte – producto acabado

Relación	Definición	Dominio	Rango	Transitividad	Simetría	Recursividad	Relación inversa	Ejemplos
	objeto del que forman parte.							
Miembro-colección	Relación que se establece entre partes que establecen una relación de proximidad espacial o conexión social con respecto al todo, sin necesidad de que cumplan una función específica ni que estén dispuestos estructuralmente de un modo específico.	Entidades	Entidades	No	No	Sí	Colección-miembro	Mosaico – tesela
Porción-masa	Relación entre una parte y el todo donde la naturaleza de la porción es la misma que la de la masa.	Entidad	Entidad	Sí	No	Sí	Masa-porción	--
Material-objeto	Relación entre una parte intrínseca y el	Entidad	Entidad	Sí	No	Sí	Objeto-material	sustancia viscosa – adhesivo; frita –

Relación	Definición	Dominio	Rango	Transitividad	Simetría	Recursividad	Relación inversa	Ejemplos
	todo al que pertenece, del que no puede separarse por ser consustancial al mismo.							esmalte; arena natural – base de arena
Etapa-proceso	Relación que se establece entre cada una de las fases de una actividad y dicha actividad en conjunto.	Actividad	Actividad	Sí	No	Sí	Proceso-etapa	etapa de gran fuego – cocción; desolidarización – colocación; ensayo – control de calidad
Característica-actividad	Relación que se establece entre las características de una actividad y el proceso complejo en el que se inscriben.	Propiedad	Actividad	No	No	Sí	Actividad- característica	--
Lugar-área	Relación que se establece entre un área espacial y las localizaciones concretas dentro de esa área.	Entidad	Entidad	Sí	No	No	Área-lugar	--

Relación	Definición	Dominio	Rango	Transitividad	Simetría	Recursividad	Relación inversa	Ejemplos
Parte-parte	Relación que se establece entre las partes que juntas componen un todo.	Entidad Actividad	Entidad Actividad	Sí	Sí	Sí	Ella misma	cara vista – arista; esmalte – soporte; encolado – dosificación; engobe – vidriado
Concepto-lugar	Relación que se establece entre dos conceptos que se ubican juntos en el espacio.	Entidad Actividad	Entidad	Sí	Sí	Sí	Lugar-concepto	cocción-horno; adhesivo-adherendo
Concepto-lugar al que se dirige	Relación que se establece entre dos conceptos para indicar la dirección en la que se produce una actividad determinada.	Entidad Actividad	Entidad	Sí	Sí	Sí	Lugar al que se dirige-concepto	--
Concepto-concepto simultáneo	Relación que se establece entre dos actividades o entidades que se ubican en el mismo espacio de tiempo.	Entidad Actividad	Entidad Actividad	Sí	Sí	Sí	Ella misma	--
Concepto anterior-concepto	Relación que se establece entre dos conceptos	Entidad Actividad	Entidad Actividad	Sí	No	Sí	Concepto posterior-concepto	--

Relación	Definición	Dominio	Rango	Transitividad	Simetría	Recursividad	Relación inversa	Ejemplos
posterior	que se sitúan en el tiempo de forma consecutiva.						anterior	
Proceso-agente	Relación que se establece entre un proceso y la entidad que lleva a cabo dicho proceso.	Actividad	Entidad	No	No	Sí	Agente-proceso	--
Proceso-producto	Relación que se establece entre el proceso y el producto final de dicho proceso.	Actividad	Entidad	No	No	Sí	Producto-proceso	moldeo – baldosa cerámica; reacción química – adhesivo de resina de reacción; doble encolado – macizado
Proceso-paciente	Relación que se establece entre un proceso y la entidad sobre la que se lleva a cabo.	Actividad	Entidad	No	No	Sí	Paciente-proceso	alcatado – azulejo; colocación – baldosa cerámica; enlucido – superficie
Proceso-instrumento	Relación entre un proceso y el instrumento empleado para	Actividad	Entidad	No	No	Sí	Instrumento-proceso	colocación en capa gruesa – cemento aditivado;

Relación	Definición	Dominio	Rango	Transitividad	Simetría	Recursividad	Relación inversa	Ejemplos
	llevarlo a cabo.							control de resistencia al cuarteo – autoclave; colocación con junta – mortero
Proceso-estado	Relación entre un proceso y el estado final que alcanza el paciente de dicho proceso.	Actividad	Entidad Propiedad	No	No	No	Estado-proceso	--
Proceso-método	Relación que se establece entre un proceso y el método que se emplea para llevarlo a cabo.	Actividad	Actividad	No	No	Sí	Método-proceso	fabricación – extrusión
Causa-efecto	Relación que se establece entre una causa y el efecto que produce.	Entidad Actividad Propiedad	Entidad Actividad Propiedad	No	No	Sí	Efecto-causa	fricción – abrasión; porosidad abierta – absorción de agua; calafateado – sellado hermético
Objeto-uso	Relación entre un objeto y la función o el uso para el que está	Entidad	Actividad Propiedad	No	No	Sí	Uso-objeto	baldosa – revestimiento de suelos; capa de

Relación	Definición	Dominio	Rango	Transitividad	Simetría	Recursividad	Relación inversa	Ejemplos
	destinado.							desolidarización – aislamiento del soporte barniz – uso decorativo
Fenómeno-medida	Relación que se establece entre un fenómeno o una característica y la unidad de medida empleada para expresar dicho fenómeno o característica.	Entidad Actividad Propiedad	Entidad	No	No	Sí	Medida-fenómeno	dureza mohs – escala de mohs; carga de rotura – newtons (n)
Objeto-característica	Relación entre una entidad y una característica que lo define.	Entidad	Propiedad	Sí	No	Sí	Característica-objeto	baldosa cerámica – absorción de agua; azulejo – resistencia a la helada
Relación asociativa	Relación que se establece por la correlación entre dos o más conceptos cuya naturaleza no se corresponde con ninguno de los	Entidad Actividad Propiedad	Entidad Actividad Propiedad	No	No	Sí	Ella misma	adherendo – adhesivo; ensayo – propiedad de los materiales; embaldosado directo – capa de

Relación	Definición	Dominio	Rango	Transitividad	Simetría	Recursividad	Relación inversa	Ejemplos
	tipos de relación identificados anteriormente.							desolidarización
Relación de sinonimia	Relación que se establece por equivalencia entre dos o más denominaciones que se refieren al mismo concepto.	Entidad Actividad Propiedad	Entidad Actividad Propiedad	Sí	Sí	Sí	Ella misma	cuerpo – soporte; capa aislante – capa de aislamiento; canto – arista

Tabla 45. Tabla resumen del catálogo de relaciones conceptuales.

**CAPÍTULO 7: INTRODUCCIÓN DE DATOS EN EL
EDITOR DE ONTOLOGÍAS Y RECUPERACIÓN DE LA
INFORMACIÓN**

7 Introducción de datos en el editor de ontologías y recuperación de la información

En el capítulo anterior se ha presentado el catálogo de relaciones conceptuales y su formalización e implementación en el editor de ontologías *Protégé*. Basándonos en estas propuestas hemos desarrollado la base de conocimiento *Ontoceram*, en la que se han incluido conceptos sobre los productos cerámicos acabados y sus conceptos relacionados, que nos servirá para evaluar la utilidad de nuestras propuestas para el almacenamiento y recuperación de la información sobre relaciones conceptuales.

En este capítulo, explicamos en primer lugar (apartado 7.1) cómo se ha llevado a cabo el proceso de introducción de datos en *Ontoceram*. Se detalla el proceso de extracción manual de los conceptos y sus relaciones, así como el modo en que se han estructurado los conceptos en la base de conocimiento. A continuación, se describe *Ontoceram* desde el punto de vista cuantitativo. Como producto de este proceso, se presenta un protocolo para la introducción de datos en el que se detalla la metodología seguida para la extracción de conceptos y relaciones y su introducción en *Ontoceram*, lo que servirá de ayuda para el desarrollo posterior de la base de conocimiento.

En la sección 7.2 se analizan las posibilidades de recuperación de la información sobre relaciones conceptuales estructurada en la base de conocimiento *Ontoceram* generada con el editor de ontologías *Protégé*. En el apartado 7.2.1 se proponen una serie de consultas sobre las relaciones conceptuales que se han diseñado, con el fin de valorar la idoneidad de utilizar un editor de ontologías para almacenar la información sobre las relaciones conceptuales. De este modo comprobaremos si el modelo de estructuración propuesto en el capítulo anterior resulta adecuado para recuperar dicha información. Para ejemplificar cómo pueden ejecutarse estas consultas, en el apartado 7.2.2 se presentan algunos ejemplos sobre cómo se pueden aprovechar la estructuración de las clases en forma taxonómica en el editor de *classes* de *Protégé* y los *plug-ins* de consultas proporcionados por *Protégé* (*Queries Tab* y *PAL Queries Tab*).

La introducción de datos y la ejecución de consultas en *Ontoceram* nos han permitido reflexionar sobre algunas adaptaciones del editor de ontologías

Protégé que podrían mejorar su funcionalidad para el trabajo terminográfico. Estas propuestas se recogen en el apartado 7.3.

Por último, en el apartado 7.4 se resumen los resultados presentados en este capítulo.

7.1 *Proceso de introducción de datos en Ontoceram*

Uno de los objetivos aplicados de la presente investigación es desarrollar una base de conocimiento que nos permita evaluar la adecuación del modelo propuesto para la representación de los conceptos y relaciones en un editor de ontologías. En este apartado describimos el proceso llevado a cabo para la introducción de información en *Ontoceram*.

a) Limitación del área temática

En primer lugar, y debido a que el ámbito de la industria cerámica es muy amplio, se ha limitado el campo para el estudio empírico. La base de datos terminológicos *Cerámica* (descrita en el apartado 5.2.4), utilizada como fuente de referencia sobre los conceptos de la cerámica, contiene 4.617 términos, y la introducción de todos ellos en nuestra base de conocimiento excede los límites de la presente tesis doctoral. Por ello se ha acotado el campo de trabajo, tomando como punto de partida esta base de datos terminológica, estructurada de acuerdo con un árbol de campo desarrollado teniendo en cuenta todas las fases del proceso de fabricación de los productos cerámicos. En esta investigación nos hemos centrado en los conceptos de la rama *1.1.3. Tipos de producto acabado* del árbol de campo, puesto que este tipo de productos son el resultado final de los procesos cerámicos, y por tanto son susceptibles de establecer relaciones con conceptos de otras ramas, como los procesos de fabricación, las materias primas, los productos intermedios o las características, entre otros.

A partir de la base de datos *Cerámica* se han extraído los términos correspondientes a la rama *1.1.3. Tipos de producto acabado*, lo que nos ha proporcionado una lista inicial de 340 términos. Esta lista se ha depurado para eliminar términos como *aparejo a soga* o *brillo especular*, que no son estrictamente productos acabados, con lo que se ha reducido el listado inicial a 300 conceptos. Dicho listado se ha ido ampliando con nuevos conceptos extraídos durante la fase de análisis de los conceptos y sus relaciones, que en un principio no aparecían recogidos en la base de datos terminológicos *Cerámica*.

b) Estructuración de los conceptos en *Ontoceram*

Como se ha explicado en la sección 6.3 del capítulo anterior, proponemos implementar los conceptos correspondientes a la lista de productos cerámicos acabados mediante el componente *Class* del editor *Protégé*. En los párrafos siguientes explicamos cómo se han organizado los conceptos, basándonos en el árbol de campo de la cerámica presentado en el apartado 5.2.5, el modelo conceptual de Sager y Kageura (1994) presentado en 2.2.1.1 y la organización de los niveles superiores de la ontología Mikrokosmos (Nirenburg y Raskin, 2004), presentada en el apartado 3.4.2.1.

La observación del árbol de campo nos ha permitido identificar los tipos de conceptos que se manejan en el ámbito de la industria cerámica. En ella, una o varias materias primas se someten a determinados procesos de fabricación empleando una maquinaria específica. El resultado de dichos procesos pueden ser productos acabados o productos intermedios, que a su vez se someten a nuevos procesos de fabricación para dar lugar a productos acabados que poseen unas características determinadas. De ellas dependerá el proceso de colocación y el uso definitivo del producto. Esta síntesis de los procesos de la industria cerámica nos permite identificar los siguientes tipos de conceptos: *materias primas*, *productos intermedios*, *maquinaria*, *procesos de fabricación*, *productos acabados*, *características*, *colocación* y *uso*, todos ellos representados en las diversas ramas del árbol de campo.

Estos tipos de conceptos se adecuan a las clases conceptuales propuestas por Sager y Kageura (entidades, actividades y propiedades). Así, de acuerdo con este modelo conceptual, las materias primas, los productos intermedios y acabados, la maquinaria y el lugar de colocación serían entidades. Por su parte, los procesos de fabricación y colocación y el uso pertenecen a la clase de las actividades, y las características se ubican dentro de la clase conceptual de las propiedades. Como vemos, los conceptos del ámbito de la cerámica se ajustan al modelo de clases conceptuales adoptado para la formalización de las relaciones conceptuales.

Los conceptos representados mediante el componente *Class* en el editor de ontologías *Protégé* conforman una estructura jerárquica en la que se mantiene la herencia. Para organizar los conceptos en la jerarquía, hemos utilizado la estructuración de los conceptos propuesta para la ontología Mikrokosmos

(Nirenburg y Raskin, 2004, revisado en el apartado 3.4.2.1), que se ajusta con algunas adaptaciones al modelo conceptual de Sager y Kageura. Este mismo modelo ya ha sido empleado con éxito en el desarrollo de aplicaciones terminológicas basadas en ontologías como las de los proyectos ONCOTERM y Genoma-KB (presentadas en los apartados 3.3.1 y 3.3.4, respectivamente). Por ello, proponemos reutilizar una propuesta de organización conceptual ya probada en estos trabajos, y de este modo dejamos la puerta abierta a que nuestra base de conocimiento pueda ser reutilizada en otras aplicaciones.

En la ontología Mikrokosmos existe un concepto raíz denominado ALL, y de él dependen los conceptos OBJETO, ACTIVIDAD y PROPIEDAD. En *Protégé*, el concepto raíz del que parte la jerarquía de clases se denomina :THING y se emplea por defecto como origen de toda la jerarquía. Dependiendo directamente de :THING, hemos situado los conceptos superiores ENTIDAD, ACTIVIDAD y PROPIEDAD. Por debajo de ellos, se ha mantenido la estructura propuesta de la ontología Mikrokosmos hasta el tercer nivel conceptual. En la Fig. 35 se observa cómo se han introducido en *Ontoceram* los conceptos de los niveles superiores de la ontología Mikrokosmos.

Con el fin de distinguir los conceptos tomados de la ontología Mikrokosmos de los introducidos a partir del estudio empírico, los procedentes de Mikrokosmos se han marcado tipográficamente con mayúsculas, mientras que los derivados del análisis se escriben con mayúscula inicial, como hemos explicado en el apartado 6.3 del capítulo anterior.

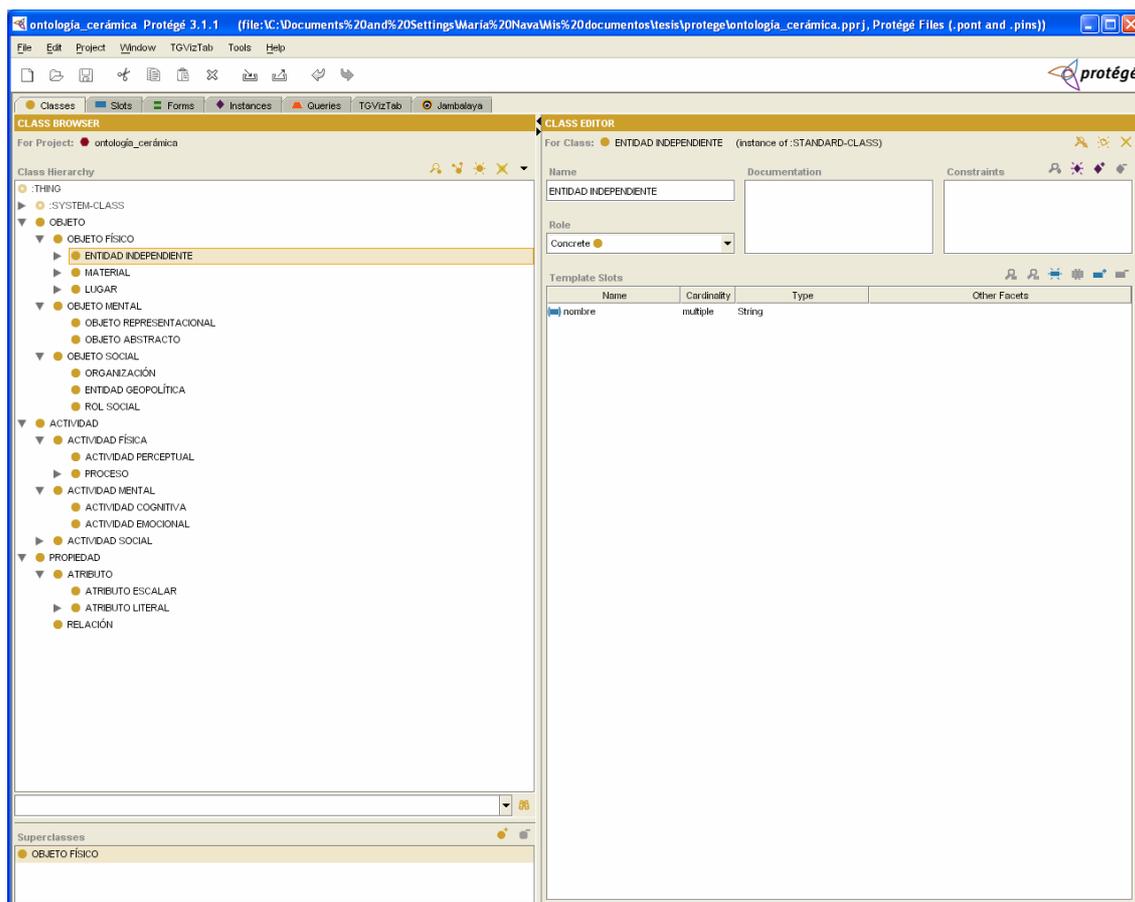


Fig. 35. Estructura de la ontología Mikrokosmos implementada en *Ontoceram*.

Para mayor claridad, a continuación reproducimos la organización de los conceptos propuesta para la ontología Mikrokosmos, situada en *Ontoceram* por debajo del concepto raíz `:THING`⁶³:

1. ENTIDAD

a. ENTIDAD FÍSICA

i. MATERIAL

ii. ENTIDAD INDEPENDIENTE

iii. LUGAR

b. ENTIDAD MENTAL

i. ENTIDAD

REPRESENTACIONAL

ii. ENTIDAD ABSTRACTA

⁶³ En la ontología *Mikrokosmos* se habla de *OBJECTS*. En nuestra implementación, con el fin de acercar las propuestas de Nirenburg y Raskin (2004) y la de Sager y Kageura (1994), hemos cambiado esta denominación por *ENTIDAD*.

- c. ENTIDAD SOCIAL
 - i. ORGANIZACIÓN
 - ii. ENTIDAD GEOPOLÍTICA
 - iii. ROL SOCIAL
- 2. ACTIVIDAD
 - a. ACTIVIDAD FÍSICA
 - i. ACTIVIDAD PERCEPTUAL
 - ii. PROCESO
 - b. ACTIVIDAD MENTAL
 - i. ACTIVIDAD COGNITIVA
 - ii. ACTIVIDAD EMOCIONAL
 - c. ACTIVIDAD SOCIAL
- 3. PROPIEDAD
 - a. ATRIBUTO
 - i. ATRIBUTO ESCALAR
 - ii. ATRIBUTO LITERAL
 - b. RELACIÓN

Los conceptos relacionados con los productos cerámicos extraídos a partir del trabajo empírico se han ubicado en la jerarquía de *Classes* bajo estos conceptos superiores, conformando la jerarquía de conceptos en *Ontoceram*. Por ejemplo, el concepto *Producto cerámico acabado* se ubica como subordinado de ENTIDAD INDEPENDIENTE, los distintos procesos de fabricación de las piezas cerámicas (*Moldeo, Esmaltado, etc.*) se consideran subordinados de PROCESO y las características de los productos acabados (como por ejemplo, *Resistencia a la heladicidad*) se sitúan dependiendo de ATRIBUTO LITERAL.

c) Introducción de las relaciones conceptuales que establecen los productos cerámicos acabados

Una vez introducidos los conceptos sobre los productos cerámicos acabados en la jerarquía de *Classes* del editor, se ha procedido a establecer las relaciones conceptuales y los conceptos con los que se vinculan. Para identificar las relaciones se ha seguido el proceso descrito en el estudio inicial que hemos llevado a cabo para validar el catálogo de relaciones conceptuales (presentado en detalle en el apartado 6.2.3).

Así, a partir de la lista de trescientos productos cerámicos introducidos en *Ontoceram*, se ha analizado su definición en los diccionarios especializados y los contextos definitorios de la base terminológica *Cerámica*. En dichas definiciones hemos identificado los conceptos relacionados con cada concepto buscado. La información sobre conceptos relacionados contenida en los diccionarios y la base terminológica se ha completado con la extraída a partir de las concordancias de cada término en el corpus *TXT CERAM* obtenidas con el programa *WordSmith Tools*. Mediante este análisis hemos identificado los conceptos relacionados con cada producto cerámico acabado de la lista y hemos determinado la naturaleza de la relación que los vincula.

La jerarquía de conceptos se ha enriquecido a medida que se han ido identificando nuevos conceptos que aún no estaban incluidos en la base de conocimiento.

Las relaciones se ha implementado siguiendo la propuesta presentada en el apartado 6.4, es decir, se han incorporado por medio del componente *Slot* que representa cada relación, o bien mediante la inclusión de cada concepto bajo su superordinado en la jerarquía de clases

En cuanto al orden en el que se ha llevado a cabo el análisis de los conceptos de la lista de productos cerámicos acabados, éste se inició a partir de *Pieza cerámica*, ya que consideramos que se trata de un concepto de base en el sentido propuesto por Otman (1996), es decir, se trata de un concepto situado en un nivel intermedio en la jerarquía, que ofrece máxima información conceptual, y a partir del cual puede desarrollarse el análisis de los superordinados y subordinados. Esto nos ha permitido explotar el mecanismo de herencia entre clases en el editor de ontologías. Así, los conceptos subordinados de *Pieza cerámica* heredan las relaciones asignadas a este concepto, y por tanto basta con restringir, ampliar o modificar los valores de cada relación en la clase que representa al concepto subordinado. De este modo conseguimos desarrollar la base de datos conceptual de un modo económico aprovechando los mecanismos de herencia, es decir, sin necesidad de introducir de nuevo todas las relaciones conceptuales para cada concepto.

Este proceso ha dado como resultado una estructura jerárquica de conceptos en la que las relaciones conceptuales se hacen explícitas según el modelo propuesto en el capítulo 6.

En la Fig. 36 observamos una captura de pantalla que muestra el concepto *Baldosa cerámica*. El concepto representado en *Protégé* por la clase *Baldosa cerámica* es un subordinado directo de *Pieza cerámica*, concepto con el que mantiene una relación subordinado-superordinado. En la parte derecha de la pantalla vemos que se trata de una *class* concreta en el sentido de *Protégé* (es decir, que puede poseer *Instances*), y a esta *clasa* se le han asignado una serie de *slots* de relación que permiten vincularla con otros conceptos contenidos en *Ontoceram*. Los *slots* de relación atribuidos a *Baldosa cerámica* son *concepto-concepto simultáneo*, *concepto-lugar*, *objeto-característica*, *objeto-componente*, *objeto-material*, *objeto-uso* y *paciente-proceso*, cuyos valores son otros conceptos representados mediante el componente *Class* en la base de datos. Se ha asignado asimismo un *slot* denominado *nombre*, que es el que nos permitirá definir las diversas denominaciones del concepto por medio del componente *Instance*, como veremos en la siguiente fase.

En todos los casos los *slots* de relación se han heredado de la clase superordinada *Pieza cerámica*, si bien sus valores se han restringido para limitarlos a los conceptos relacionados con *Baldosa cerámica* identificados durante el análisis. Así, por ejemplo, para la relación definida por el *slot objeto-componente*, se han añadido en el nivel de la clase los siguientes valores de conceptos relacionados: *Soporte*, *Cara vista*, *Cara posterior*, *Esmalte*, *Membrana de separación*, *Nervio de extrusión* y *Engobe*. Estos conceptos también se han incluido en su lugar correspondiente en la jerarquía de *classes* de la base *Ontoceram*.

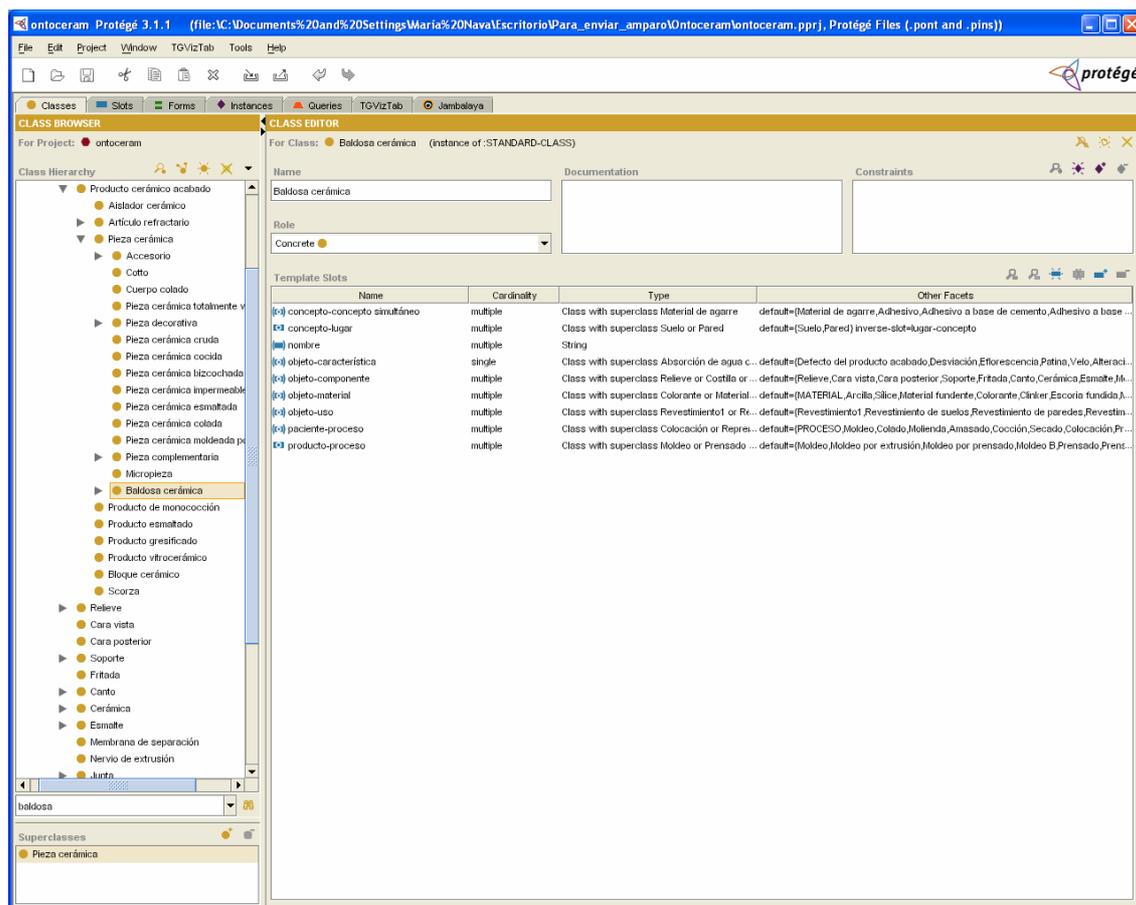


Fig. 36. Representación de la clase *Baldosa cerámica* en *Ontoceram*.

d) Introducción de las denominaciones de los conceptos mediante el componente *Instance*

Una vez incluidos en *Ontoceram* los productos cerámicos y sus conceptos relacionados, se ha procedido a incorporar las distintas denominaciones que se emplean para referirse a los conceptos a través del componente *Instance* del editor *Protégé*. Como se ha explicado en el apartado 6.3.1.2, este componente se emplea en el presente trabajo para introducir las variantes denominativas y sinónimos. Cada *Instance* se corresponde con una denominación para el concepto del que depende. Esto nos permite, por un lado, reflejar las distintas posibilidades denominativas de los conceptos (sinónimos, variantes, abreviaturas y siglas) y por otro, efectuar consultas en la base de conocimiento mediante el *plug-in Queries Tab*, tal y como se verá en el apartado 7.2.1.

Por ejemplo, para la clase *Baldosa cerámica* se han introducido cuatro posibles denominaciones, a saber: *baldosa cerámica*, *baldosa*, *baldosín* y *rasilla*, que se emplean para referirse al concepto *Baldosa cerámica* en los textos

consultados. En la Fig. 37 se muestran las cuatro instancias almacenadas en la base de datos conceptual *Ontoceram* que dependen de *Baldosa cerámica*:

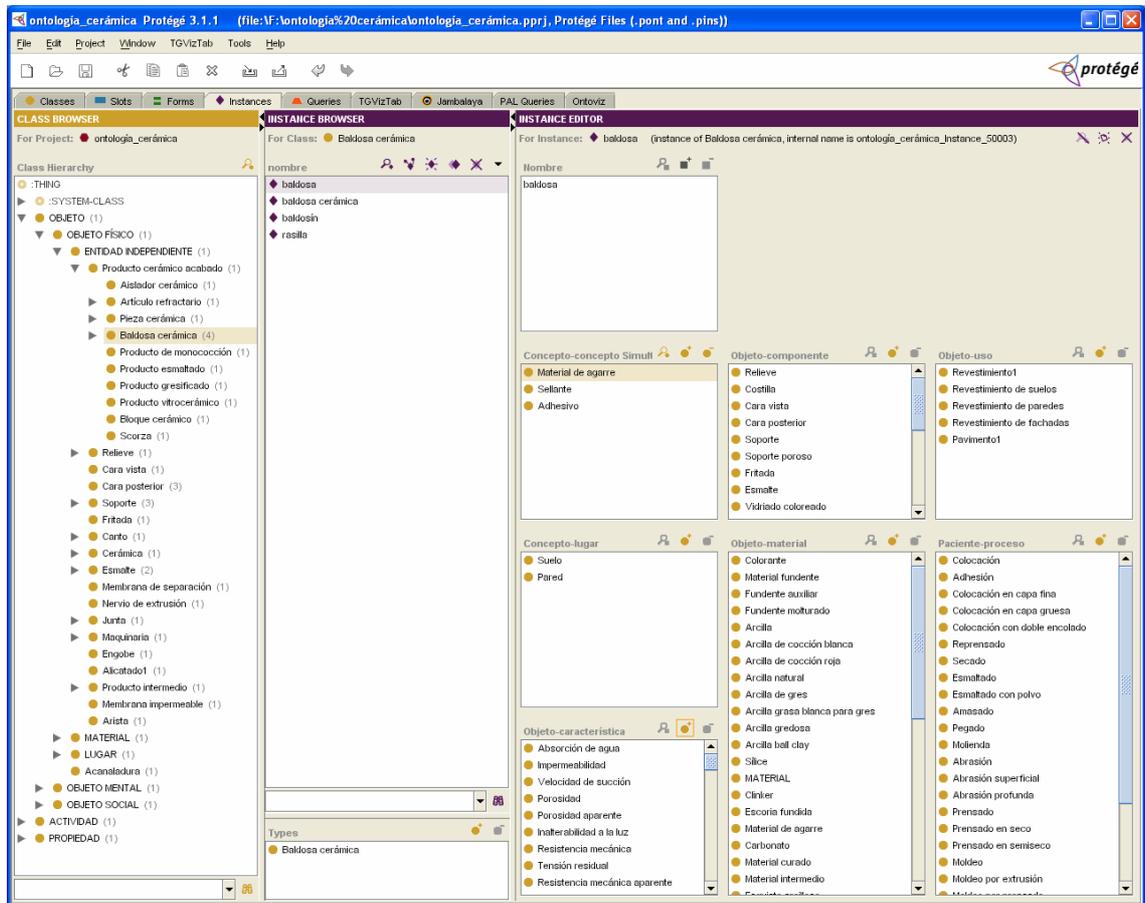


Fig. 37. Instancias para la clase *Baldosa cerámica*.

e) Descripción cuantitativa de la base de conocimiento *Ontoceram*

Al final del proceso de introducción de conceptos, relaciones y denominaciones hemos obtenido como resultado una base de datos conceptual en la que las relaciones se hacen explícitas de acuerdo con el modelo presentado en el capítulo 6. Esta base de conocimiento se ofrece para su consulta en el CD anexo a esta tesis doctoral.

Durante la elaboración del presente trabajo empírico se han almacenado en *Ontoceram* 621 conceptos a través del componente *Class*. La mayoría de ellos son productos cerámicos acabados, si bien se incluyen otros conceptos relacionados como los materiales, los procesos de fabricación, los usos y las características de los mismos. El número total de denominaciones incluidas mediante el componente *Instances* asciende a 668.

En cuanto al número de *slots* empleados en nuestra base de conocimiento, se han definido un total de 58. La mayoría de ellos (46) son *slots* que expresan relaciones, que reflejan las contenidas en el catálogo presentado en el capítulo 6 y sus inversas, en caso de que existan.

El resto de los *slots* se han definido para incorporar a los *slots* de relación las facetas correspondientes a sus propiedades matemáticas y meronímicas, y para permitir la introducción de denominaciones mediante las instancias. Las relaciones superordinado-subordinado y su inversa (subordinado-superordinado) no se han introducido como *slots*, sino que se hacen explícitas a través de la jerarquía de clases.

En la Tabla 46 se recogen los datos cuantitativos del contenido de *Ontoceram*:

<i>Clases (conceptos)</i>	621
<i>Slots</i>	58
<i>Slots de relación</i>	46
<i>Instancias (denominaciones)</i>	668

Tabla 46. Descripción cuantitativa de la base de conocimiento *Ontoceram*.

f) Tratamiento de algunos aspectos relacionados con la estructuración de los datos

Además de las dificultades conceptuales surgidas durante la introducción de los datos relacionadas con el ámbito de la industria cerámica, comentadas en la sección 5.1, durante el proceso de introducción de los datos hemos tenido que resolver algunas cuestiones derivadas de la estructuración de la base de conocimiento, que se presentan a continuación:

1. *Tratamiento de la herencia*: El editor de ontologías *Protégé* posibilita la herencia múltiple, es decir, contempla la posibilidad de que un misma *class* herede sus propiedades a partir de más de un superordinado. Si bien la herencia múltiple permitiría recoger fenómenos como que una *Baldosa esmaltada* puede ser al mismo tiempo subordinado de *Baldosa cerámica* y de *Baldosa extruida* o *Baldosa prensada*, en nuestro trabajo hemos considerado únicamente la herencia simple, es decir, la herencia de un solo antecesor. De este modo evitamos incompatibilidades conceptuales

entre los valores heredados de dos superordinados distintos. Por ejemplo, *Baldosa extrudida* y *Baldosa prensada* se producen por un método de moldeo distinto (moldeo por extrusión y por prensado, respectivamente). Si *Baldosa prensada* heredase estos dos valores –que en principio no son compatibles– a partir de dos superordinados, la información contenida en la base de datos podría ser inconsistente. Consideramos que, al mantener la herencia simple, la información recogida en la base de conocimiento resulta más coherente.

2. *Tratamiento de la polisemia*: En nuestro trabajo empírico se han identificado algunos términos polisémicos, generalmente en los casos en los que una actividad y el producto resultante de la misma comparten denominación. Por ejemplo, *Pavimento* y *Revestimiento* pueden ser considerados tanto una pieza cerámica (o un conjunto de piezas cerámicas) como el uso que se le da a la pieza o incluso el proceso de revestir o pavimentar. En estos casos, una misma denominación se emplea para referirse a distintos conceptos. Para reflejar estos casos de polisemia (que por otra parte no son muy frecuentes), hemos tenido en cuenta que en *Protégé* se mantiene la denominada *unique name assumption*, es decir, que dos *Classes* distintas no pueden tener el mismo nombre. Para superar esta dificultad hemos añadido un dígito al modo en el que se denominan los conceptos. Así, *Pavimento1* se refiere al pavimento como producto cerámico acabado y *Pavimento2* al uso que se les da a determinadas piezas cerámicas. En las denominaciones correspondientes a estos conceptos, sin embargo, se ha introducido la denominación sin el dígito. De este modo, es posible consultar la base de datos a partir de las denominaciones expresadas mediante el componente *Instance* sin necesidad de tener en cuenta el hecho de que puedan referirse a conceptos distintos.
3. *Tratamiento de las relaciones inversas*: Como hemos señalado en el capítulo 6, para cada relación hemos introducido su relación inversa en caso de que esta se cumpla. A pesar de que *Protégé* permite insertar *slots* inversos, esto no implica que las relaciones

en sentido contrario se establezcan de forma automática, por lo que es necesario introducir las relaciones conceptuales en los dos sentidos de forma manual. Por ejemplo, si hemos definido la relación *material-objeto* entre *Arcilla* y *Baldosa cerámica*, hemos de introducir también la relación *objeto-material* entre *Baldosa cerámica* y *Arcilla*. De momento no se han completado todos los valores de las relaciones inversas en las dos direcciones, sino que nos hemos limitado a reflejarlas en el sentido en el que han aparecido durante el análisis. Más adelante será necesario considerar esta doble direccionalidad, o bien proponer algún tipo de mejora que permita que el editor de ontologías las refleje de forma automática al introducir la relación en uno de los dos sentidos posibles.

4. *Ubicación de algunos conceptos en la jerarquía:* Algunos conceptos, como las partes de los productos acabados (por ejemplo *Cara vista* o *Relieve*) se sitúan en la ontología a la misma altura que *Producto cerámico acabado*, puesto que ambas se consideran subordinadas del concepto superior ENTIDAD INDEPENDIENTE. Podríamos haber introducido un concepto “paraguas” que permitiera agrupar a todas las partes de los productos acabados como *Parte del producto acabado*. Sin embargo, se ha considerado más adecuado no incluir conceptos no lexicalizados en la base de conocimiento.

7.1.1 Protocolo para la identificación de relaciones conceptuales e introducción de datos en *Ontoceram*

Como producto de esta investigación se ha desarrollado un protocolo que detalla los pasos que deben seguirse para identificar las relaciones conceptuales e introducir los datos *Ontoceram*, que servirá de ayuda en el desarrollo posterior de esta o de otras bases de conocimiento con ayuda del editor de ontologías *Protégé*, siguiendo la metodología empleada en el presente trabajo empírico.

Este protocolo podrá utilizarse como guía para desarrollar una ontología cerámica completa en el marco del proyecto ONTODIC. Se presenta tal y

como se podría entregar a los colaboradores en el proyecto, describiendo de forma precisa cada uno de las etapas del análisis.

1. Introducción

Con este protocolo se pretende proporcionar pautas para identificar las relaciones conceptuales que se establecen entre los conceptos implicados en los procesos de la industria cerámica, y para introducirlos en la base de conocimiento *Ontoceram*, desarrollada empleando el editor de ontologías *Protégé*.

A continuación se explica en detalle cómo se realiza la extracción manual de conceptos y relaciones conceptuales de los conceptos de la cerámica a partir del corpus *TXT CERAM* y la base de datos terminológica *Cerámica*, cómo se almacena la información extraída antes de introducirla en la base de datos conceptual *Ontoceram*, y por último, cómo se introducen los conceptos, relaciones y denominaciones. El protocolo se estructura en los siguientes apartados:

- Descripción de los materiales textuales y programas informáticos que se utilizarán para la extracción los conceptos y sus relaciones.

- Procedimiento para la extracción y almacenamiento temporal de los conceptos, sus denominaciones y sus conceptos relacionados en una tabla creada con el programa *Excel*.

- Procedimiento de introducción de los conceptos, relaciones y denominaciones en la base de conocimiento *Ontoceram* utilizando el editor *Protégé*.

2. Materiales textuales y programas informáticos

a) Lista de términos sobre los productos cerámicos

El análisis conceptual de las relaciones se realizará a partir de una lista de términos relacionados con los productos cerámicos proporcionada a los colaboradores por los coordinadores del proyecto.

b) Base de datos terminológicos Cerámica

La base de datos terminológicos *Cerámica*, desarrollada con el programa de gestión terminológica *Multiterm*, se empleará para extraer información sobre los conceptos cerámicos mediante el análisis de los contextos lingüísticos y definatorios contenidos en cada entrada de la base de datos.

c) Materiales textuales: textos del corpus TXT CERAM

La búsqueda de los conceptos relacionados se lleva a cabo mediante el análisis manual de las concordancias de los términos cerámicos en el corpus

textual *TXT CERAM*. A través del análisis de los contextos en los que aparece cada término se identificarán conceptos relacionados y se determinará la naturaleza de la relación que se establece entre ellos. Para la búsqueda de concordancias se empleará el programa *Concord* de la herramienta de análisis textual *WordSmith Tools*.

d) Tabla diseñada con el programa Excel para recoger los conceptos y sus conceptos relacionados

Esta tabla se empleará para almacenar los conceptos y especificar los conceptos relacionados con ellos extraídos durante el análisis, así como el tipo de relación que se establece entre ellos.

En ella se indican, en la primera columna, los conceptos extraídos, y en las columnas subsiguientes los conceptos relacionados con ellos. En la fila superior de la tabla se indica el tipo de relación que los vincula, basada en un catálogo de relaciones conceptuales definido previamente.

e) Tabla resumen del catálogo de relaciones conceptuales

Esta tabla recoge de forma resumida el catálogo de relaciones conceptuales contempladas en la base de datos *Ontoceram*. Para cada relación se indican: el nombre de la relación conceptual, la descripción de la relación en lenguaje natural, el dominio de la relación (clases de conceptos a partir de los cuales ésta puede establecerse), el rango de la relación (clases de conceptos con los que puede establecerse), sus propiedades (transitividad, simetría, recursividad), la relación inversa, si es que existe, y algunos ejemplos de la misma⁶⁴.

f) Base de conocimiento Ontoceram

Esta base de datos recoge los conceptos, sus relaciones conceptuales y sus denominaciones. Se almacena en un archivo denominado *Ontoceram.pprj*, creado utilizando el editor de ontologías *Protégé*.

3. Procedimiento para la extracción de conceptos y relaciones: almacenamiento en una tabla creada con el programa *Excel*

⁶⁴ La tabla resumen de las relaciones conceptuales es la que se ha incluido al final del capítulo 6. Por razones de espacio no la incluimos de nuevo en esta descripción del protocolo. Durante el desarrollo de la base de datos conceptual se entregará a los colaboradores.

a) En primer lugar, tomamos uno de los términos sobre productos cerámicos de la lista proporcionada con este protocolo, por ejemplo *pavimento de gres*.

b) A continuación, buscamos la entrada del término en la base de datos *Cerámica*. En la Fig. 38 se muestra la entrada correspondiente al término *pavimento de gres* en la base de datos *Cerámica*:

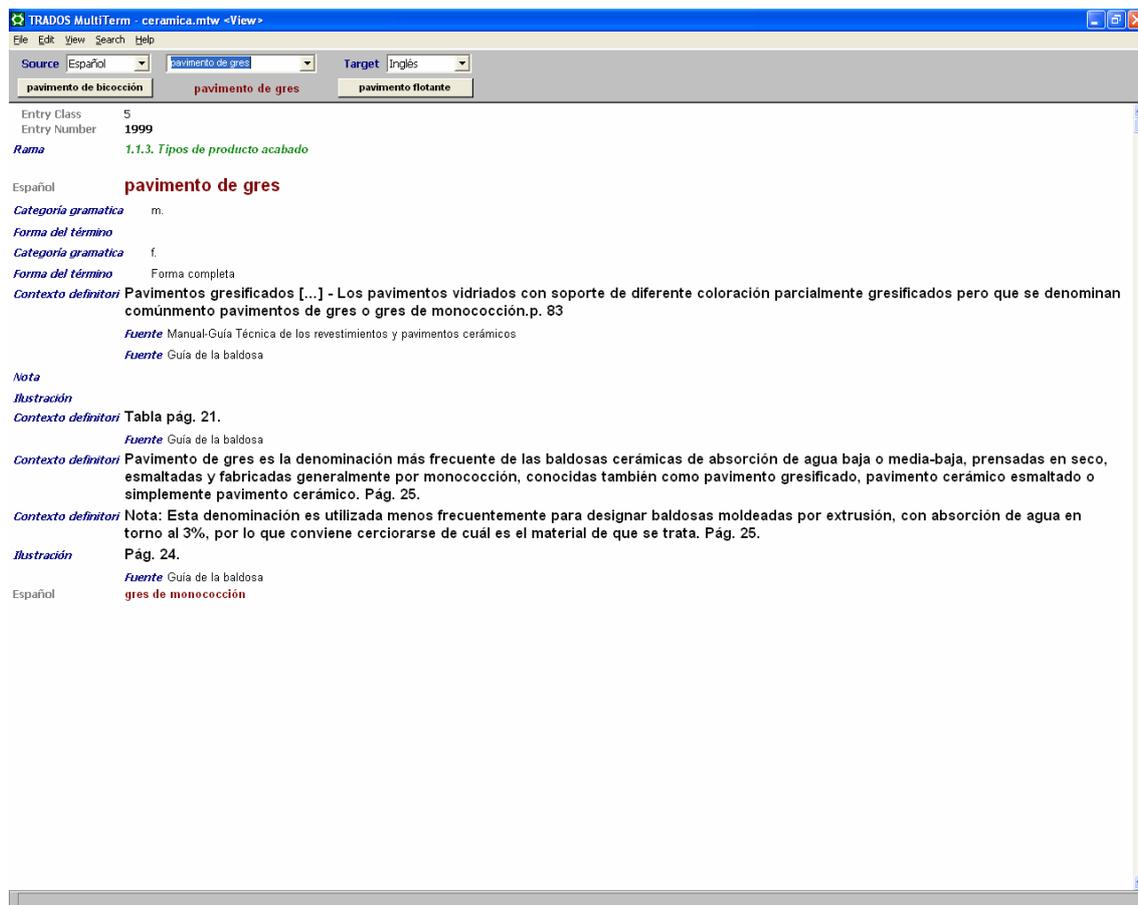


Fig. 38. Ficha terminológica de *pavimento de gres* en la base de datos *Cerámica*.

En esta entrada podemos observar que el término *pavimento de gres* se ubica en la rama *1.1.3 Tipos de producto acabado* del árbol de campo. A partir de los tres contextos definitorios contenidos en la ficha podemos identificar los siguientes conceptos relacionados, junto a los que se indica el tipo de relación que los vincula, deducible por el contexto⁶⁵:

- i. *pavimento vidriado*: relación subordinado-superordinado

⁶⁵ Para establecer la direccionalidad de la relación (es decir, el sentido en el que se establece) se ha considerado que el concepto buscado, esto es, *pavimento de gres* se sitúa a la izquierda de la misma.

- ii. *soporte*: relación objeto-componente
- iii. *baldosa cerámica*: relación subordinado-superordinado
- iv. *absorción de agua*: relación objeto-característica
- v. *prensado en seco*: relación paciente-proceso
- vi. *esmaltado*: relación paciente-proceso
- vii. *monococción*: relación paciente-proceso
- viii. *gres de monococción*: relación de sinonimia
- ix. *pavimento gresificado*: relación de sinonimia
- x. *pavimento cerámico esmaltado*: relación de sinonimia
- xi. *pavimento cerámico*: relación de sinonimia
- xii. *moldeo por extrusión*: relación paciente-proceso

Como puede verse, en los contextos se identifican dos superordinados posibles para *pavimento de gres*: *baldosa cerámica* y *pavimento vidriado*. Dado que *pavimento vidriado* es a su vez un tipo de baldosa cerámica, consideramos que el superordinado más directo es *pavimento vidriado*.

c) Una vez identificados los conceptos relacionados a partir de los contextos definatorios en la entrada terminológica, completamos la tabla de *Excel*, situando el concepto inicial *pavimento de gres* en la primera columna y los conceptos relacionados en las columnas subsiguientes, por debajo del tipo de relación que los vincula, indicado en la primera fila. Dado que las dimensiones de la tabla de relaciones en *Excel* exceden los límites del papel, en este documento se presenta un extracto de la tabla que refleja únicamente las relaciones identificadas. En la Tabla 47 se presentan los conceptos relacionados con *pavimento de gres*:

<i>Concepto/tipo de relación</i>	<i>Subordinado-superordinado</i>	<i>Objeto-característica</i>	<i>Objeto-componente</i>	<i>Paciente-proceso</i>	<i>Sinónimos</i>
Pavimento de gres	Pavimento vidriado	Absorción de agua	Soporte	Prensado en seco Monococción Esmaltado Extrusión	Pavimento gresificado Pavimento cerámico esmaltado Pavimento cerámico Gres de monococción

Tabla 47. Tabla que recoge los conceptos relacionados con *Pavimento de gres*.

d) A continuación extraemos las concordancias de *Pavimento de gres* en el corpus *TXTCERAM* utilizando la herramienta *Concord* de *WordSmith Tools*. Para ello, seguimos los pasos que se detallan a continuación:

1. Abrimos el programa WordSmith Tools y seleccionamos la herramienta Concord en el menú Tools.
2. Seleccionamos los textos mediante la función Start del menú File, o bien presionamos el icono . Se abrirá la ventana que se muestra en la Fig. 39:

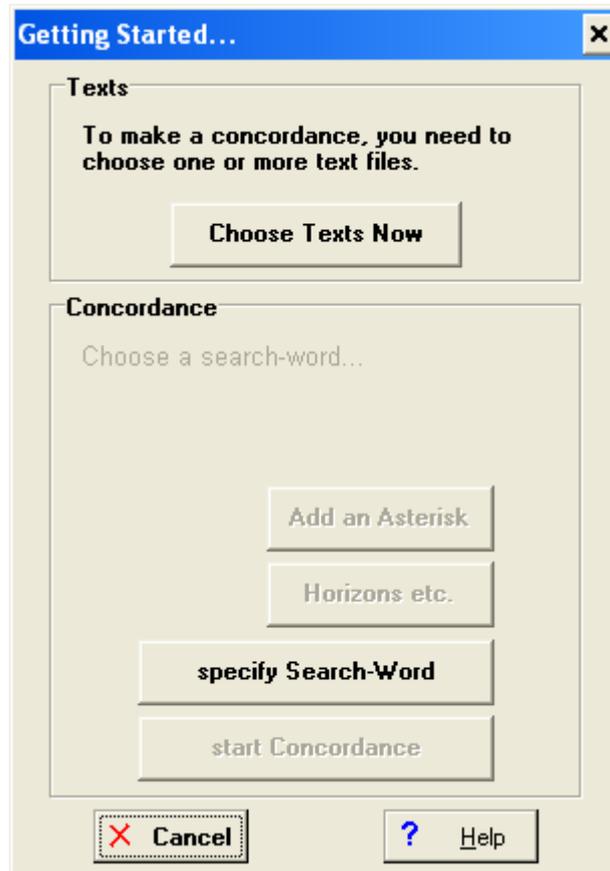


Fig. 39. Menú inicial de la herramienta *Concord*.

Presionamos el botón *Choose Texts Now* y seleccionamos la ruta que indica la ubicación de los textos del corpus *TXT CERAM*, que en nuestro caso se encuentran en la carpeta *corpus* del disco local:

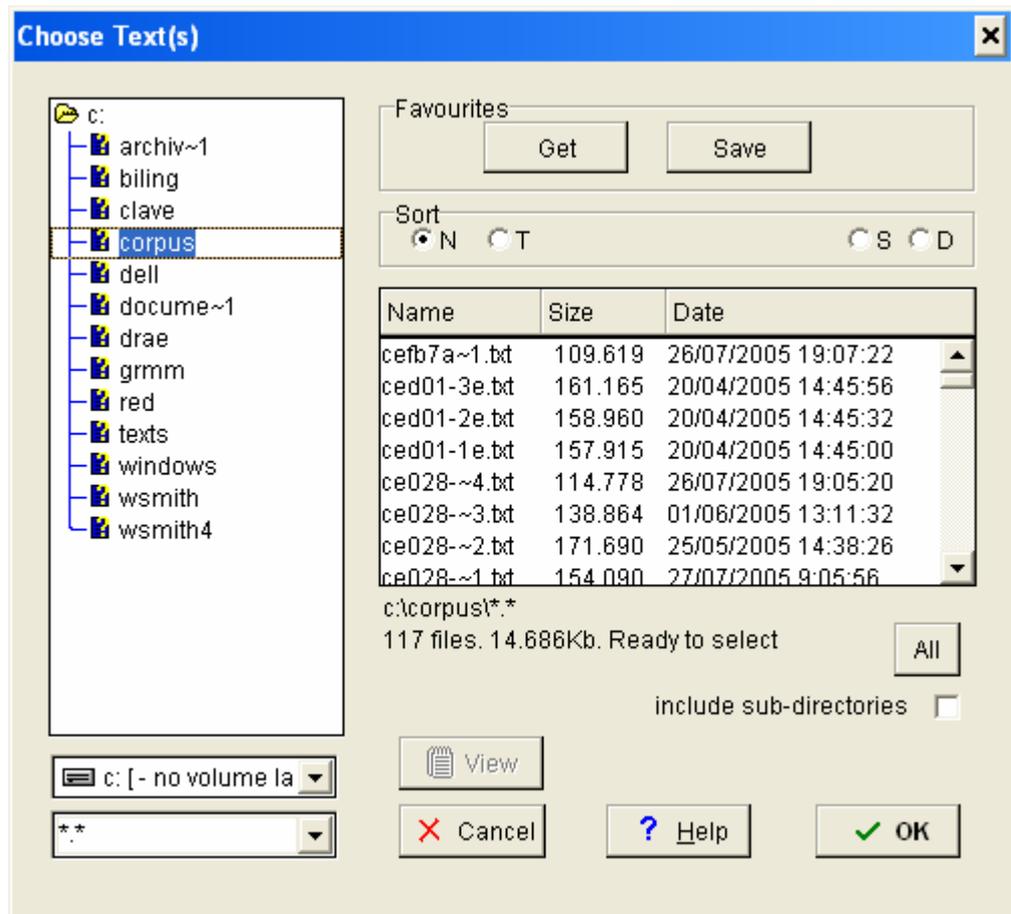


Fig. 40. Menú de selección de los textos del corpus *TXTCERAM*.

Seleccionamos los 117 archivos del corpus pulsando el botón *All* y a continuación *Store*. El programa nos devuelve al menú *Getting Started* mostrado anteriormente. Seleccionamos el botón *Specify search word* y llegamos a la ventana que se muestra en la Fig. 41:

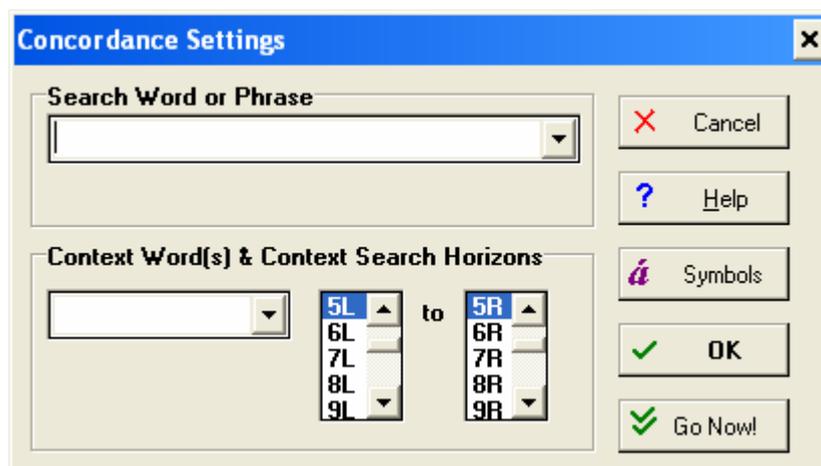


Fig. 41. Ventana de selección de la palabra o frase de búsqueda en *Concord*.

En ella escribimos el término *pavimento* de gres*. El uso del asterisco como comodín posibilita la búsqueda del término en plural. Pulsamos el botón *Go now*.

El programa nos devuelve ala ventana que se muestra en la Fig. 42, en la que podemos observar los 22 contextos del corpus *TXTCERAM* en los que aparece la expresión *pavimento cerámico* o su variante en plural. Utilizando el botón *Grow* (■), situado en la segunda fila de botones, podemos expandir los contextos hasta observar frases completas:

N	Concordance	Set	Tag	Word No.	File	%
1	ntar al producto final, junto al fenómeno de gresificación, una porosidad muy baja. Dada la gran variedad de baldosas gresificadas, por su composición, características físicas, de aspecto y dimensionales, dividiremos el producto en tres clases fundamentales, atendiendo a su capacidad de absorción de agua. • Mosaico, baldosas y pavimentos de gres porcelánico con una capacidad de absorción de agua igual o inferior al 0.5%, medida en pieza individual. Por sus características, este tipo pertenecerá siempre al grupo BI de las normas ENP. • Baldosas gresificadas, con capacidad de absorción de agua superior al 0.5% e inferior o igual al 3%. Pertenecen a los grupos AI y B1. Baldosas de gres rústico. (Foto: Sergio Daho) • Baldosas gresificadas, con capacidad de absorción de agua comprendida			1.632		6
2	stimiento de interiores. Ocre amarillento - rojizo 4 - 15 Pavimentos de bicocción. 14-16 9-11 Blla, Bllb, Blll, Alla, Allb Revestimiento de interiores. Blanco - Blanco grisáceo 15 - 22 Azulejo de pasta blanca 10-20 5-10 Blll Revestimiento de interiores. Diversos colores desde el Blanco grisáceo al Marrón oscuro 1 - 6 Pavimento de gres. 18-23 8-10 Bll, Blla Revestimiento de interiores y exteriores. Diversos colores y tonalidades 0 - 7 Klinker vidriado 20-25 25-30 Al, Alla, Allb, Bll Revestimiento de exteriores. Rojo 10 - 20 Baldosín vidriado 6-10 6-8 Alll Revestimiento de interiores. Ocre, Marrón oscuro 1 - 4 Baldosas de gres 16-20 7-10 Al, Alla Revestimientos y pavimentos exteriores e industriales. Conjunto de piezas de pavimento esmaltado de interior de viviendas de div			7.555		59
3	Dos ejemplos de una misma patología: la pérdida de aspecto por una incorrecta selección del pavimento. En los años sesenta y setenta se utilizaron pavimentos vidriados en locales públicos con bajas resistencias mecánica y química, algunos con vidriados adecuados sólo a revestimientos. El resultado... aquí está a la vista. Pavimento de gres porcelánico pulido, de formato 26x42 cm. Fuente: Catálogo Castellarano Fandre Cerámiche SpA. Las operaciones de pulido o el desgaste continuado a causa de la elevada abrasión a que se someten conllevan no solamente la pérdida de aspecto —prácticamente inobservable— sino la aforación de esa porosidad residual, perdiendo esa propiedad tan preciada como es la resistencia a las manchas. Las reclamaciones vienen por esta cau			14.737		50
4	siguientes características: 1) se trata de estructuras cristalinas estables a altas temperaturas. Según su estabilidad térmica tendríamos que diferenciar entre pigmentos de baja temperatura (utilizables en materiales cerámicos tales como decoración en vajilla doméstica en segunda o tercera cocción, revestimiento cerámico y hasta pavimento de gres) y de alta temperatura (utilizables en cerámica fina, loza, porcelana y también para pigmentación en monococción bajo cubierta). 2) ser materiales insolubles en el vidriado matriz al que se adiciona el colorante. 3) ser materiales con buenas propiedades físicas que no modifique las características de resistencia mecánica ala abrasión, cuarteo y a la acción de los agentes atmosféricos de la matriz. Estructuralmente, un pigmento cerámico está f			9.580		47
5	ntos básicos: 1.Termoestable Es decir, tiene una estructura cristalina estable a elevadas temperaturas. Según su estabilidad, podemos distinguir entre pigmentos de Baja temperatura y de alta temperatura. Los primeros son utilizables en decoración de vajillería domestica en segunda o tercera cocción, revestimiento cerámico o incluso pavimento de gres. Los de elevada temperatura se utilizan para cerámica fina, loza, y porcelanas como en monococción, en decoración bajo cubierta. 2.Insoluble en la matriz vítrea El pigmento se añade a un esmalte cerámico de naturaleza básicamente vítrea y debe mantener su entidad estructural en la cocción produciendo la coloración homogénea en la matriz. 3.No modificar las propiedades de la matriz. El pigmento debe ser un material con buenas propieda			1.658		13
6	trabilidad ante las manchas grasas son básicas en este caso, y hacen del producto cerámico esmaltado el más solicitado revestimiento para este uso. Y ahora también el más solicitado pavimento, debido a la existencia en el mercado de suelos a juego con el alicatado y, también, a la riqueza colorística que últimamente ofrecen los pavimentos de gres. Un pavimento cerámico en cocinas debe ofrecer fuertes prestaciones ante las manchas y la acción de lejías y detergentes activos; si el producto reúne tales condiciones, no hay pavimento que lo supere para este uso. Si la intensidad de utilización es baja, también las exigencias ante el producto podrán ser menores, pero debe ponderarse el factor de riesgo: no es lo mismo una baja intensidad de utilización que una intensa utilización pocas			9.822		37
7	es 18-20 7-10 Al, Alla Revestimientos y pavimentos exteriores e industriales. Conjunto de piezas de pavimento esmaltado de interior de viviendas de diversos formatos. Pavimento esmaltado de interior de viviendas y locales comerciales con rodapié y taco. Pavimento esmaltado de interior de viviendas y locales comerciales. Pavimento de gres porcelánico de altas prestaciones para locales comerciales y viviendas. Pavimento de interior esmaltado y no esmaltado de elevadas prestaciones para locales comerciales y viviendas. Pavimento de interior no esmaltado de alta absorción de agua. Pavimento rústico no esmaltado. Pavimento de exterior esmaltados y no esmaltado de diversos formatos. Pavimento de exterior antideslizante y piezas especiales. Pavimento de exteriores y			7.656		60
8	trata de una máquina de concepción mecánica simple, y que por lo tanto no precisa de un mantenimiento excesivamente especializado. Desventajas: Esta prensa, por sus características, no es capaz de garantizar un prensado rigurosamente uniforme. Esto la descalifica para su adaptación a procesos del tipo de fabricación de pavimentos de gres, en el que, como ya sabemos, es muy importante el mantenimiento de las variables de prensado (presión y humedad) estrictamente constantes, o al menos variables dentro de límites aceptables y controlados. La prensa de fricción (Figura 4-21) dispone de un volante mandado a través de una transmisión a fricción; la banda de cuero que rodea el volante entra en contacto con los discos con los que fricciona y de los que recibe el movimiento			2.160		21
9				688		2

Fig. 42. Concordancias de *Pavimento* cerámico* en el corpus *TXTCERAM*.

El análisis pormenorizado de estos contextos nos proporciona nuevos conceptos relacionados con *pavimento de gres*. A continuación se especifican dichos conceptos relacionados junto al tipo de relación que los vincula:

- i. *pigmento de baja temperatura*: relación objeto-material
- ii. *suelo*: relación concepto-lugar
- iii. *gresificado*: relación paciente-proceso

Estos nuevos conceptos relacionados se añaden a la tabla obtenida en el paso anterior, de modo que se completa tal y como se muestra en la Tabla 48:

Concepto/tipo de relación	Subordinado-superordinado	Objeto-característica	Objeto-componente	Objeto-material	Paciente-proceso	Concepto-lugar	Sinónimos
Pavimento de gres	Pavimento vidriado	Absorción de agua	Soporte	Pigmento de baja temperatura	Prensado en seco Monococción Esmaltado Extrusión Gresificado	Suelo	Pavimento gresificado Pavimento cerámico esmaltado Pavimento cerámico Gres de monococción

Tabla 48. Tabla completa de los conceptos relacionados con *pavimento de gres*.

4. Introducción de los datos en el editor de ontologías *Protégé*

Una vez completado el análisis de los conceptos relacionados con *pavimento de gres*, podemos pasar a la fase de introducción del concepto y sus conceptos relacionados en la base de conocimiento *Ontoceram*.

a) En primer lugar, abrimos el programa *Protégé* desde *Inicio>Programas*. Seleccionamos la opción *Open Existing Project* y aparece la pantalla que se recoge en la Fig. 43:

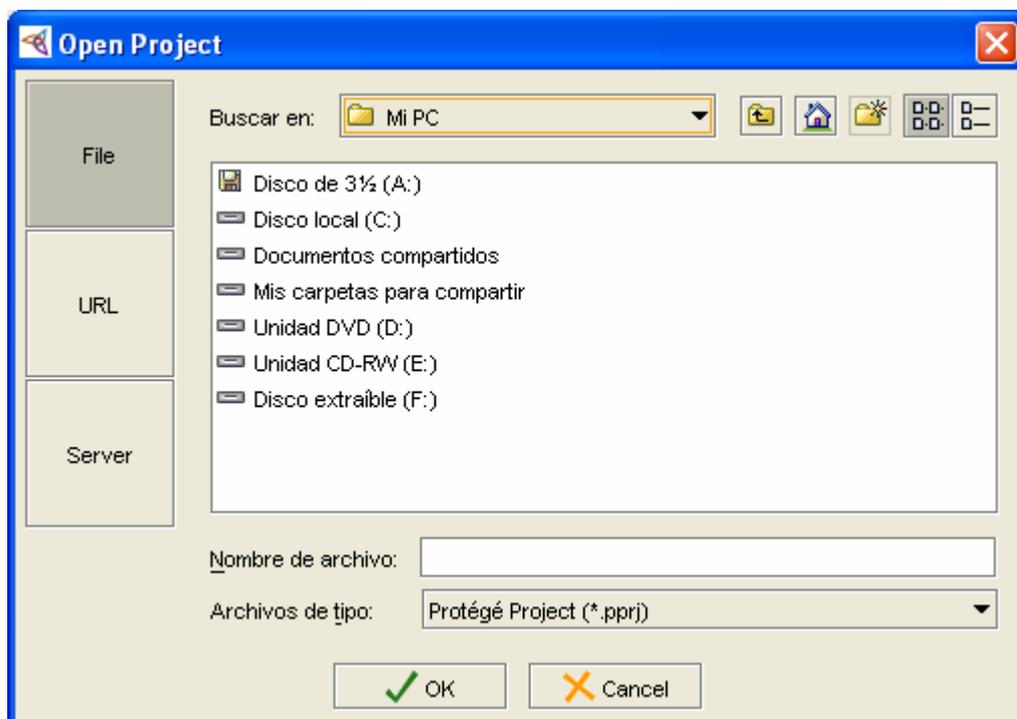


Fig. 43. Ventana *Open Project* del editor de ontologías *Protégé*.

Indicamos la ruta para llegar a la ubicación de la base de datos conceptual *Ontoceram.pprj* y abrimos el archivo, con lo que llegamos a la ventana que se muestra en la Fig. 44:

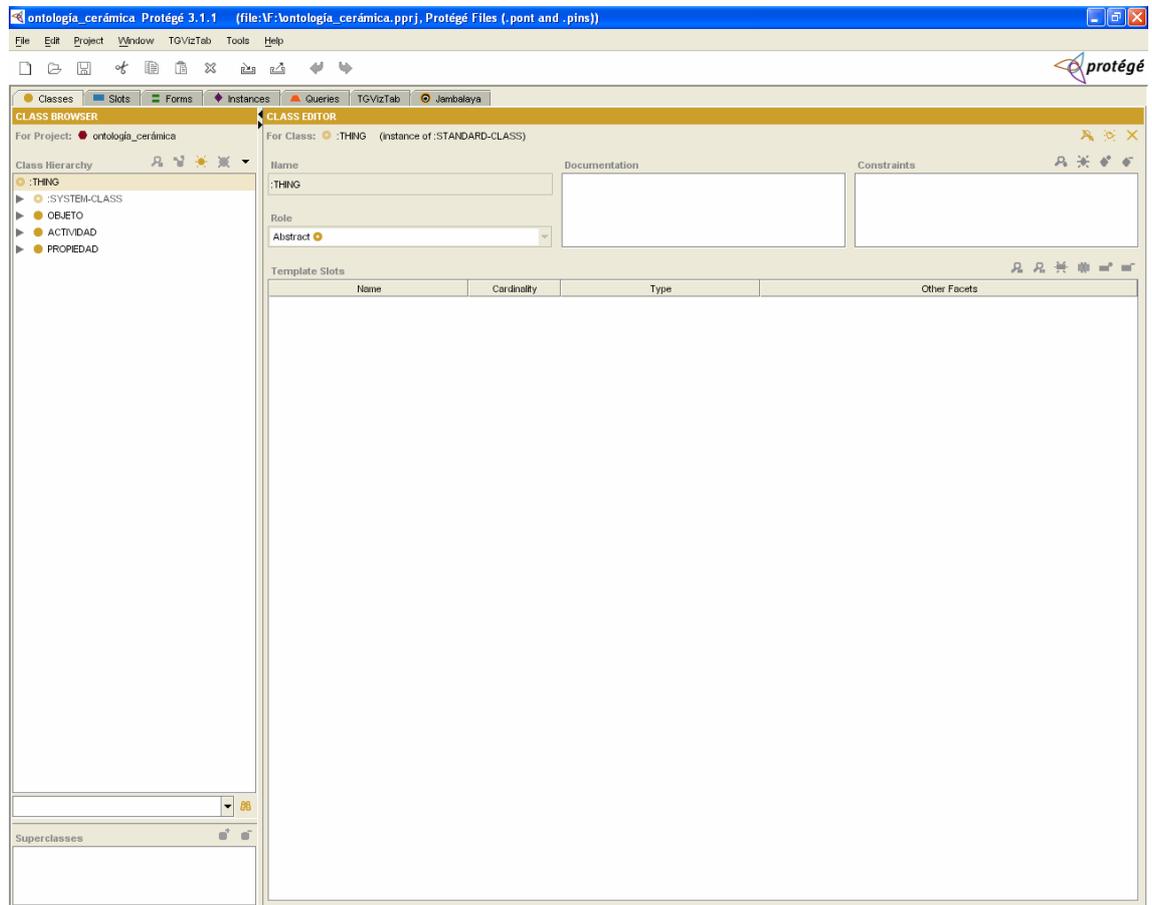


Fig. 44. Apariencia de la base de datos *Ontoceram* al abrirla en el editor de ontologías.

En la parte inferior izquierda de la pantalla, utilizamos la casilla de búsqueda para comprobar si el concepto que queremos introducir (*Pavimento de gres*) ya se encuentra contenido en la base de datos o no. Si es así, abrimos el marco correspondiente a la clase *Pavimento de gres*, en el que aparecen a la derecha los *slots* que se atribuyen al concepto, bien porque los haya heredado de sus superordinados, o bien porque se hayan modificado en una fase anterior del trabajo. Si la clase *Pavimento de gres* aún no está contenida en *Ontoceram*, seleccionamos el superordinado apropiado (en nuestro caso *Pavimento vidriado*), y creamos la nueva clase mediante el botón *Create Class*. En la Fig. 45 ofrecemos la captura de pantalla que muestra la configuración de la *Class* que representa al concepto *Pavimento cerámico*:

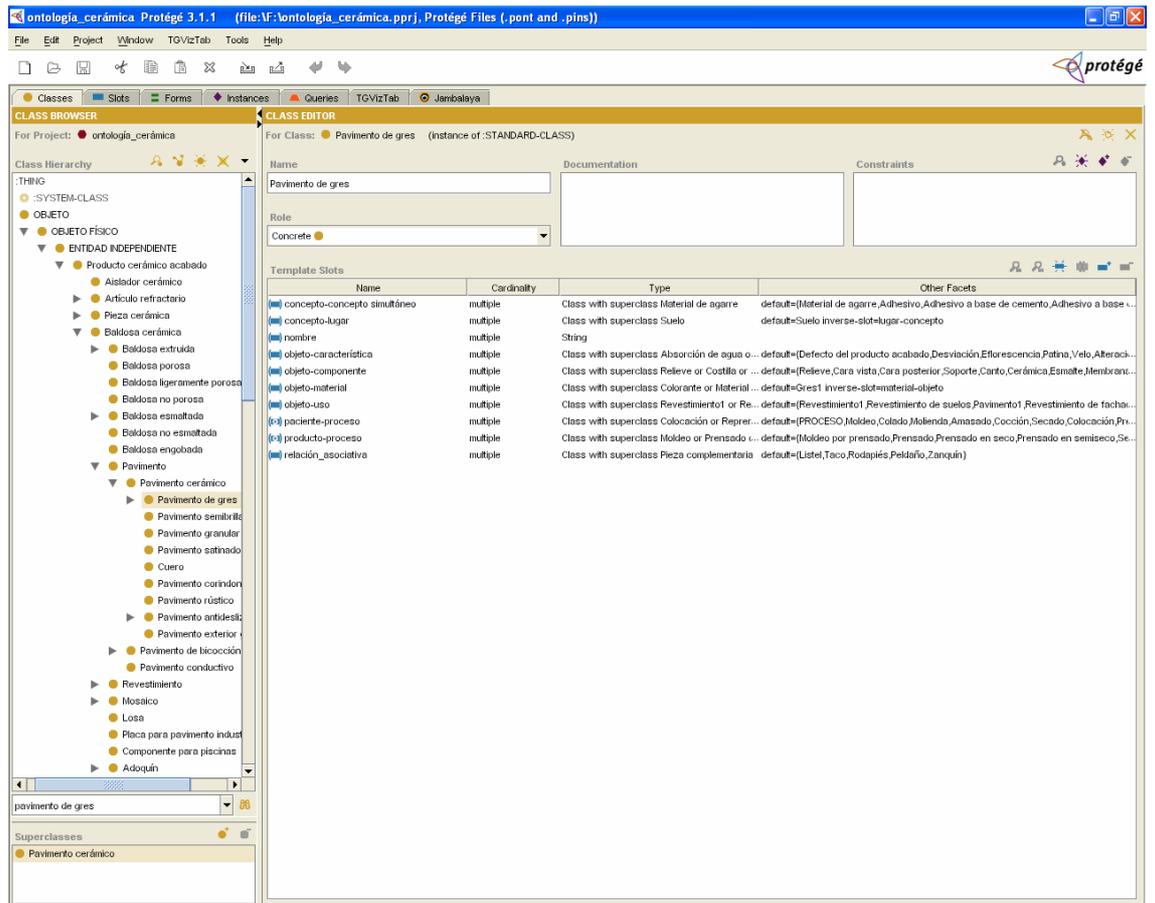


Fig. 45. Marco de la *Class* que representa al concepto *Pavimento de gres*.

Basándonos en la información contenida en la tabla de Excel que hemos completado en la fase anterior, añadimos los conceptos relacionados como valores de los *slots* correspondientes. Para ejemplificar este proceso, a continuación se explica cómo se introducen los conceptos relacionados con *Pavimento de gres* mediante la relación *paciente-proceso*, que son *Prensado en seco*, *Esmaltado*, *Monococción*, *Extrusión* y *Gresificado*. Hacemos doble clic sobre el *slot* correspondiente a la relación *paciente-proceso* y seleccionamos la opción *View slot at class*, puesto que queremos modificar los valores de ese *slot* para la clase *Pavimento de gres*. Obtenemos la pantalla que se muestra en la Fig. 46:

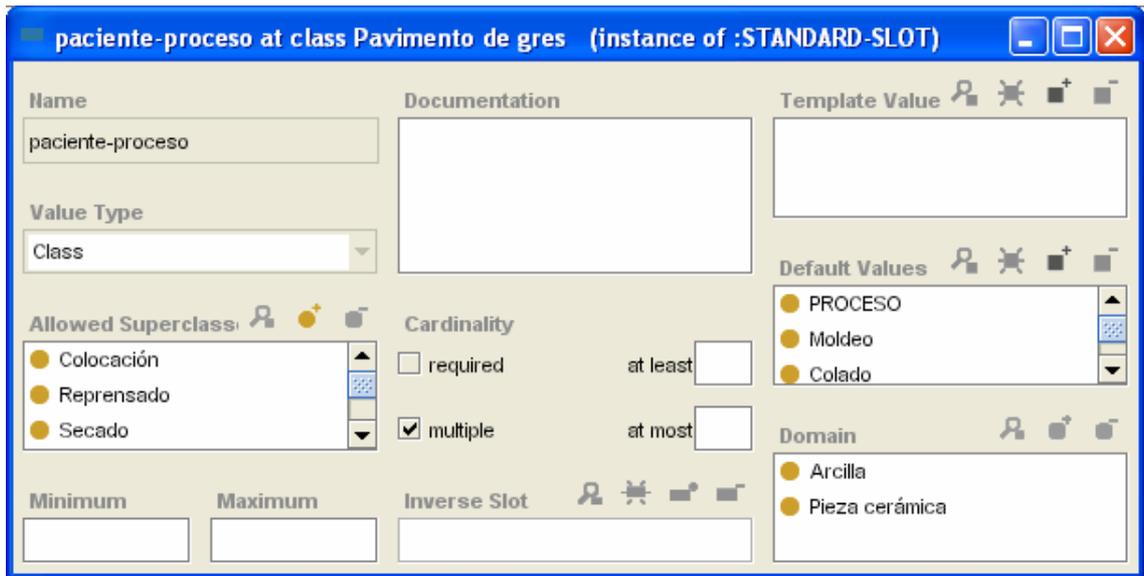


Fig. 46. Ventana del *slot* de relación *paciente-proceso* en la clase *Pavimento de gres*.

Este *slot* ya posee algunos valores que, o bien se han heredado a partir de su superordinado (*Pavimento vidriado*), o bien se han introducido en una fase anterior del trabajo. Comprobamos si los conceptos relacionados que hemos identificado aparecen como valores posibles, y en caso contrario los añadimos en la faceta *Allowed Superclass*.

En caso de que los conceptos relacionados identificados aún no estén incluidos, haremos clic en el botón que nos permite agregar clases y seleccionaremos el proceso correspondiente entre los que ya han sido incluidos en la base de conocimiento. Si alguno de los procesos aún no está incluido en la base de datos conceptual, deberemos crear una nueva clase que represente a ese concepto. Por ejemplo, el proceso *Gresificado* aún no está incluido como clase en *Ontoceram*. Para incluirlo, seleccionamos en el navegador de clases el superordinado correspondiente, que será PROCESO. Elegimos la opción *Create Class* para crear una clase subordinada, le asignamos el nombre *Gresificado* en la casilla *Name* situada en la parte derecha (teniendo en cuenta que los conceptos representados por clases se escriben con mayúscula inicial) y guardamos la nueva clase, con lo que aparecerá entre las clases que se pueden relacionar con *Pavimento de gres* mediante la relación *paciente-proceso*.

b) Una vez establecidas todas las relaciones del concepto *Pavimento de gres* mediante los *slots*, introducimos las denominaciones correspondientes a este concepto mediante el componente *Instance* de *Protégé*.

Para introducir las denominaciones tenemos que seleccionar la pestaña *Instances* en el programa *Protégé*, y buscar en el navegador de clases la clase *Pavimento de gres*. La Fig. 47 muestra las dos denominaciones ya contenidas en la base de datos conceptual, *pavimento de gres* y *pavimento gresificado*:

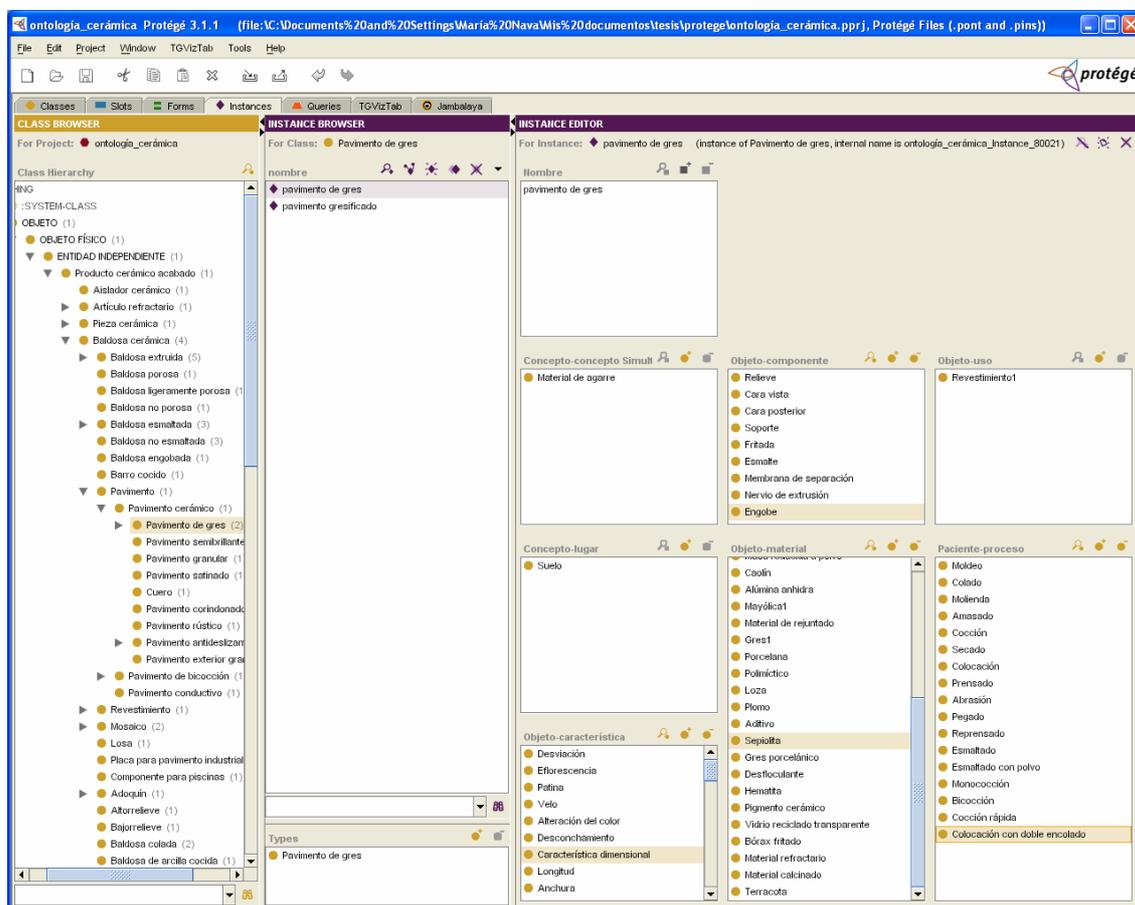


Fig. 47. Instancias que representan las denominaciones del concepto *Pavimento de gres*.

En nuestro análisis hemos identificado las siguientes posibilidades denominativas para referirse al concepto *Pavimento de gres*: *pavimento de gres*, *gres de monococción*, *pavimento gresificado*, *pavimento cerámico esmaltado* y *pavimento cerámico*. Todas ellas deben recogerse como denominaciones del concepto *Pavimento de gres*. Para añadir instancias procedemos del siguiente modo:

- i. Con la clase *Pavimento de gres* seleccionada en el navegador de clases (*Class Browser*), seleccionamos el botón *Create Instance* en el navegador de instancias que aparece en la parte central de la pantalla.

ii. En el editor de instancias situado en la parte derecha de la pantalla, editamos la denominación, que aparece en la parte superior, y escribimos la primera de las denominaciones aún no incluidas (*gres de monococción*).

Repetimos el mismo proceso para añadir cada uno de los sinónimos (*pavimento cerámico esmaltado* y *pavimento cerámico*). En la Fig. 48 recogemos la captura de pantalla que muestra las cinco denominaciones posibles para el concepto *Pavimento de gres* en nuestra base de datos conceptual:

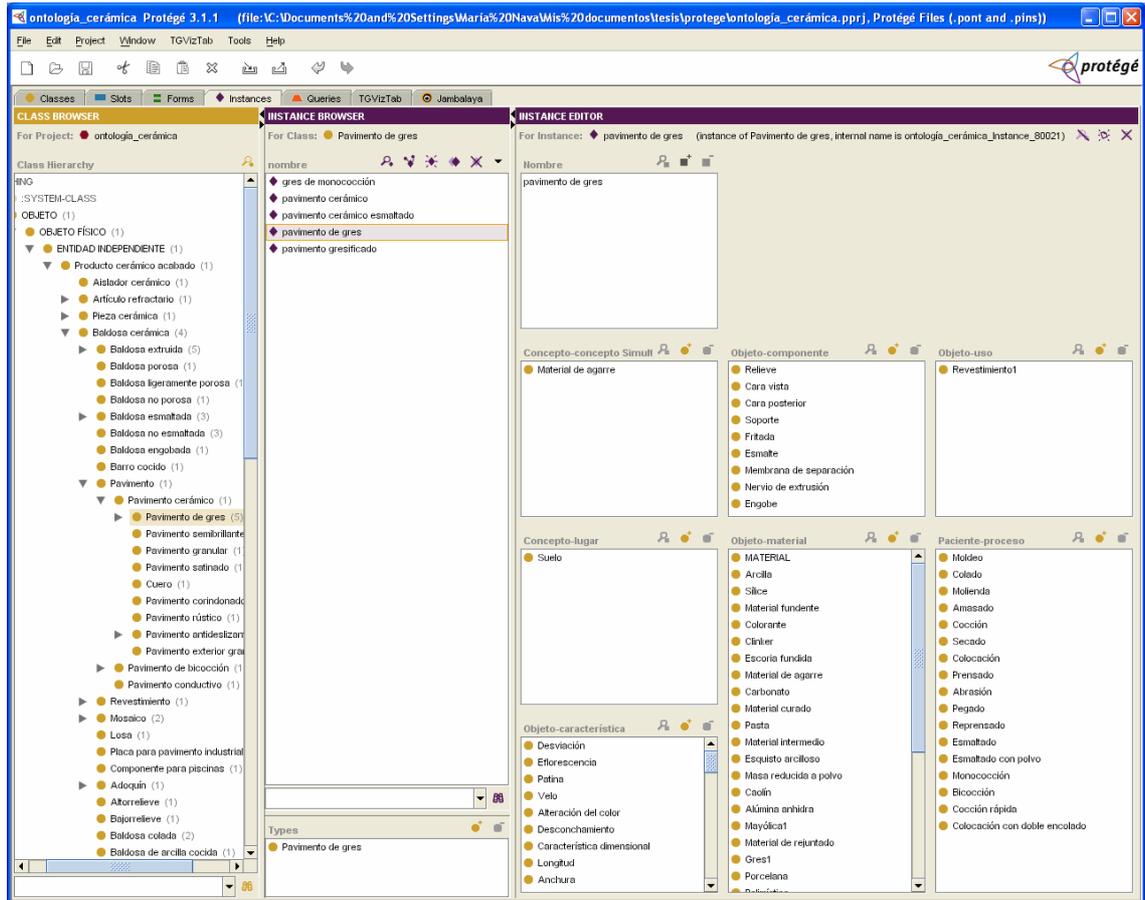


Fig. 48. Lista completa de las denominaciones empleadas para referirse al concepto *Pavimento de gres*.

Mediante la repetición del proceso descrito en este protocolo para todos los conceptos de la industria cerámica y documentados en los textos, se irá completando la base de datos conceptual hasta configurar una verdadera ontología de la cerámica.

7.2 Recuperación de la información sobre las relaciones conceptuales en Ontoceram

En la última fase de nuestro trabajo empírico hemos analizado las posibilidades de aprovechamiento de la base de conocimiento para la recuperación de información sobre los conceptos de manera onomasiológica, es decir, a partir de la información sobre las relaciones conceptuales formalizada en ella.

En primer lugar, nos hemos preguntado qué tipo de información conceptual sobre relaciones nos podría interesar recuperar a partir de la base de conocimiento. Podríamos querer recuperar los conceptos más generales a los cuales pertenece la noción buscada, los conceptos que son una particularización del buscado, o las propiedades típicas de esta noción. También podríamos necesitar información sobre las relaciones que existen entre dos conceptos, la lista de conceptos subordinados de uno determinado, o aquellos que poseen una propiedad particular. Por último, la consulta podría orientarse a encontrar denominaciones para un concepto determinado, en cuyo caso (búsqueda propiamente onomasiológica) podríamos querer buscar una denominación a partir de los elementos de la definición (pertenencia a una clase, propiedades, relaciones) o la adecuación de una descripción a una denominación determinada.

En este apartado pretendemos mostrar el modo en que puede recuperarse la información sobre relaciones conceptuales estructurada en la base de datos conceptual *Ontoceram*. Para ello, en el apartado 7.2.1 proponemos diez consultas que hacen referencia a las relaciones conceptuales entre los productos cerámicos. En cada consulta, explicamos el modo en el que la estructuración de la información en *Ontoceram* permite dar una respuesta adecuada a la consulta, y cómo se ha implementado dicha información en nuestra base de datos conceptual.

En el apartado 7.2.2 se presentan algunos ejemplos que muestran el modo en el que se puede dar respuesta a algunas de estas consultas con las funciones que el editor de ontologías *Protégé* pone a nuestro alcance. Veremos cómo utilizar la pestaña de consultas (*Queries Tab*) para la recuperación de denominaciones explicitadas por medio del componente *Instance*, y presentaremos algún ejemplo de consultas más complejas empleando el lenguaje axiomático PAL, a través de la pestaña *PAL Queries Tab*.

7.2.1 Consultas sobre relaciones conceptuales en *Ontoceram*

En este apartado presentamos diez consultas sobre las relaciones conceptuales que se establecen entre los productos cerámicos con el fin de determinar si el modelo de estructuración de la información sobre las relaciones en la base de conocimiento *Ontoceram* que hemos planteado en este trabajo de investigación es válido para la recuperación de la información.

Hemos procurado que las consultas planteadas sean representativas de varios tipos de relación, y que estén expresadas en lenguaje natural, tal y como lo haría una persona que no conozca el catálogo de relaciones.

A continuación presentamos cada una de las diez consultas y explicamos cómo se ha incorporado en la base de conocimiento la información conceptual sobre las relaciones que nos permitirá dar respuesta a cada una de ellas. De este modo esperamos comprobar si los editores de ontologías desarrollados desde la ingeniería del conocimiento pueden ser utilizados como herramienta para el almacenamiento y recuperación de la información sobre relaciones conceptuales.

Consulta 1



¿Qué partes componen una baldosa cerámica?

Mediante esta consulta pretendemos saber cuáles son los componentes del objeto *baldosa cerámica*. En nuestra base de conocimiento se ha incorporado la relación *componente funcional-objeto*, que se establece entre los elementos que desempeñan una función dentro del todo en el que se integran. Esta relación se ha implementado en la base de datos mediante dos *slots* que reflejan la doble direccionalidad de la relación (*componente funcional-objeto* y *objeto-componente funcional*). Durante el trabajo empírico hemos identificado las partes que componen una baldosa cerámica y las hemos representado en el editor de ontologías por medio del componente *Class*. Entre *baldosa cerámica* y sus componentes se establece un vínculo por medio del *slot objeto-componente funcional*. En el concepto *Baldosa cerámica*, el *slot objeto-componente* posee los valores *Soporte*, *Cara vista*, *Arista*, *Cara posterior*, *Membrana de separación*, *Nervio de extrusión*, *Engobe*, *Costilla* y *Relieve*, que también se han incluido

como conceptos en la base de conocimiento. Esta sería la respuesta a nuestra primera consulta.

Si recuperamos el formalismo de representación de las relaciones conceptuales $a R b$, *baldosa cerámica* sería el dominio de la relación y las partes que la componen constituyen el rasgo de la misma. La Tabla 49 muestra la información conceptual que se ha incorporado en la base de conocimiento y que nos permite dar respuesta a esta consulta:

Nombre de la relación	Objeto-componente funcional
Definición en lenguaje natural	Relación que se establece entre un todo y sus partes constituyentes funcionales.
Clases conceptuales	
- Dominio	Baldosa cerámica
- Rango	Soporte, Cara vista, Arista, Cara posterior, Membrana de separación, Nervio de extrusión, Engobe, Costilla, Relieve
Propiedades matemáticas	
- Transitividad	Sí
- Simetría	No
- Recursividad	Sí
- Relación inversa	Componente funcional-objeto

Tabla 49. Información sobre las relaciones implementadas en la base de conocimiento.

Consulta 2



¿Con qué materiales se fabrica el baldosín catalán?

Con esta consulta pretendemos obtener los materiales con los que se fabrica un determinado producto acabado, en nuestro caso el *Baldosín catalán*. En nuestra base de datos conceptual hemos introducido este concepto en la jerarquía de *Classes*, y se le ha asignado el *slot* de relación *objeto-material*. Mediante este *slot* hemos establecido la relación entre el objeto *Baldosín catalán* y los materiales con los que se fabrica, que también son conceptos incluidos en *Ontoceram*. Si buscamos *Baldosín catalán* en la base de datos y observamos los valores de la relación representada por el *slot objeto-material*, veremos que un baldosín catalán se fabrica con *Arcilla de cocción roja* a la que se pueden añadir *Colorantes*.

Consulta 3



¿Qué piezas cerámicas son resistentes a la helada?

Con esta consulta queremos conocer qué tipos de baldosas poseen una determinada característica, en este caso, cuáles de ellas poseen la propiedad *Resistencia a la helada*. En *Ontoceram* hemos introducido una relación *objeto-característica*, mediante la que hemos vinculado los productos cerámicos acabados y sus propiedades. En este caso, nos interesa ver en el concepto *Resistencia a la helada* cuál es el rango la relación *característica-objeto*, es decir, qué objetos pueden aparecer como valores de dicho *slot* de relación. De entre estos objetos habría que seleccionar aquellas que sean baldosas cerámicas. Como vemos, la información se encuentra estructurada en la base de datos, pero el modo de recuperarla no es directo. Más adelante veremos que la generación de una consulta en la pestaña *Queries Tab* de *Protégé* nos permitirá dar respuesta a esta pregunta de forma directa.

Consulta 4



¿Qué baldosas cerámicas se producen por prensado en seco?

En este caso la consulta está encaminada a obtener como resultado una lista de productos cerámicos que se fabrican mediante un proceso concreto, el *Prensado en seco*. Dado que en *Ontoceram* se ha introducido un *slot* de relación *producto-proceso*, inversa de la relación *proceso-producto*, podremos identificar los valores de esta relación para el concepto *Baldosa cerámica*, que corresponderán a aquellas baldosas que se fabrican mediante este proceso. Para dar respuesta a esta consulta podemos utilizar el *plug-in* de consultas de *Protégé Queries Tab*, como veremos en el siguiente apartado. Los tipos de baldosa cerámica producidos mediante este proceso son, según el estado de desarrollo actual de la base de datos conceptual, el *Pavimento cerámico esmaltado*, el *Pavimento cerámico* y la *Baldosa de gres porcelánico*.

Consulta 5



¿Qué tipo de producto cerámico es una baldosa de klinker?

En este caso queremos conocer cuál es el superordinado del concepto *Baldosa de klinker*. La estructura de nuestra base de datos conceptual nos permite observar fácilmente cuál es el concepto situado inmediatamente por encima de este en la jerarquía de conceptos, que en este caso es el concepto *Baldosa de gres industrial de pasta compuesta*. Por tanto, la relación subordinado-superordinado también queda establecida de forma explícita en nuestra base de conocimiento *Ontoceram*.

Consulta 6



¿Cuáles son las fases del proceso de moldeo por extrusión?

La información que queremos recuperar mediante esta consulta son las distintas fases que conforman el proceso de extrusión. En nuestra base de conocimiento hemos introducido un tipo de relación meronímica, la relación etapa-proceso, que tiene una relación inversa, la relación proceso-etapa. Sin embargo, dado que de momento sólo hemos analizado los conceptos sobre productos cerámicos acabados, aún no disponemos de la información conceptual que permitiría dar respuesta a esta consulta. Sería necesario seguir desarrollando la base de conocimiento para incorporar la información sobre los procesos de fabricación. No obstante, cuando se complete la información sobre procesos, las etapas del *Moldeo por extrusión* estarán vinculadas con este proceso mediante el *slot* de relación *proceso-etapa*.

Consulta 7



¿Qué tipo de baldosas se pueden utilizar para el revestimiento de suelos?

Mediante esta consulta queremos recuperar aquéllas baldosas cerámicas que se utilizan con un fin concreto, en este caso el revestimiento de suelos. En nuestro catálogo hemos incluido una relación de tipo circunstancial que vincula los objetos con el uso para el que están destinados, la relación *objeto-uso*. Por tanto, es posible recuperar esta información mediante la definición de una consulta sobre aquellas baldosas cerámicas cuyo uso es el revestimiento de suelos. El resultado de esta consulta sería, por el momento, el concepto *Pavimento de gres*.

Consulta 8



¿Qué efectos produce el calafateado en los productos cerámicos?

Esta consulta hace referencia a la relación causa-efecto que se establece entre un proceso, el *Calafateado* y los efectos que dicho proceso tiene sobre un objeto, en este caso sobre los productos cerámicos. En nuestro catálogo de relaciones hemos incluido una relación argumental *causa-efecto*, implementada en la base *Ontoceram* mediante el *slot* del mismo nombre. Dicho *slot*, asignado a la clase *Calafateado*, tendrá unos valores determinados, que en este caso es el *Sellado hermético*, que es una de las características que pueden presentar los productos cerámicos.

Consulta 9



¿En qué procesos de colocación se emplea el mortero de cemento portland?

Con esta consulta queremos conocer en qué procesos se utiliza un determinado instrumento, en este caso el *Mortero de cemento portland*. Dado que en nuestro catálogo hemos definido una relación *proceso-instrumento* y su relación inversa, la relación *instrumento-proceso*, sería posible dar respuesta a esta consulta, ya que el concepto *Mortero de cemento portland* tiene asignados unos valores para el *slot* de relación *instrumento-proceso*, entre los cuales se

encontrarán los procesos de colocación en los que se emplea este tipo de material de agarre. De momento no hemos desarrollado en la base de conocimiento la información sobre los procesos de colocación y los materiales de agarre que se emplean en ellos. A medida que se vaya desarrollando la base *Ontoceram* iremos incorporando este tipo de datos, que permitirán dar respuesta a esta consulta.

Consulta 10



¿Para qué sirve la capa de desolidarización?

Con esta consulta queremos saber cuál es el uso al que se destina una entidad concreta, en este caso la *Capa de desolidarización*. En nuestro catálogo hemos recogido la relación circunstancial que se establece entre un objeto y el uso para el que está destinada. Tanto la relación *objeto-uso* como su relación inversa *uso-objeto* se han implementado en la base de datos conceptual *Ontoceram* mediante sendos *slots* de relación. Por ello, una vez llevado a cabo el análisis del concepto *Capa de desolidarización* podremos asignarle el *slot objeto-uso* para vincular este concepto con el uso para el que se destina, que es el *Aislamiento del soporte*, que a su vez es una de las fases del proceso de colocación de los productos cerámicos.

El modelo de formalización e implementación del catálogo de relaciones conceptuales propuesto en este trabajo permite dar respuesta a estas consultas, que en las bases de datos terminológicas tradicionales resultarían difíciles de resolver, ya que la información se encuentra dispersa en las definiciones terminológicas de los conceptos y no se encuentran estructuradas de forma recuperable. Vemos asimismo que nuestro catálogo no sólo tiene en cuenta las relaciones jerárquicas tradicionales, sino que permite la consideración de otros tipos de relación, como las argumentales y circunstanciales, que también tienen importancia en la estructuración del conocimiento especializado.

7.2.2 Realización de consultas en el editor de ontologías *Protégé*

Al describir el editor de ontologías en el apartado 5.3.3, hemos indicado que *Protégé* posee dos funciones que permiten realizar consultas en las ontologías creadas con este editor. Estas funciones se realizan a través de las pestañas *Queries Tab* y *PAL Queries Tab*. La primera posibilita la ejecución de consultas sobre la información contenida en el componente *Instance*, que en nuestro caso representa las distintas denominaciones para un concepto, indicando una serie de criterios que deben cumplirse, mientras que la segunda permite realizar consultas más complejas sobre todos los marcos contenidos en las ontologías empleando el lenguaje axiomático PAL (*Protégé Axiom Language*).

En este apartado se muestran algunos ejemplos de cómo estas funciones de *Protégé* pueden servirnos para dar respuesta a algunas de las consultas planteadas en el apartado anterior.

a) Consultas mediante la pestaña *Queries Tab*

Meditante la pestaña *Queries tab* podemos crear, guardar y ejecutar consultas sobre las instancias contenidas en la base de conocimiento. Así, crearemos una biblioteca de consultas que quedan guardadas en el editor para su posterior uso, y es posible definir otras nuevas en función de nuestras necesidades. Las consultas generadas nos permiten recuperar las denominaciones (almacenadas en el componente *Instance*) que cumplen los criterios especificados.

A modo de ejemplo, veamos cómo podemos emplear este tipo de consultas para dar respuesta a dos de las preguntas que definíamos al principio de esta sección.

Tomemos la Consulta 3, con la que pretendíamos recuperar aquellas piezas cerámicas que son resistentes a la helada. Para dar respuesta a esta pregunta hemos creado una consulta a la que hemos denominado *piezas cerámicas resistentes a la helada*. En ella seleccionamos el concepto *Pieza cerámica*, representado en nuestra base de conocimiento mediante el componente *Class*. Seleccionamos como *slot* el que representa a la relación *objeto-característica* y como criterio seleccionamos que tenga como valor de este *slot* la característica *Resistencia a la helada*, que también está representada en nuestra base de conocimiento por una clase. Al ejecutar esta consulta obtenemos 167 resultados que corresponden con las denominaciones de todas las piezas que poseen la

característica técnica buscada. En la Fig. 49 se muestra una captura de pantalla de los resultados de la consulta, entre los que se encuentran piezas cerámicas como *baldosa de doble cocción* o *baldosa esmaltada para revestimiento*. Entre paréntesis podemos observar el concepto representado mediante el componente *Class* al que hace referencia cada una de las denominaciones encontradas.

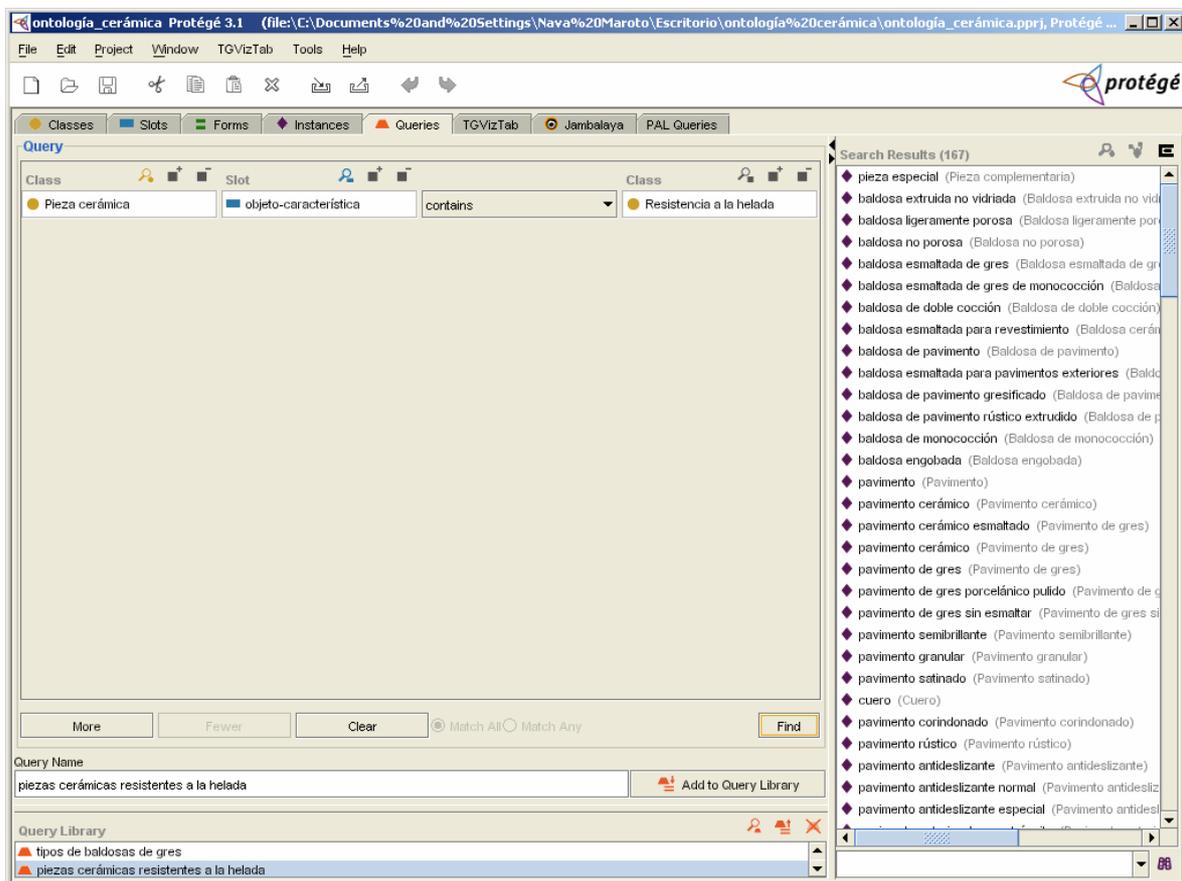


Fig. 49. Resultado de la consulta *piezas cerámicas resistentes a la helada*.

Para dar respuesta a la consulta número 4: *¿Qué baldosas cerámicas se producen por prensado en seco?*, hemos generado una consulta a la que hemos denominado *baldosas cerámicas prensadas en seco*. En ella limitamos la búsqueda al concepto representado por la *Class Baldosa cerámica*, y utilizamos como criterio que el valor del *slot producto-proceso* sea el concepto *Moldeo por prensado en seco*. La búsqueda nos proporciona tres instancias (*pavimento cerámico esmaltado*, *pavimento cerámico* y *baldosa de gres porcelánico*), correspondientes a los conceptos *Pavimento de gres* y *Baldosa de gres porcelánico*. En la Fig. 50 se muestra la captura de pantalla en la que aparece el resultado de esta búsqueda:

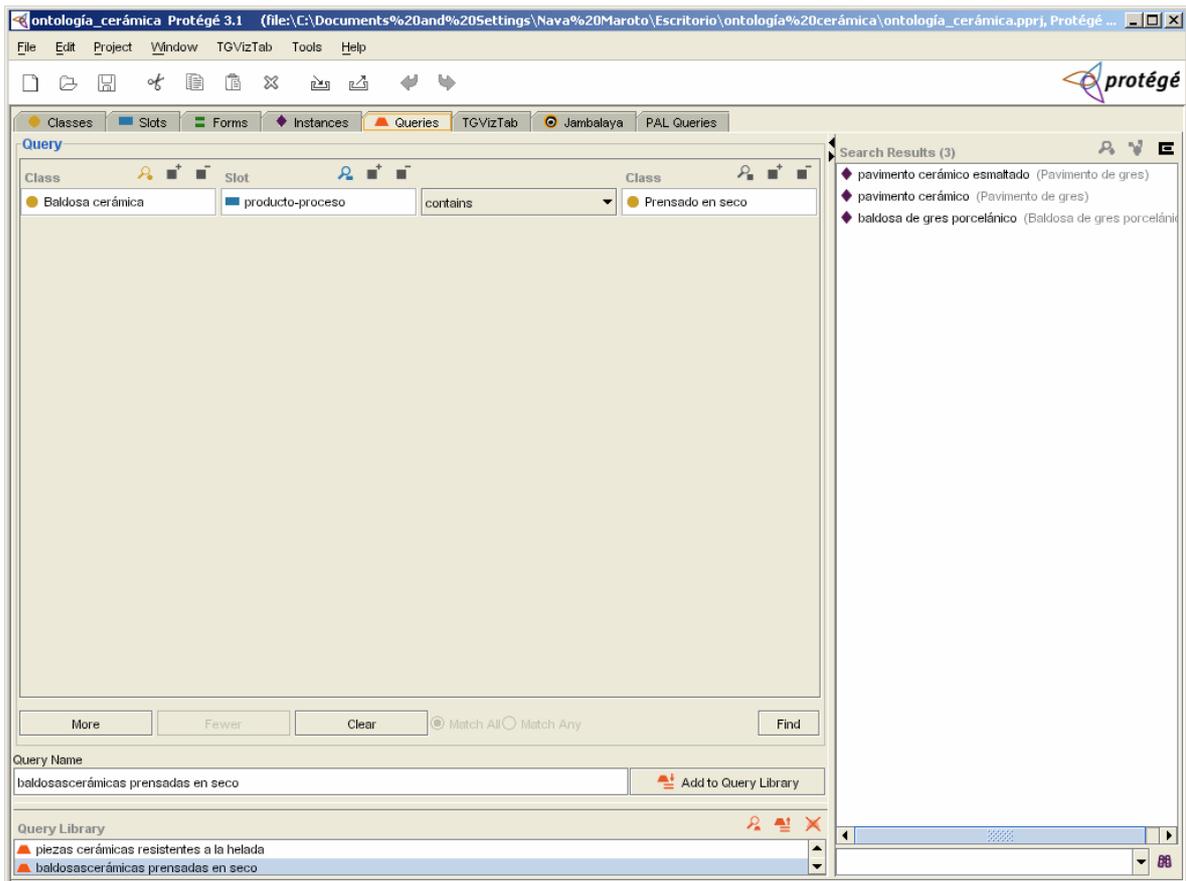


Fig. 50. Resultado de la búsqueda *baldosas cerámicas prensadas en seco*.

Como vemos, la función de consultas *Queries Tab* de *Protégé* permite realizar y almacenar búsquedas de denominaciones representadas mediante el componente *Instances* en nuestra base de conocimiento. Para ello hemos de especificar los criterios que relacionan los conceptos representados por el componente *Class* mediante *slots* de relación. Los resultados obtenidos con el estado actual de desarrollo de *Ontoceram* aún no son muy precisos, y sería necesario seguir desarrollando la base de conocimiento. Sin embargo, sirvan estos ejemplos para dar cuenta de cómo las consultas creadas y ejecutadas mediante la pestaña *Queries Tab* facilitan la recuperación la información sobre las relaciones entre conceptos estructurada con ayuda del editor *Protégé* según el modelo de implementación propuesto.

b) Consultas a través del *plug-in* de consultas en PAL

Además de las consultas sobre el componente *Instance* que facilita la pestaña *Queries Tab*, en *Protégé* podemos realizar consultas o escribir restricciones empleando el lenguaje axiomático PAL (*Protégé Axiom Language*).

Este lenguaje amplía las posibilidades de modelado del conocimiento mediante el editor de ontologías, puesto que permite escribir y almacenar restricciones lógicas y consultas sobre todos los tipos de marcos contenidos en las ontologías diseñadas con *Protégé*, es decir, podemos efectuar consultas sobre el contenido de los componentes *Class*, *Slot* e *Instance*. Una restricción o consulta es una afirmación que cumple determinadas variables lógicas, y que se aplica a un conjunto de valores contenidos en la ontología. Por tanto, una consulta desarrollada con el lenguaje axiomático PAL consiste en un conjunto de variables relacionadas mediante operadores lógicos. *Protégé* incluye un *plug-in* denominado *PAL Queries Tab* que permite diseñar consultas empleando dicho lenguaje.

La explicación exhaustiva de la sintaxis en PAL excede los límites de este trabajo, ya que requiere unos conocimientos avanzados sobre lenguajes de programación. En este apartado nos limitamos a reproducir dos consultas diseñadas con ayuda de D. Ismael Sanz, miembro del grupo de investigación de Bases de Conocimiento Temporal del Departamento de Ingeniería y Ciencias de los Computadores de la Universitat Jaume I.

La primera de ellas es la número 5 ***¿Qué tipo de producto cerámico es una baldosa klinker?*** Para dar respuesta a esta consulta es necesario definir en la base de conocimiento una instancia dependiente de la clase del sistema PAL-QUERY, que se integra en la base de conocimiento al activar el *plug-in* *PAL Queries Tab*. Esta instancia posee las facetas *Name*, que recoge el nombre de la consulta, *Description*, en la que se incluye una descripción de la consulta en lenguaje natural, *Range*, que hace referencia al rango de la consulta, es decir, al tipo de clases a las que se puede aplicar, y *Statement*, que recoge la consulta en lenguaje PAL propiamente dicha. En la Fig. 51 se muestran los valores de la consulta *producto-acabado-baldosa-klinker*, que nos permite recuperar los conceptos superordinados de *Baldosa klinker*.



Fig. 51. Consulta en PAL *producto-acabado-baldosa-klinker*.

Esta consulta se almacena en la base de conocimiento y puede ejecutarse desde la pestaña *PAL Queries Tab*, que debe estar activada en el editor de ontologías. La ejecución de la consulta proporciona los resultados que se muestran en la Fig. 52. En ella podemos ver que *Baldosa de klinker* es un subordinado de *Baldosa de gres industrial de pasta compuesta*, que a su vez es un tipo de *Baldosa de gres*, y así sucesivamente hasta llegar al superordinado *Pieza cerámica*. Como puede verse, el resultado de esta consulta es la cadena de la jerarquía de conceptos en la que se ubica la clase *Baldosa de klinker*.

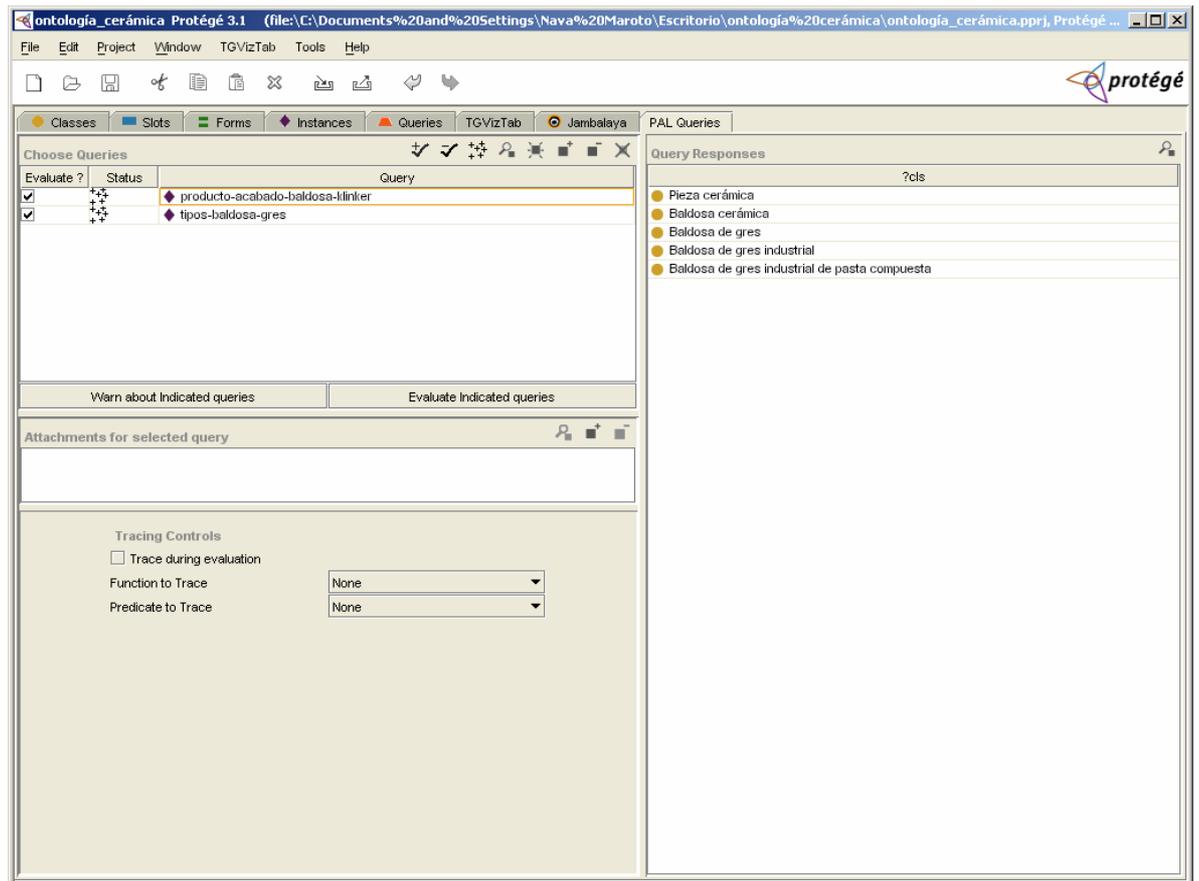


Fig. 52. Pestaña de consultas en lenguaje axiomático PAL.

La principal ventaja de este tipo de restricciones es que permiten realizar consultas sobre cualquiera de los componentes de las bases de datos conceptuales desarrolladas con *Protégé*, es decir, las búsquedas pueden realizarse sobre *Classes*, *Slots* e *Instances*.

Sin embargo, la dificultad añadida que puede suponer el conocimiento del lenguaje PAL dificulta en gran medida la definición de consultas de este tipo. Como veremos en el apartado siguiente, creemos que sería muy útil poder diseñar un plug-in en *Protégé* que agrupara todos los tipos de consultas que hemos visto en esta sección y facilitara su definición por parte de un usuario no experto en lenguajes axiomáticos.

7.3 Propuestas para la adaptación del editor de ontologías al trabajo terminológico automatizado

Si bien *Protégé* es una plataforma fácil de utilizar y que permite la edición de bases de conocimiento de forma relativamente intuitiva, consideramos que se podrían proponer algunas mejoras en el programa, de cara a su utilización para el desarrollo de verdaderas ontologías terminológicas.

- a) En primer lugar, sería muy útil que, cuando establecemos una relación entre dos conceptos representados mediante el componente *Class* a través de un *slot* de relación, el programa fuera capaz de establecer la relación inversa automáticamente una vez que se haya definido cuál es su *slot* inverso. Es decir, que cuando introducimos la relación *objeto-componente funcional* entre los conceptos *Baldosa cerámica* y *Cara vista*, el programa sea capaz de introducir la relación *componente funcional-objeto* entre *Cara vista* y *Baldosa cerámica*.
- b) En segundo lugar, cabría desarrollar una aplicación integrada que permitiera realizar todo tipo de consultas sobre la información almacenada en la base de conocimiento. Como hemos visto, *Protégé* dispone de dos funciones que facilitan la consulta de los datos almacenados a través de las pestañas *Queries Tab* y *PAL Queries Tab*. Sin embargo, la primera sólo nos da como resultados posibles los datos almacenados en el componente *Instance*, y el empleo de la segunda requiere un conocimiento profundo del lenguaje axiomático PAL. Consideramos que todas estas consultas deberían poder hacerse a partir de un único *plug-in* en el que las consultas se pudieran plantear en lenguaje natural, o bien de un modo más sencillo.
- c) En tercer lugar, sería recomendable que el programa *Protégé* fuera capaz de hacer razonamientos sobre las propiedades matemáticas implementadas como facetas de los *slots* de relación.

- d) Por último, de cara al trabajo terminográfico multilingüe, cabría plantearse la posibilidad de que las denominaciones, que en nuestro modelo se representan mediante el componente *Instance*, pudieran introducirse en diversas lenguas, tal vez definiendo una faceta específica para cada una de las lenguas consideradas.

7.4 Recapitulación

En este capítulo se ha descrito el proceso de introducción de datos en la base de conocimiento sobre los productos cerámicos acabados *Ontoceram*, y hemos explorado algunas posibilidades de recuperación de la información sobre relaciones conceptuales contenida en ella.

Nuestra base de conocimiento consta, por el momento, de 621 conceptos que representan productos cerámicos acabados y sus conceptos relacionados. La estructuración de la información propuesta permite establecer un total de 26 relaciones conceptuales distintas entre los conceptos, en los dos sentidos posibles en el caso de las relaciones que poseen una relación inversa. La mayoría de estas relaciones se han implementado mediante *slots* en el editor de ontologías, en los que se hacen explícitas sus propiedades y el tipo de conceptos entre los que pueden establecerse. Asimismo, en la base de conocimiento se han incorporado las denominaciones que se emplean para referirse a cada concepto a través del componente *Instance*. La información se encuentra, por tanto, estructurada de tal modo que puede ser posteriormente recuperada por medios informáticos. Hemos podido comprobar que el modelo de formalización de la información sobre relaciones conceptuales facilita la recuperación de la información de forma onomasiológica. La definición de consultas propuesta por *Protégé* a través de los *plug-ins* de consulta nos ayuda en esta tarea, si bien su uso requiere conocimientos informáticos que dificultan el aprovechamiento de las mismas por un usuario no experto. Por ello, hemos propuesto algunas mejoras en el programa que facilitarían la recuperación de la información contenida en la base de conocimiento.

Por último, hemos propuesto algunas mejoras que servirán para hacer más económico el proceso de introducción de datos y permitirán el desarrollo de bases en las que se consideren las denominaciones en varias lenguas.

En conclusión, en este capítulo hemos demostrado que nuestro modelo de formalización e implementación de las relaciones conceptuales a través de un editor de ontologías puede servir de apoyo en el trabajo terminográfico orientado a la consulta onomasiológica.

**CAPÍTULO 8: RECAPITULACIÓN Y
CONCLUSIONES**

8 Recapitulación y conclusiones

En este último capítulo se resumen los principales resultados y contribuciones de esta tesis doctoral. En primer lugar, retomaremos la observación inicial de la que hemos partido y los objetivos que nos planteábamos al comenzar este trabajo. A continuación, en el apartado 8.1 recopilamos lo presentado en la revisión teórica presentada en los capítulos 2, 3 y 4, que nos ha servido para afrontar el trabajo empírico, descrito en los capítulos 5, 6 y 7, cuyos resultados se recogen en el apartado 8.2. Por último, en el apartado 8.3 se presentan las conclusiones generales de nuestro trabajo y las nuevas vías de investigación que se abren a la luz de los resultados obtenidos.

En la introducción de esta tesis doctoral partíamos de la observación inicial de que, a pesar de la importancia que la terminología otorga a la estructuración sistemática de los conceptos en las áreas de especialidad a través de las relaciones conceptuales, los sistemas y las representaciones conceptuales tradicionales sobre los que se sustentan los trabajos terminológicos resultan insuficientes para una estructuración del conocimiento que permita la recuperación de forma verdaderamente onomasiológica y automatizada. Por un lado, las relaciones conceptuales suelen hacerse explícitas a través de definiciones en lenguaje natural, y esto dificulta la recuperación de la información sobre ellas. Además, no siempre se tienen en cuenta los distintos puntos de vista desde los que se pueden clasificar los conceptos, atendiendo a las diversas perspectivas posibles, es decir, la multidimensionalidad. Por otro lado, las clasificaciones de las relaciones conceptuales tradicionales no suelen definir las relaciones conceptuales de manera formal, sino que se centran en las relaciones jerárquicas como elemento para la estructuración de sistemas conceptuales, desestimando la relevancia de las relaciones no jerárquicas, cuya importancia se atisba, pero no se explota en los trabajos terminológicos.

La ingeniería del conocimiento se ha ocupado en los últimos años del desarrollo de ontologías, entendidas como una conceptualización explícita de la estructuración de un ámbito concreto del saber, consensuada y compartida por un grupo de expertos y reutilizable en diversas aplicaciones informáticas. Las ontologías son uno de los modos más empleados para representar la estructura de

un área de especialidad, y por tanto cumplen una de las funciones básicas que se atribuyen a la terminología: la representación del conocimiento. La ingeniería ontológica ha desarrollado herramientas y métodos para la elaboración de ontologías, que pueden ser aplicables al trabajo terminológico.

Esta observación ha hecho que nos planteemos como principal objetivo de este trabajo la utilización de las herramientas desarrolladas desde la ingeniería del conocimiento, en particular de los editores de ontologías, como herramienta para representar las relaciones conceptuales en el trabajo terminológico automatizado. Los editores de ontologías pueden utilizarse para superar las deficiencias de los sistemas tradicionales de representación conceptual en terminología.

Para lograr el objetivo general, nos hemos fijado cuatro objetivos específicos, que son los siguientes:

1. Desarrollo de un catálogo de relaciones conceptuales completo y aplicable a cualquier área de especialidad
2. Propuesta de un modelo de formalización de las relaciones del catálogo
3. Propuesta de un modelo de implementación de las relaciones en un editor de ontologías
4. Desarrollo de una base de datos conceptual en la que información sobre relaciones conceptuales esté estructurada de forma explícita y sea recuperable mediante un editor de ontologías

Una vez alcanzados estos objetivos hemos valorado la utilidad de nuestras propuestas de formalización e implementación de las relaciones conceptuales para el almacenamiento y recuperación de la información a través de un editor de ontologías.

8.1 *Recapitulación de la revisión teórica*

Por un lado, en el capítulo 2 hemos constatado la importancia conferida a las relaciones conceptuales en la terminología clásica como elemento estructurador del conocimiento especializado. Sin embargo, las relaciones no se estructuran de manera formal, sino que se consideran primitivas axiomáticas, y se pasa por alto la naturaleza compleja de algunas de ellas, como por ejemplo las

relaciones meronímicas. Además, suelen hacerse explícitas a través de las definiciones de lenguaje natural.

Por otro lado, observamos que la atención de las propuestas de clasificación de las relaciones en la terminología clásica se centra en las relaciones jerárquicas, es decir, en las relaciones lógicas y partitivas. Si bien se reconoce la importancia de las relaciones no jerárquicas, estas se presentan de una manera poco sistemática, como una especie de “cajón de sastre” en el que tienen cabida relaciones de diversa naturaleza, sin entrar a considerar su representación de forma explícita.

Frente a la terminología clásica, nuevos enfoques, como el de Sager y Kageura (1994), o la propuesta de Feliu (2004) desde la teoría comunicativa de la terminología proporcionan elementos para la formalización de las relaciones conceptuales, y ofrecen catálogos de relaciones completos y estructurados. En concreto, Sager y Kageura (1994) proponen la consideración de las clases de conceptos (entidades, propiedades, actividades y relaciones) que participan en las relaciones conceptuales como elemento para su formalización. Feliu (2004), por su parte, revisa el formalismo para la representación de las relaciones conceptuales de Otman (1996) y lo expresa mediante la fórmula $a \mathbf{R} b, n$, donde a y b son los conceptos relacionados y \mathbf{R} representa el vínculo que los une. El parámetro n nos permite reflejar que algunas relaciones conceptuales pueden establecerse entre un concepto a y n conceptos b , siendo n mayor o igual que uno. En nuestro trabajo hemos retomado esta fórmula y el modelo de clases conceptuales para caracterizar las relaciones conceptuales.

En el capítulo 2 hemos reflejado la complejidad de algunas relaciones, como la hiponimia, la meronimia y las relaciones causales, y hemos revisado las definiciones de las propiedades de las relaciones aportadas desde la Semántica Léxica. Todos los aspectos presentados en este capítulo nos han proporcionado una panorámica amplia de las relaciones conceptuales. Gracias a esta revisión hemos sido capaces de proponer un catálogo de relaciones conceptuales completo en el trabajo empírico.

En el capítulo 3 se han presentado las propuestas de representación de las relaciones conceptuales hechas proporcionadas en los últimos dos decenios desde la Terminología y la Lingüística. Así, se han podido constatar las limitaciones de

los sistemas conceptuales tradicionales, que dificultan una estructuración verdaderamente formalizada de la información conceptual.

La creación de bases de conocimiento terminológicas y el recurso a las ontologías para la estructuración de los conceptos en terminología suponen un avance considerable hacia una representación del conocimiento que tenga en cuenta distintos puntos de vista a la hora de recuperar la información, es decir, que considere la multidimensionalidad. Estos nuevos modos de representación dan cuenta de las relaciones no jerárquicas, e incluso aportan metodologías detalladas para la creación de ontologías terminológicas, como es el caso de CAOS (Madsen *et al*, 2001) y la termontografía (Kerremans y Temmerman, 2003). No obstante, casi todas estas formas de estructurar la terminología con la ayuda de ontologías emplean herramientas creadas *ad hoc* para la terminología, como ONTOTERM® (Moreno Ortiz, 2002), que es un sistema integrado que permite la estructuración conceptual orientada al trabajo terminológico.

En el último apartado del capítulo 3 se han revisado las bases léxicas como WordNet y EuroWordNet, así como las propuestas de la semántica de marcos y la semántica ontológica, que, procedentes de la Lingüística, suponen un avance en la representación de las relaciones semánticas entre palabras por medios informáticos. De la semántica ontológica hemos tomado prestados para nuestro trabajo empírico los niveles superiores de la ontología Mikrokosmos, que posteriormente nos han servido para estructurar los conceptos en nuestra base de conocimiento de los productos cerámicos acabados.

La observación de los métodos de trabajo y las propuestas de formalización mediante bases de conocimiento en Terminología y Lingüística nos han sido de gran utilidad para afrontar nuestro trabajo empírico, si bien nuestro objetivo es la optimización y adaptación de las herramientas creadas desde la ingeniería del conocimiento para el trabajo terminológico automatizado.

Una vez revisadas las formas de representación de las relaciones conceptuales en la Terminología y la Lingüística, en el capítulo 4 se han presentado los principios y herramientas de la ingeniería del conocimiento, y en concreto de la ingeniería ontológica. El objeto de estudio de esta nueva rama del saber son las ontologías, entendidas como una conceptualización explícita y consensuada de los fenómenos del mundo, que puede ser reutilizada y compartida por distintos grupos de personas y aplicaciones informáticas. Se han presentado

brevemente los principales lenguajes ontológicos, que permiten definir clases organizadas en forma de taxonomía, atributos y relaciones entre ellas, con especial atención al mantenimiento de la herencia entre clases dentro de la jerarquía de clases.

Se han descrito las principales metodologías para el desarrollo de ontologías, presentando de manera más exhaustiva la metodología METHONTOLOGY (Gómez Pérez *et al*, 2004), que propone que la elaboración de una ontología se realice llevando a cabo una serie de tareas que garantizan su desarrollo sistemático. Esta metodología tiene muchos rasgos comunes con el proceso seguido en el trabajo terminológico, y nos ha aportado algunas claves para proponer nuestra propia metodología de trabajo.

Se han descrito y comparado los distintos tipos de herramientas diseñadas desde la ingeniería ontológica, centrándonos en los editores de ontologías. De entre los editores disponibles, nos hemos decantado en nuestro trabajo por el editor *Protégé* en su versión de marcos, ya que es fácilmente accesible y ampliable, tiene un elevado número de usuarios y permite gestionar ontologías de gran tamaño con independencia del lenguaje ontológico empleado. Además, *Protégé* posee una arquitectura abierta y modular mediante la que es posible ampliar sus funciones por medio de *plug-ins*. Por otro lado, la facilidad de manejo y la accesibilidad de esta herramienta han determinado que la consideremos un instrumento adecuado para su integración en el proceso de gestión terminológica.

8.2 Recapitulación y aportaciones del trabajo empírico

Tras revisar las propuestas de clasificación y representación de las relaciones conceptuales, y una vez presentadas la ingeniería ontológica y sus herramientas, hemos afrontado el trabajo empírico. Dicho trabajo se recoge en los capítulos 5, 6 y 7, cuyos principales resultados y aportaciones se resumen a continuación.

En el capítulo 5 hemos presentado las fases del trabajo terminológico y los recursos y herramientas empleados para llevarlo a cabo. Las fases en las que se ha desarrollado se corresponden con los objetivos específicos que nos habíamos planteado en este trabajo, y son las siguientes:

En primer lugar, hemos elaborado de un catálogo de relaciones conceptuales completo, teniendo en cuenta las clasificaciones revisadas en el capítulo 2. El catálogo consta de 26 relaciones conceptuales estructuradas en los siguientes cinco bloques, atendiendo a la naturaleza de la relación:

1. Relaciones lógicas

- a) Superordinado-subordinado
- b) Subordinado-subordinado

2. Relaciones meronímicas

- a) Componente funcional-objeto
- b) Miembro-colección
- c) Porción-Masa
- d) Material-objeto
- e) Etapa-proceso
- f) Característica-actividad
- g) Lugar-área
- h) Parte-parte

3. Relaciones secuenciales

- a) Concepto-lugar
- b) Concepto-lugar al que se dirige
- c) Concepto-concepto simultáneo
- d) Concepto anterior-concepto posterior

4. Relaciones argumentales y circunstanciales

- a) Proceso-agente
- b) Proceso-paciente
- c) Proceso-producto
- d) Proceso-estado
- e) Causa-efecto
- f) Proceso-instrumento
- g) Proceso-método
- h) Objeto-uso

5. Otras relaciones

- a) Fenómeno-medida
- b) Objeto-característica
- c) Relación de sinonimia

d) Relación asociativa

Para la estructuración del catálogo se han tomado como referencia los trabajos de Sager (1990), Feliu (2004) y Dancette y L'Homme (2004). Hemos propuesto denominar las relaciones teniendo en cuenta las clases de conceptos que intervienen en ellas, con la finalidad de que su nombre sea transparente.

Las relaciones se han definido en primer lugar en lenguaje natural, y posteriormente las hemos descrito de manera formal, teniendo en cuenta los siguientes aspectos:

- a) Las clases conceptuales que intervienen en la relación, que pueden ser entidades, actividades y propiedades.
- b) Las propiedades de las relaciones, en particular, la simetría, la transitividad y la recursividad. En el caso de las relaciones meronímicas, se ha tenido también en cuenta si las relaciones son homeoméricas, funcionales, simultáneas y separables o no.
- c) Si poseen una relación inversa o no. En caso de que la posean, también se ha considerado la relación en el sentido contrario al propuesto inicialmente.

Teniendo en cuenta los componentes y funciones del editor de ontologías *Protégé*, hemos propuesto un modelo de implementación de los conceptos, relaciones y denominaciones en las base de conocimiento creada co dicho editor. Así, los conceptos se representan en *Protégé* mediante el componente *Class*, las denominaciones mediante el componente *Instance* y las relaciones mediante los denominados *Slots*.

Al final del capítulo 6 se ofrece una tabla (Tabla 45) que reproduce el catálogo completo de relaciones en el que se especifica cómo se formalizan e implementan éstas en el editor de ontologías. Para cada relación se facilita la siguiente información:

- a) Descripción del significado de la relación en lenguaje natural
- b) Clases conceptuales que participan a uno y otro lado de la relación
- c) Propiedades matemáticas de cada relación y propiedades específicas de las relaciones meronímicas
- d) Implementación en *Protégé*
- e) Ejemplos extraídos de nuestro análisis empírico

Las principales aportaciones que, en nuestra opinión, ofrece nuestro catálogo de relaciones conceptuales son las siguientes:

- a) El catálogo propuesto es exhaustivo y válido para cualquier área de especialidad. Además, en caso necesario, el modelo de descripción de las relaciones conceptuales propuesto posibilita la aplicación del catálogo con nuevas relaciones conceptuales específicas del ámbito temático.
- b) La división del catálogo en cinco grupos de relaciones permite superar la escasa estructuración de las relaciones no jerárquicas que hemos detectado en algunas de las clasificaciones de las relaciones conceptuales clásicas.
- c) En nuestra propuesta de formalización las relaciones se presentan de un modo estructurado, lo que facilitará su implementación en un editor de ontologías.
- d) El modo en el que se han denominado las relaciones resulta transparente, y esto facilita la comprensión de la naturaleza de las mismas.
- e) El hecho de considerar las relaciones en sus dos direcciones posibles enriquece la representación conceptual, y permitirá recuperar la información sobre las relaciones conceptuales en cualquiera de los dos sentidos de forma onomasiológica.

En el capítulo 7 se ha descrito el proceso de introducción de los datos en la base de conocimiento *Ontoceram* en la que hemos recogido los conceptos que representan a los productos cerámicos acabados y sus conceptos relacionados, vinculados mediante las relaciones expresadas mediante *slots*, así como las denominaciones identificadas durante el análisis empírico.

Nuestra base de datos, creada con el editor de ontologías *Protégé*, recoge en este momento 621 conceptos sobre productos cerámicos acabados y 665 denominaciones posibles para dichos conceptos, entre los que se establecen las relaciones conceptuales propuestas en nuestro catálogo.

El proceso de introducción de datos en *Ontoceram* se ha llevado a cabo siguiendo una metodología que queda recogida en un protocolo que servirá como guía para la ampliación de esta u otras bases de conocimiento (presentado en el

apartado 7.1.1. El análisis de las relaciones conceptuales se realiza a partir de la observación de los contextos procedentes del corpus *TXTCERAM* y de la base de datos *Cerámica*. Los conceptos relacionados identificados durante este proceso de análisis se reflejan en una tabla de *Excel*, en la que se organiza la información recogida y se deja preparada para su introducción en la base de datos conceptual. Esta tabla, inspirada en las tablas desarrolladas en algunas de las tareas propuestas en METHONTOLOGY, permite recoger de un modo visual cada concepto junto con sus conceptos relacionados.

Una vez recogida la información *in vivo*, se ha llevado a cabo la introducción de ésta en la base de conocimiento *Ontoceram*. La organización de los conceptos de forma jerárquica y el mantenimiento de la herencia de atributos y relaciones –representados mediante *slots*– entre los conceptos situados en un nivel superior de la jerarquía y los que dependen de aquellos permite desarrollar la base de datos conceptual de un modo económico. Además, al restringir el rango de cada relación, es decir, la clase de conceptos que puede aparecer a la derecha de la relación, se simplifica el proceso de introducción de los valores de cada relación.

En la segunda parte del capítulo 7 se ha evaluado la adecuación de nuestras propuestas de formalización e implementación. Para ello se han diseñado una serie de consultas onomasiológicas que ejemplifican los distintos tipos de relación contemplados en nuestro catálogo. De este modo hemos comprobado que la manera en que se ha estructurado la información conceptual sobre las relaciones proporciona, al menos potencialmente, la respuesta a estas consultas. Hemos explorado brevemente las funciones de búsqueda que ofrece el editor de ontologías *Protégé*, mediante las pestañas de consultas *Queries Tab* y *PAL Queries Tab*.

La fase de introducción de datos y la posterior valoración de la adecuación del modelo presentado para la recuperación onomasiológica de la información sobre relaciones conceptuales nos han llevado a proponer algunas adaptaciones del editor de ontologías *Protégé* que podrían facilitar el trabajo terminológico.

Así, creemos que sería útil desarrollar un *plug-in* que facilitara la realización de consultas, integrando las consultas sobre *instances*, *classes* y *slots* en una única interfaz en la que sea necesario recurrir a un lenguaje axiomático para plantear consultas complejas y en lenguaje natural.

Otra de las adaptaciones que facilitaría su empleo para el desarrollo de ontologías terminológicas es la posibilidad de que el editor tenga en cuenta automáticamente las relaciones inversas cuando se introducen las relaciones en un sentido, ya que esto economizaría en gran medida la labor de introducción de datos.

Podría adaptarse el editor de ontologías para introducir denominaciones para los conceptos en varias lenguas, asociando etiquetas de idioma a los conceptos.

Por último, creemos que sería muy útil poder combinar el editor de ontologías con otras herramientas para la gestión de la terminología, como las bases de datos terminológicas tradicionales desarrolladas con programas como *Multiterm*.

De momento se ha introducido en la base de datos conceptual información sobre la transitividad, la recursividad y la simetría de las relaciones de las relaciones. Sin embargo, esta información aún no es aprovechable para hacer inferencias y razonamientos automatizados. Tomemos como ejemplo la relación componente funcional-objeto, que hemos definido como una relación transitiva. La funcionalidad y del editor mejoraría si pudiera deducir que si una *arista* es uno de los componentes de las *baldosas cerámicas* y las *baldosas cerámicas* fueran a su vez componentes de un objeto superior, entonces *arista* sería también un componente de dicho objeto superior.

La principal aportación teórica de esta tesis doctoral es la formalización de un catálogo de relaciones completo y aplicable a distintos ámbitos temáticos, es decir, en el que las relaciones quedan definidas tanto en lenguaje natural como desde el punto de vista formal. Cada relación se caracteriza por las clases conceptuales a las que pertenecen los conceptos vinculados y por las propiedades matemáticas que posee.

Desde el punto de vista metodológico, esta tesis aporta un método de trabajo para el análisis manual de las relaciones conceptuales a partir de las definiciones propuestas en diccionarios y otras fuentes terminológicas y de las concordancias de los términos extraídas a partir de un corpus textual digitalizado. De este modo se combinan el análisis de los conceptos *in vivo* e *in vitro*.

Asimismo, proponemos un modelo práctico para la implementación de las relaciones formalizadas en el editor de ontologías *Protégé*, que podría aplicarse en otros trabajos. La estructuración de la información conceptual mediante el editor de ontologías nos permite realizar consultas verdaderamente onomasiológicas, teniendo en cuenta los múltiples puntos de vista que se pueden adoptar a la hora de recuperar la información, es decir, las distintas dimensiones desde las que pueden observarse los conceptos.

Consideramos que los editores de ontologías pueden integrarse, con algunas adaptaciones, en el proceso terminológico automatizado, ya que facilitan la estructuración del conocimiento de manera explícita y recuperable por medios informáticos.

Las bases de conocimiento desarrolladas con el editor de ontologías podrían emplearse no sólo en el trabajo terminológico, sino tal vez en otras aplicaciones que puedan servirse de ontologías. Por ejemplo, podrían ser de utilidad para sistemas automatizados para determinar cuál es el producto cerámico más adecuado para un uso concreto. O tal vez podrían emplearse como base para aplicaciones informáticas de gestión de los distintos productos cerámicos disponibles en un almacén. En definitiva, el desarrollo de una verdadera ontología de la industria cerámica podría tener aplicación no sólo desde el punto de vista de la gestión de la terminología, sino también de los expertos.

8.3 Conclusiones y perspectivas de investigación futura

Consideramos que en el presente trabajo se han cumplido los objetivos que se habían propuesto en la introducción. El objetivo general era formalizar e implementar las relaciones entre los conceptos de un ámbito de especialidad empleando un editor de ontologías estándar, en nuestro caso *Protégé*.

Asimismo, se han cumplido los objetivos específicos. Se ha desarrollado un catálogo de relaciones conceptuales bastante completo y que es de aplicación en cualquier ámbito temático. A continuación se han formalizado las relaciones de dicho catálogo, y se ha propuesto un modelo de implementación de las mismas empleando el editor de ontologías.

En tercer lugar, se han introducido la información sobre los conceptos relacionados con los productos cerámicos acabados en la base de conocimiento *Ontoceram*, y por último se han propuesto una serie de consultas onomasiológicas

a las que podemos dar respuesta con la información introducida en la base de datos conceptual. Estas consultas nos han permitido comprobar la utilidad de nuestra propuesta, así como proponer adaptaciones al editor de ontologías que mejorarán su utilización para la estructuración conceptual en el trabajo terminográfico.

La investigación realizada en el marco de esta tesis doctoral es necesariamente limitada, y existen algunos aspectos en los que sería interesante profundizar en el futuro.

En primer lugar, hemos limitado la elaboración de la base de conocimientos a los productos cerámicos acabados. En el marco del proyecto ONTODIC está previsto el desarrollo de una ontología terminológica del ámbito de la cerámica. Nuestra propuesta de implementación de las relaciones conceptuales en el editor de ontologías y la metodología aquí presentada pueden resultar de utilidad para el desarrollo de una ontología cerámica completa (Alcina, en prensa).

Un segundo aspecto en el que sería necesario profundizar es la formalización de las características de los conceptos en el editor de ontologías. En nuestro trabajo hemos recogido la relación objeto-característica, que vincula una entidad con sus propiedades, pero no hemos tenido en cuenta los valores concretos que estas características pueden tener. Así, por ejemplo, hemos considerado la característica *absorción de agua* para todos los productos cerámicos acabados, pero no hemos considerado los valores concretos que puede tener esta propiedad en cada tipo de producto cerámico, que será un valor numérico expresado en porcentaje.

Hasta el momento hemos propuesto un modelo de implementación que permite representar las propiedades matemáticas de las relaciones, como la simetría, la transitividad y la recursividad, pero creemos que sería conveniente profundizar en la forma de razonar a partir de dichas propiedades matemáticas, con el fin de facilitar tanto el desarrollo como la consulta de las bases de datos conceptuales con ayuda del editor.

En nuestro trabajo, el análisis conceptual y la detección de relaciones se han llevado a cabo de forma manual. No obstante, en el marco de los proyectos TXTCERAM y ONTODIC se han iniciado trabajos para la elaboración de un catálogo de marcadores léxicos que permitan extraer la información sobre las relaciones conceptuales propuestas en nuestro catálogo.

Por último, nuestro trabajo empírico propone adaptaciones del editor de ontologías que mejorarían su empleo en el trabajo terminológico automatizado. Creemos necesario profundizar en el planteamiento de dichas mejoras en el programa, en su integración con otros programas empleados en la gestión terminológica, con el fin de desarrollar un entorno de trabajo automatizado para el procesamiento de la Terminología que tenga en cuenta tanto los aspectos de representación conceptual como los aspectos lingüísticos y comunicativos de los términos.

9. BIBLIOGRAFÍA

9 Bibliografía

- AENOR (1979): *Vocabulario de la Terminología*. Concuenda con: ISO 1087. Madrid: IRANOR, 18 pp.
- (1991): *Principios y métodos de la terminología*. Concuenda con: ISO 704-1981. Madrid: AENOR. 16 pp.
- AHMAD, Kursid y Heather FULFORD (1992): *Knowledge Processing 4: Semantic Relations and Their Use in Elaborating Terminology*. Guildford: University of Surrey.
- AHMAD, Kursid y S. HOOK (1992): *Knowledge Processing 10. Conceptual Graphs and Term Elaboration: Explicating (Terminological) Knowledge*. Guildford: University of Surrey.
- AHMAD, Kursid (2001): "The Role of Specialist Terminology in Artificial Intelligence and Knowledge Acquisition". En: Sue Ellen WRIGHT y Gerhard BUDIN (eds.) (2001): *Handbook of Terminology Management*. Vol. 2. Amsterdam: John Benjamins, pp. 809-844.
- ALCINA, Amparo y Silvia GAMERO (eds) (2002): *La traducción científico-técnica y la terminología en la sociedad de la información*. Castelló: Universitat Jaume I.
- ALCINA, Amparo, Victoria SOLER y Anna ESTELLÉS (2005): "Internet como instrumento para la documentación en terminología y traducción. Hacia las plataformas de recursos electrónicos para el traductor especializado". En: SALES, Dora (ed.) (2005): *La biblioteca de Babel: documentarse para traducir*. Granada: Comares, pp. 221-241.
- ALCINA, Amparo (en prensa): "Metodología y tecnologías para la elaboración de diccionarios terminológicos onomasiológicos". En: ALCINA, Amparo (ed.) (en prensa): *Terminología y sociedad del conocimiento*. Berna: Peter Lang.
- ALONGE, Antonietta (1996): *Definition of the Links and Subsets for Verbs*. Amsterdam: EuroWordNet.

- ANTIA, Bassey E (2000): *Terminology and Language Planning: An Alternative Framework of Practise and Discourse*. Amsterdam/Philadelphia: John Benjamins.
- (2001): "Terminological Investigations into Specialized Knowledge and Texts. A Case Study of Legislative Discourse." *Terminology* 7.1: pp. 5-29.
- ARANO, Silvia (2005): "Los tesauros y las ontologías en la biblioteconomía y la documentación" *Hipertext.net*.3.
- ARNTZ, Reiner y Heribert PICT (1995): *Introducción a la terminología*. Madrid: Fundación Germán Sánchez Ruipérez.
- ARNTZ, Reiner, Heribert PICT y Felix MAYER (2002): *Einführung in Die Terminologearbeit*. Hildesheim: Georg Olms Verlag.
- ARPÍREZ VEGA, Julio César (2000): *Webode 1.0. User's Manual. Draft Version 0.1*. Madrid: Laboratorio de Inteligencia Artificial, Universidad Politécnica de Madrid.
- AUSSENAC, Nathalie (1999): "Géditerm: Un Logiciel Pour Gérer Des Bases De Connaissances Terminologiques." *Terminologies Nouvelles*, 19: 111-123.
- BAKER, Collin F., Charles J. FILLMORE y Beau CRONIN (2003): "The Structure of the Framenet Database." *International Journal of Lexicography* 16.3: 281-296.
- BARRIÈRE, Caroline (1997): *From a Children's First Dictionary to a Lexical Base of Conceptual Graphs*. Simon Fraser University, tesis doctoral inédita.
- (2001): "Investigating the Causal Relation in Informative Texts". *Terminology* 7.2: 135-154.
- . (2002): "Hierarchical Refinement and Representation of the Causal Relation." *Terminology* 8.1: 91-111.
- (2004): "Building a Concept Hierarchy from Corpus Analysis." *Terminology* 10.2: 241-63.

- (2004): "Knowledge-Rich Contexts Discovery." En: *Seventeenth Canadian Conference on Artificial Intelligence (AI'2004)*. Londres, Ontario: National Research Council of Canada.
- BERNERS-LEE, Tim, James HENDLER y Ora LASSILA (2001): "The Semantic Web". *Scientific American* 284.5: 34-44.
- BEJOINT, Henri y Philippe THOIRON (eds.) (2000): *Le sens en terminologie*. Lyon: Presses universitaires de Lyon.
- (2000): "Le Sens Des Termes". En: BEJOINT, Henri y Philippe THOIRON (eds.) (2000): *Le sens en terminologie*. Lyon: Presses universitaires de Lyon, pp. 5-19.
- BIÉBOW, Brigitte y Sylvie SZULMAN (1999): "TERMINAE: A Method and a Tool to Build a Domain Ontology". En: FENSEL, Dieter y Rudi STUDER (eds) (1999): *11th European Workshop on Knowledge Acquisition*. London: Springer Verlag, pp. 49-66.
- BODSON, Claudine (2004): *Termes et relations sémantiques en corpus spécialisés : rapport entre patrons de relations sémantiques (PRS) et types sémantiques (TS)*. Université de Montreal, tesis doctoral inédita.
- BOULANGER, Jean-Claude (1991): "Une lecture socioculturelle de la terminologie". *Cahiers de Linguistique Sociale*, 18: pp.13-30.
- (1995): "Présentation: images et parcours de la socioterminologie". *Meta* 40.2, pp. 194-205.
- BOURIGAULT, Didier, Christian JACQUEMIN y Marie Claude L'HOMME (eds.) (2001): *Recent Advances in Computational Terminology. Natural Language Processing*. Amsterdam/Philadelphia: John Benjamins.
- (1997): "Multidimensional Classification of Concepts and Terms" En: WRIGHT, SUE ELLEN y Gerhard BUDIN (eds.): *Handbook of Terminology Management*. Amsterdam: John Benjamins, pp. 133-43.

- BOWKER, Lynne y Jennifer PEARSON (2002): *Working with Specialized Language. A Practical Guide to Using Corpora*. London: Routledge.
- BUDIN, Gerhard y E. OESER (1995): "Controlled Conceptual Dynamics: From "Ordinary Language" to Scientific Terminology -and Back." *Terminology Science and Research* 6.2: pp. 3-17.
- CABRÉ, María Teresa (1993): *La terminología. Teoría, metodología, aplicaciones*. Barcelona: Antártida.
- CABRÉ, María Teresa (1995): "La terminología hoy: concepciones, tendencias y aplicaciones." *Ciencia da Informaçao* 24.3 .
- (1996): "Lexicología y variación: hacia un modelo integrado". En: *V Simposio Iberoamericano de terminología RITERM*. Ciudad de México, 1996.
- (1996): *Terminologia: Selecció de textos d'Eugen Wüster*. Barcelona: Universitat de Barcelona.
- CABRÉ, María Teresa (1999): *Representación y comunicación*. Barcelona: IULA.
- (2000): "Terminologie et linguistique: la théorie des portes" *Terminologies Nouvelles*. 21 : pp. 10-15.
- CABRÉ, María Teresa y Judit FELIU (2001): *Terminología y cognición*. Barcelona: IULA.
- CABRÉ, María Teresa, Judit FELIU y Jordi VIVALDI (2004): "Base de connaissances Genoma: le rôle de l'ontologie". Workshop on Terminology, Ontology and Knowledge Representation. Université Jean-Moulin, Lyon 3, 2004.
- CABRÉ, María Teresa (2004): "TERMINTEGRAL: una plataforma para la construcción de bases terminológicas y ontologías". *Linguistica Antverpiensia*, NS nº 3: 245-61.
- CABRÉ, María Teresa *et al* (2004): "The GENOMA-KB Project: Towards the Integration of Concepts, Terms, Textual Corpora and Entities". *IV International Conference on Language Resources and Evaluation*. LREC 2004. Lisboa: European Language Resources Association, pp. 87-90.

- CHAFFIN, Roger y Douglas J. HERRMANN (1984): "The Similarity and Diversity of Semantic Relations". *Memory and Cognition* 12.2: pp. 134-141.
- (1988): "The Nature of Semantic Relations: A Comparison of Two Approaches." En: EVENS, M. W. (1988): *Relational Models of the Lexicon*. Cambridge: Cambridge University Press, pp. 289-334.
- CHAFFIN, Roger, Douglas J. HERRMANN y Morton E. WINSTON (1988): "An Empirical Taxonomy of Part-Whole Relations: Effects of Part-Whole Relation Type on Relation Identification". *Language and Cognitive Processes* 3.1: pp. 17-48.
- CIVERA GARCÍA, Pilar y Amparo ALCINA (2000): "elaboración de una herramienta de comunicación plurilingüe para el sector cerámico: un diccionario terminológico multilingüe en soporte informático". XL Congreso de la Sociedad Española de Cerámica y Vidrio. Onda (Castellón), 2000.
- CIVERA GARCÍA, Pilar (2002): "Traducción científico-técnica y terminología en el sector de la industria cerámica". En: ALCINA, Amparo y Silvia GAMERO (eds.) (2002): *La traducción científico-técnica y la terminología en la sociedad de la información*. Castellón: Publicacions de la Universitat Jaume I, pp. 167-76.
- CLARK, Kendall G. (2003): "The Semantic Web Is Closer Than You Think". Disponible en: <http://www.xml.com/pub/a/2003/08/20/deviant.html>. [Consulta: junio de 2006].
- CLAVERÍA, Gloria y Joan TORRUELLA (1993): "Formación de términos en los léxicos especializados de la lengua española". En: SAGER, Juan Carlos (1993). *Curso práctico sobre el procesamiento de la terminología*. Madrid: Fundación Germán Sánchez Ruipérez, pp. 315-344.
- CLIMENT, Salvador, Horacio RODRÍGUEZ y Julio GONZALO (1996): *Definition of the Links and Subsets for Nouns of the EuroWordNet Project*. Amsterdam: EuroWordNet.
- CONDAMINES, Anne y Josette REBEYROLLE (2001): "Searching for and Identifying Conceptual Relationships Via a Corpus-Based Approach to a Terminological Knowledge Base (CKTB): Method and Results". En BOURIGAULT, Didier,

Christian JACQUEMIN y Marie Claude L'HOMME (eds.) (2001): *Recent Advances in Computational Terminology. Natural Language Processing*. Amsterdam/Philadelphia: John Benjamins, pp. 127-148.

CONDAMINES, Anne (2002): "Corpus Analysis and Conceptual Relation Patterns". *Terminology* 8.1: pp. 141-162.

CORDÓN GARCÍA, José Antonio, Jesús LÓPEZ LUCAS y José Raúl VAQUERO PULIDO (2001): *Manual de investigación bibliográfica y documental. Teoría y práctica*. Madrid: Pirámide.

CRUSE, D. Allan (1979): "On the Transitivity of the Part-Whole Relation". *Journal of Linguistics*, 15: pp. 29-38.

--- (1986): *Lexical Semantics*. Cambridge: Cambridge University Press.

--- (2002): "Hyponymy and Its Varieties". En: GREEN, Rebecca, Carol A. BEAN y Sungo Hyon MYAENG (eds.) (2002): *The Semantics of Relationships: An Interdisciplinary Approach*. Dordrecht: Kluwer Academic Publishers, pp. 3-21.

--- (2004): *Meaning in Language*. 2nd ed. Oxford: Oxford University Press.

CURRÁS, Emilia (2005): *Ontologías, taxonomía y tesauros. Manual de construcción y uso*. Gijón: TREA.

DAHLBERG, Ingetraut (1982): *Classification Systems and Thesauri*. Frankfurt: Indeks Verlag.

DANCETTE, Jeanne y C. RHETORE (2000): *Dictionnaire alytique de la distribution. Analytical Dictionary of Retailing*. Montreal: Presses de l'Université de Montréal.

DANCETTE, Jeanne y Marie Claude L'HOMME (2002): "The Gate to Knowledge in a Multilingual Specialized Dictionary: Using Lexical Functions for Taxonomic and Partitive Relations." EURALEX 2002. Copenhage, pp. 597-606.

--- (2004): "Building Specialized Dictionaries Using Lexical Functions". *Linguistica Antverpiensia*, NS, n° 3: pp. 113-31.

- DAVIDSON, Laura. (1997): *Knowledge Extraction Technology for Terminology*. Tesis de Master. Universidad de Ottawa.
- DEPECKER, Loïc (2000): "Le signe entre signifié et concept". En: BEJOINT, Henri y Philippe THOIRON (eds.) (2000): *Le sens en terminologie*. Lyon: Presses universitaires de Lyon, pp. 86-126.
- DÍEZ ORZAS, Pedro Luis (1999): *La relación de meronimia en los sustantivos del léxico español: contribución a la semántica computacional*. Estudios de Lingüística Española, 2.
- DIK, Simon C. (1989): *The Theory of Functional Grammar. Part I: The Structure of the Clause*. Dordrecht: Foris.
- DUINEVELD, A. J. *et al* (2000): "Wonder Tools? A Comparative Study of Ontological Engineering Tools". *International Journal of Human-Computer Studies*, 52.6 : pp. 1111-1133.
- ESPANHA GOMES, Hagar y Maria Luiza DE ALMEIDA CAMPOS (1996): "Systematic Aspects of Terminology". *Meta* XLI.2: pp. 247-254.
- ESTELLÉS, Anna, Amparo ALCINA y Victoria SOLER (2006): "Retrieving Terminological Data from the TXTCERAM Tagged Domain Corpus: A First Step Towards a Terminological Ontology". *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC'06)*. Génova: ELRA.
- ESTELLÉS, Anna y Amparo ALCINA (en prensa): "Etiquetado del corpus TXTCERAM orientado a la extracción de información conceptual". *XXIV Congreso de AESLA: "Aprendizaje de lenguas, uso del lenguaje y modelación cognitiva: perspectivas aplicadas entre disciplinas"*, Madrid.
- EVENS, Martha W. (ed.) (1988): *Relational Models of the Lexicon. Representing Knowledge in Semantic Networks*. Cambridge: Cambridge University Press.
- FABER, Pamela (1999): "Conceptual Analysis and Knowledge Acquisition in Scientific Translation". *Terminologie et Traduction*, 2: pp. 97-123.

- FABER, Pamela y Maribel TERCEDOR SÁNCHEZ (2001): "Codifying Conceptual Information in Descriptive Terminology Management". *Meta* XLVI.1: pp. 192-204.
- FABER, Pamela, Clara Inés LÓPEZ RODRÍGUEZ y Maribel TERCEDOR SÁNCHEZ (2001): "Utilización de técnicas de corpus en la representación del conocimiento médico". *Terminology* 7.2: pp.167-198.
- FABER, Pamela (2002): "Investigar en terminología". En: FABER, Pamela y Catalina JIMÉNEZ (2002): *Investigar en terminología*. Granada: Comares, pp. 3-23.
- FABER, Pamela y Catalina JIMÉNEZ (2002): *Investigar en terminología*. Granada: Comares.
- FABER, Pamela (2002): "ONCOTERM: sistema bilingüe de información y recursos oncológicos". En: ALCINA, Amparo y Silvia GAMERO (eds) (2002): *La traducción científico-técnica y la terminología en la sociedad de la información*. Castelló: Universitat Jaume I, pp. 177-188.
- FABER, Pamela *et al* (2006): "Process-Oriented Terminology Management in the Domain of Coastal Engineering". *Terminology* 12.2: pp. 189-213.
- FAULSTICH, Enilde (1998/1999): "Principes Formels Et Fonctionnels De La Variation En Terminologie". *Terminology* 5.1: pp. 93-106.
- (2002): "Variação em terminologia. Aspectos de socioterminologia". En: GUERRERO RAMOS, Gloria y Manuel Fernando PÉREZ LAGOS (eds.) (2002): *Panorama actual de la terminología*. Granada: Comares, pp. 65-91.
- FELBER, Helmut (1984): *Terminology Manual*. París: UNESCO-Infoterm.
- FELBER, Helmut y Gerhard BUDIN (1989): *Terminologie in Theorie Und Praxis*. Tübingen: Günter Narr.
- FELIU, Judit (2000): *Relacions Conceptuals I Variació Funcional: Elements Per a Un Sistema De Detecció Automàtica*. Universitat Pompeu Fabra, tesina inédita.

- FELIU, Judit y María Teresa CABRÉ (2002): "Conceptual Relations in Specialized Texts: New Typology and an Extraction System Proposal". *6th International Conference on Terminology and Knowledge Engineering*. Nancy, Francia.
- FELIU, Judit, Jordi VIVALDI y María Teresa CABRÉ (2002): *Ontologies: A Review*. Barcelona: Institut Universitari de Lingüística Aplicada.
- FELIU, Judit (2004): *Relacions conceptuals i terminologia: anàlisi i proposta de detecció semiautomàtica*. Universitat Pompeu Fabra, tesis doctoral inédita.
- FELLBAUM, Christiane (ed.) (1998): *WordNet: An Electronic Lexical Database. Language, Speech and Communication*. Cambridge, Massachusetts: MIT Press.
- (1998): "A Semantic Network of English Verbs". En: FELLBAUM, Christiane (ed.) (1998): *Wordnet: An Electronic Lexical Database. Language, Speech and Communication*. Cambridge, Massachusetts: MIT Press, pp. 69-104.
- (2002): "On the Semantics of Troponymy". En: GREEN, Rebecca, Carol A. BEAN y Sungo Hyon MYAENG (eds.) (2002): *The Semantics of Relationships: An Interdisciplinary Approach*. Dordrecht: Kluwer Academic Publishers, pp. 23-34.
- FILLMORE, Charles J (1982): "Frame Semantics". *Linguistics in the Morning Calm*. Seúl: Hanshin, pp. 111-137.
- (1985): "Frames and the Semantic of Understanding", *Quaderni di Semantica* 6.2: pp. 222-54.
- FILLMORE, Charles J., Christopher R. JOHNSON y Miriam R. L. PETRUCK (2003): "Background to Framenet". *International Journal of Lexicography* 16.3: 235-250.
- GAMBIER, Yves (1987): "Problèmes terminologiques des pluies acides: pour une socio-terminologie". *Meta* 32.3: pp. 314-320.
- (1991): "Présupposés de la terminologie: vers une remise en cause". *Cahiers de Linguistique Sociale* 18: pp. 31-58.

- GARCIA, Daniela (1998): *Analyse automatique des textes pour l'organisation causale des actions. Réalisation du système informatique COATIS*. Universidad de la Sorbona IV, tesis doctoral inédita.
- GARCÍA DE QUESADA, Mercedes (2001): *Estructura definicional terminográfica en el subdominio de la oncología clínica*. Estudios de Lingüística Española. Disponible en: <http://elies.rediris.es/elies14/index.html#indice>. [Consulta: julio de 2006].
- GARCÍA JIMÉNEZ, Antonio (2004): "Instrumentos de representación del conocimiento: tesauros versus ontologías". *Anales de Documentación*, 7: pp. 79-95.
- GAUDIN, François (2003): *Socioterminologie: une approche sociolinguistique de la terminologie*. Bruxelles: de Boeck-Ducolot.
- GILLAM, Lee y Mariam TARIQ (2004): "Ontology Via Terminology?". En IBEKWE-SAN JUAN, Fidelia (2004): *Workshop on Terminology, Ontology and Knowledge Representation*. Lyon: Université de Lyon.
- GILLAM, Lee, Mariam TARIQ y Kursid AHMAD (2005): "Terminology and the Construction of Ontology". *Terminology* 11.1: pp. 55-81.
- GOMEZ PEREZ, Asunción (1999): "Développements récents en matière de conception, de maintenance et d'utilisation des ontologies". *Terminologies Nouvelles* 19: pp. 9-20.
- GÓMEZ PÉREZ, Asunción, Mariano FERNÁNDEZ LÓPEZ y Óscar CORCHO (2004): *Ontological Engineering*. London: Springer Verlag.
- GÓMEZ PÉREZ, Asunción *et al* (2006) [en línea]: *Ontoroadmap*. Disponible en: <http://babage.dia.fi.upm.es/ontoweb/wp1/OntoRoadMap/index.html>. [Consulta: junio de 2006].
- GRABAR, Natalia *et al* (2004): "Repérage de relations terminologiques transversales en corpus". Congreso TALN 2004. Fez, Marruecos.
- GRABAR, Natalia y Pierre ZWEIGENBAUM (2004): "Lexically-Based Terminology Structuring". *Terminology* 10.1: pp. 23-53.

- GREEN, Rebecca, Carol A. BEAN y Sungo Hyon MYAENG (eds.) (2002): *The Semantics of Relationships: An Interdisciplinary Approach*. Dordrecht: Kluwer Academic Publishers.
- GRUBER, Thomas R. (1993): *Toward Principles for the Design of Ontologies Used for Knowledge Sharing*. Palo Alto: Stanford Knowledge Systems Laboratory.
- GRÜNINGER, M. y M.S. FOX (1995): "Methodology for the Design and Evaluation of Ontologies" *IJCAI'95. Workshop on Basic Ontological Issues in Knowledge Sharing*. Montréal. 6.1-6.10.
- GUARINO, Nicola y Pierdaniele GIARETTA (1995): "Ontologies and Knowledge Bases: Towards a Terminological Clarification". En: MARS, N.J.I.(ed.) (1995): *Towards Very Large Knowledge Bases: Knowledge Building and Knowledge Sharing*. Amsterdam: IOS Press, pp. 25-32.
- GUARINO, Nicola y C. WELTY (2000): "A Formal Ontology of Properties". En: DIENG, R. y O. CORBY (2000): *12th International Conference in Knowledge Engineering and Knowledge Management (EKAW'00)*. Juan-Les-Pins, Francia: Springer Verlag, pp. 97-112.
- GUERRERO RAMOS, Gloria y Manuel Fernando PÉREZ LAGOS (2002): *Panorama actual de la terminología*. Interlengua. Granada: Comares.
- HALIDAY, Michael Alexander Kirkwood (1985): *An Introduction to Functional Grammar*. Londres: Edward Arnold.
- HANN, Michael (1978): *The Application of Computers to the Production of Systematic, Multilingual, Specialised Dictionaries and the Accessing of Semantic Information Systems*. Manchester: UMIST.
- HEARST, Marti A. (1992): "Automatic Acquisition of Hyponyms from Large Text Corpora". En: BOITET, CHRISTIAN (1992): *COLING-92. 15th International Conference on Computational Linguistics*. Nantes, Francia, pp. 539-545.
- HORRIDGE, Matthew *et al* (2004): *A Practical Guide to Building Owl Ontologies Using the Protégé-Owl Plugin and Co-Ode Tools*. Manchester: University of Manchester.

- IBEKWE-SAN JUAN, Fidelia (2004): "Workshop on Terminology, Ontology and Knowledge Representation". *Terminology* 10.2: pp. 333-335.
- IRIS, Madelyn A., Bonnie E. LITOWITZ y Martha W. EVENS (1988): "Problems of the Part-Whole Relation". En: EVENS, Martha W. (ed.) (1988): *Relational Models of the Lexicon. Representing Knowledge in Semantic Networks*. Cambridge: Cambridge University Press, pp. 261-288.
- ISO (2000a): ISO 704. *Travail terminologique: Principes et méthodes*. Ginebra: ISO, 41 pp.
- (2000b): ISO 1087-1. *Terminology Work: Vocabulary. Part 1: Theory and Application*. Ginebra: ISO, 41 pp.
- JACKENDOFF, Ray S. (1983): *Semantics and Cognition*. Cambridge, Mass.: MIT.
- KACTUS (1996): *The Kactus Booklet Version 1.0.* Disponible en: <http://hcs.science.uva.nl/projects/NewKACTUS/Reports.html>. [Consulta: julio de 2006].
- KAGEURA, Kyo (1997): "Multifaceted/Multidimensional Concept Systems". En: WRIGHT, Sue Ellen y Gerhard BUDIN (eds.) (1997): *Handbook of Terminology Management*. Amsterdam: John Benjamins, pp. 119-132.
- (1998/1999): "Theories 'of' Terminology: A Quest for a Framework for the Study of Term Formation". *Terminology* 5.1: pp. 21-40.
- (2002): *The Dynamics of Terminology: A Descriptive Theory of Term Formation and Terminological Growth*. Amsterdam: John Benjamins.
- KERREMANS, Koen y Rita TEMMERMAN (2004): "Towards Multilingual, Termontological Support in Ontology Engineering". En IBEKWE-SAN JUAN, Fidelia (2004): *Workshop on Terminology, Ontology and Knowledge Representation*. Lyon: Université de Lyon.
- KERREMANS, Koen (2004): "Categorisation Frameworks in Termontography". *Linguistica Antverpiensia*, NS, n° 3: pp. 263-277.
- KERREMANS, Koen, Isabelle DESMEYTERE y Rita TEMMERMAN (2005): "Application-Oriented Terminography in Financial Forensics". *Terminology* 11.1: pp. 83-106.
- KHOO, Christopher, Syin CHAN y Yun NIU (2002): "The Many Facets of the Cause-Effect Relation". En: GREEN, Rebecca, Carol A. BEAN y Sungo Hyon MYAENG

- (eds.) (2002): *The Semantics of Relationships: An Interdisciplinary Approach*. Dordrecht: Kluwer Academic Publishers, pp. 51-70.
- KNOWLEDGE SYSTEMS LABORATORY (1997) [en línea]: *A Glossary of Ontology Terminology*. 1997. Disponible en: <http://www-ksl-svc.stanford.edu:5915/doc/frame-editor/glossary-of-terms.html>, [Consulta: junio de 2006].
- KOCOUREK, Rotislav (1991): *La langue française de la science et de la technique*. Wiesbaden: Brandstetter.
- L'HOMME, Marie Claude (2004): "A Lexico-Semantic Approach to the Structuring of Terminology". *COLING 2004*. Ginebra, pp. 7-14.
- L'HOMME, Marie Claude (2007): "Using Explanatory and Combinatorial Lexicology to Describe Terms". En WANNER, L (2007): *Selected Lexical and Grammatical Topics in the Meaning-Text Theory. In Honour of Igor Mel'cuk*. Amsterdam/Philadelphia: John Benjamins, pp. 11-50.
- LAKOFF, George (1987): *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind*. Chicago: University of Chicago Press.
- LAKOFF, George y M. JOHNSON (2001): *Metáforas de la vida cotidiana*. Madrid: Cátedra.
- LAMARCA LAPUENTE, María Jesús (2007) [en línea]: *Hipertexto: el nuevo concepto de documento en la cultura de la imagen*. Disponible en: <http://www.hipertexto.info/index.htm>. [Consulta: junio de 2007].
- LAURÉN, C., J. MYKING y Heribert PICHT (eds.) (1998): *Terminologie unter der Lupe*. Vol. 9. Viena: TermNet.
- LENAT, D.B. y R.V. GUHA (1990): *Building Large Knowledge-Based Systems: Representation and Inference in the Cyc Project*. Boston, MA: Addison-Wesley.
- LORENTE, Mercé (2005): "Ontología sobre economía y recuperación de la información". *Hipertext.net*, nº3.
- LYONS, John (1977): *Semantics. Vol. I y II*. Cambridge: Cambridge University Press.
- LYONS, John (1984): *Introducción al lenguaje y a la lingüística*. Barcelona: Teide.
- MADSEN, Bodil Nistrup, Bolette Sandford PEDERSEN y Hanne Erdman THOMSEN (2001): "Defining Semantic Relations for Ontoquery". En: JENSEN, Anker y Peter SKADHAUGE (eds.) (2001): *First International OntoQuery Workshop*. Holding: University of Southern Denmark, 57-88.

- MADSEN, Bodil Nistrup, Hanne Erdman THOMSEN y Carl VIKNER (2002a): "The CAOS Project - Computer Aided Ontology Structuring". En GALIA, Angelova, Dan CORBETT y Uta PRISS (eds.) (2002): *Foundations and Applications of Conceptual Structures - Contributions to ICCS 2002*. Bulgarian Academy of Sciences, Bulgaria, pp. 29-33.
- MADSEN, Bodil Nistrup, Hanne Erdman THOMSEN y Carl VIKNER (2002b): "Computer Assisted Ontology Structuring". En MELBY, Alan (ed.) (2002): *Proceedings of TKE '02 - Terminology and Knowledge Engineering*. Frankfurt: INRIA.
- (2004): "Comparison of Principles Applying to Domain Specific Versus General Ontologies". En OLTRAMARI, Alessandro *et al* (2004): *OntoLex 2004: Ontologies and Lexical Resources in Distributed Environments*. ELRA, pp. 90-95.
- MADSEN, Bodil Nistrup, Hanne Erdman THOMSEN y Carl VIKNER (2004): "Principles of a System for Terminological Concept Modelling". En: LINO, M^a Teresa *et al* (2004): *Proceedings of the Fourth International Conference on Language Resources and Evaluation*. ELRA, pp. 15-19.
- MADSEN, Bodil Nistrup, Hanne Erdman THOMSEN y Carl VIKNER (2005): "Multidimensionality in Terminological Concept Modelling". En: MADSEN, Bodil Nistrup y Hanne Erdman THOMSEN (2005): *Terminology and Content Development*. Copenhagen: Litera, pp. 161-74.
- MAHESH, Kavi y Sergei NIRENBURG (1995): "A Situated Ontology for Practical NLP". *Proceedings on Basic Ontological Issues in Knowledge Sharing. International Joint Conference on Artificial Intelligence (UCAI-1995)*.
- MAHESH, Kavi y Sergei NIRENBURG (1995): "Semantic Classification for Practical Natural Language Processing". *Sixth ASIS SIG/CR Classification Research Workshop: An Interdisciplinary Meeting. Chicago, IL*.
- MALAISE, Véronique, Pierre ZWEIGENBAUM y Bruno BACHIMONT (2004): "Repérage et exploitation d'énoncés définitoires. Un corpus pour l'aide à la construction d'ontologies". *TALN 2004*. Fez, Marruecos.
- (2004): "Extraction d'informations sémantiques pour l'aide à la construction d'ontologies différentielles". En: IBEKWE-SAN JUAN, Fidelia (2004): *Journées d'étude Terminologie, ontologie et représentation des connaissances*. Lyon.
- (2005): "Mining Defining Contexts to Help Structuring Differential Ontologies". *Terminology 11.1*: pp. 21-53.

- MARKOWITZ, Judith (1988): "An Exploration into Graded Set Membership". En: EVENS, Martha W. (ed.) (1988): *Relational Models of the Lexicon. Representing Knowledge in Semantic Networks*. Cambridge: Cambridge University Press, pp. 239-260.
- MARSHMAN, Elizabeth, Tricia MORGAN e Ingrid MEYER (2002): "French Patterns for Expressing Concept Relations". *Terminology* 8.1: pp. 1-29.
- MEL'CUK, Igor y Alexander ZHOLKOVSKY (1988): "The Explanatory Combinatorial Dictionary". En: EVENS, Martha W. (ed.) (1988): *Relational Models of the Lexicon. Representing Knowledge in Semantic Networks*. Cambridge: Cambridge University Press, pp. 41-74.
- MEL'CUK, Igor (1997): *Vers une linguistique sens-texte. Leçon inaugurale*. (Lección impartida el viernes, 10 de enero de 1997), Collège de France, Chaire Internationale. Disponible en: <http://www.olst.umontreal.ca/pdf/melcukColldeFr.pdf>. [Consulta: julio de 2006].
- MEYER, Ingrid *et al* (1992): "Towards a New Generation of Terminological Resources: An Experiment in Building a Terminological Knowledge Base". *COLING 92*. Nantes: pp. 956-60.
- MEYER, Ingrid y Kristen MACKINTOSH (1996): "Refining the Terminographer's Concept-Analysis Methods: How Can Phraseology Help?". *Terminology* 3.1: pp. 1-26.
- MEYER, Ingrid, Victoria ZALUSKI y Kristen MACKINTOSH (1997): "Metaphorical Internet Terms". *Terminology* 4.1: pp. 1-33.
- MEYER, Ingrid, Karen ECK y Douglas SKUCE (1997): "Systematic Concept Analysis within a Knowledge-Based Approach to Terminology". En: WRIGHT, Sue Ellen y Gerhard BUDIN (eds.) (1997): *Handbook of Terminology Management*. Amsterdam: John Benjamins, pp. 98-118.
- MEYER, Ingrid y Kristen MACKINTOSH (2000): "When Terms Move into Our Everyday Lives; an Overview of De-Terminologization". *Terminology* 6.1: pp. 111-38.
- MEYER, Ingrid (2001): "Extracting Knowledge-Rich Contexts for Terminography". En: BOURIGAULT, Didier, Christian JACQUEMIN y Marie Claude L'HOMME (eds.) (2001): *Recent Advances in Computational Terminology. Natural Language Processing*. Amsterdam/Philadelphia: John Benjamins, pp. 279-302.

- MHAANNA, Ekaterina (2005): "Comparing the Concept Relations and Their Grouping in UMLS and in Ontoquery". En: MADSEN, Bodil Nistrup y Hanne ERDMAN (2005): *Copenhagen 7th International Conference on Terminology and Knowledge Engineering, TKE 2005*. Copenhagen: Association for Terminology and Knowledge Transfer, pp. 139-50.
- MILLER, George A. (1998): "Nouns in Wordnet". En: FELLBAUM, Christiane (ed.) (1998): *Wordnet: An Electronic Lexical Database. Language, Speech and Communication*. Cambridge, Massachusetts: MIT Press, pp. 23-46.
- MILLER, Katherine J. (1998): "Modifiers in Wordnet". En: FELLBAUM, Christiane (ed.) (1998): *Wordnet: An Electronic Lexical Database. Language, Speech and Communication*. Cambridge, Massachusetts: MIT Press, pp. 47-67.
- MONTERDE REY, Ana María (2003): "Creación de un programa informático generador de sistemas de conceptos como apoyo a la enseñanza de la terminología". En: GALLARDO SAN SALVADOR, Natividad (ed.) (2003): *Terminología y traducción: un bosquejo de su evolución*. Granada: Atrio, pp. 91-106.
- MONTERO MARTÍNEZ, Silvia (2002): *Estructuración conceptual y formalización terminológica de frasemas en el subdominio de la oncología*. Universidad de Valladolid, tesis doctoral inédita.
- MORENO ORTIZ, Antonio (2002): "Representación de la información terminológica en ONTOTERM®: un sistema gestor de bases de datos terminológicas basado en el conocimiento". En: FABER, Pamela y Catalina JIMÉNEZ (2002): *Investigar en terminología*. Granada: Comares, pp. 25-70.
- MURPHY, M. Lynne (2003): *Semantic Relations and the Lexicon*. Cambridge: Cambridge University Press.
- NECHES, Robert *et al* (1991): "Enabling Technology for Knowledge Sharing" *AI Magazine* 13.3: pp. 36-56.
- NIRENBURG, Sergei y Victor RASKIN (2004): *Ontological Semantics*. Cambridge, Mass.: MIT.
- NKWENTI AZEH, B (1994): "Positional and Combinational Characteristics of Terms: Consequences for Corpus-Based Terminography" *Terminology* 1.1: pp. 61-95.
- NOY, Natalya F. y Deborah L. MCGUINNESS (2001): *Ontology Development 101: A Guide to Creating Your First Ontology*. Stanford Knowledge Systems Laboratory Technical Report.

- NOY, Natalya F. (2005) [en línea]: *Creating Synonyms*. Disponible en: <http://protege.cim3.net/cgi-bin/wiki.pl?CreatingSynonyms>. [Consulta: julio de 2006].
- NUOPPONEN, Anita (1994): "On Causality and Concept Relationships". En: DRASKAU, J.K. y Heribert PICT (1994): *Terminology Science and Terminology Planning, IITF-Workshop on Theoretical Issues of Terminology Science*. Vienna: TermNet, 217-230.
- (1994): *Begreppssystem För Terminologisk Analysis (Concept Systems for Terminological Analysis)*. Universidad de Vaasa, tesis doctoral inédita.
- (2005): "Concept Relations: An Update of a Concept Relation Classification". En: MADSEN, Bodil Nistrup y Hanne ERDMAN (2005): *Copenhage 7th International Conference on Terminology and Knowledge Engineering, TKE 2005*. Copenhage: Association for Terminology and Knowledge Transfer, pp. 127-38.
- ONCOTERM (2002): "ONCOTERM: sistema bilingüe de información y recursos oncológicos". En: GUERRERO RAMOS, Gloria y Manuel Fernando PÉREZ LAGOS (2002): *Panorama actual de la terminología*. Interlengua. Granada: Comares, pp. 221-237.
- OSTER, Ulrike (2003): *Los términos de la cerámica en alemán y en español. Análisis semántico orientado a la traducción de los compuestos nominales alemanes*. Universitat Jaume I, tesis doctoral inédita.
- (2004): "From Relational Schemas to Subject-Specific Semantic Relations. A Two-Step Classification of Compound Terms". *Annual Review of Cognitive Linguistics* n° 2: 235-59.
- (2005): *Las relaciones semánticas de términos polilexemáticos: estudio contrastivo alemán-español*. Frankfurt am Main: Peter Lang.
- (2006): "Classifying Domain-Specific Intraterm Relations: A Schema-Based Approach". *Terminology* 12.1: pp. 1-17.
- OTMAN, Gabriel (1996): *Les représentations sémantiques en terminologie*. Paris: Masson.
- (1997) : "Les bases de connaissances terminologiques: les banques de terminologie de seconde génération". *Meta* XLII, n° 2: pp. 244-256.
- PASLARU BONTAS, Elena y Malgorzata MOCHOL (2005): "Towards a Reuse-Oriented Methodology for Ontology Engineering". En: MADSEN, Bodil Nistrup y Hanne ERDMAN (2005): *Copenhage 7th International Conference on Terminology and*

Knowledge Engineering, TKE 2005. Copenhagen: Association for Terminology and Knowledge Transfer, pp. 175-87.

PAVEL, Silvia y Diane NOLET (2001): *Handbook of Terminology*. Gatineau: Minister of Public Works and Government Services.

PEARSON, Jennifer (1998): *Terms in Context*. Amsterdam/Filadelfia: John Benjamins.

PÉREZ, Chantal (2002): "Terminología basada en corpus: principios teóricos y metodológicos". En: FABER, Pamela y Catalina JIMÉNEZ (2002): *Investigar en terminología*. Granada: Comares, pp. 127-166.

--- (2002) [en línea]: *Explotación de los corpórea textuales informatizados para la creación de bases de datos terminológicas basadas en el conocimiento*. Estudios de Lingüística Española (ELiEs). Disponible en: <http://elies.rediris.es/elies182002>, [Consulta: julio de 2006].

POLGUERE, Alain (1988): "La théorie sens-texte". *Dialangue* 8-9: pp. 9-30.

PRIBBENOW, Simone (2002): "Meronymic Relationships: From Classical Mereology to Complex Part-Whole Relations". En: GREEN, Rebecca, Carol A. BEAN y Sungo Hyon MYAENG (eds.) (2002): *The Semantics of Relationships: An Interdisciplinary Approach*. Dordrecht: Kluwer Academic Publishers, pp. 35-50.

ROGERS, Margaret (2004): "Multidimensionality in Concepts Systems: A Bilingual Textual Perspective". *Terminology* 10.2: pp. 215-40.

ROSCH, Eleanor (1978): *Cognition and Categorization*. Hillsdale: Lawrence Erlbaum Associates.

RUPPENHOFER, Josef *et al* (2005) [en línea]: *Framenet: Theory and Practice*. Disponible en:

http://framenet.icsi.berkeley.edu/index.php?option=com_wrapper&Itemid=126.

[Consulta: junio de 2006].

SAGER, Juan Carlos, David DUNGWORTH y Peter F. McDONALD (1980): *English Special Languages: Principles and Practice in Science and Technology*. Wiesbaden: Oscar Brandstetter.

SAGER, Juan Carlos (1993): *Curso práctico sobre el procesamiento de la terminología*. Madrid: Fundación Germán Sánchez Ruipérez.

--- (1994): *Language Engineering and Translation. Consequences of Automation*. Amsterdam/Philadelphia: John Benjamins.

SAGER, Juan Carlos y Kyo KAGEURA (1994): "Concept Classes and Conceptual Structures: Their Role and Necessity in Terminology". *ALFA* 7/8: pp. 191-216.

- SAGER, Juan Carlos (1998/1999): "In Search of a Foundation: Towards a Theory of the Term" *Terminology* 5.1: pp. 41-57.
- (2000): *Essays on Definition*. Amsterdam/Filadelphia: John Benjamins.
- (2000): "Pour une approche fonctionnelle de la terminologie". En: BEJOINT, Henri y Philippe THOIRON (eds.) (2000): *Le sens en terminologie*. Lyon: Presses universitaires de Lyon, pp. 40-60.
- SÁNCHEZ-GIJÓN, Pilar (2004): "La extracción de conocimiento y terminología a partir de corpus *ad hoc*: el uso de documentos digitales de la web pública". *Linguistica Antverpiensia*, NS, nº 3: pp. 179-203.
- SEGUELA, Patrick (1998): "Adaptation semi-automatique d'une base de marqueurs de relations sémantiques sur des corpus spécialisés". *Terminologies Nouvelles*, nº19: pp. 52-60.
- SIERRA, Gerardo y John MCNAUGHT (2000): "Design of an Onomasiological Search System: A Concept-Oriented Tool for Terminology". *Terminology* 6.1: pp. 1-34.
- SOLER, Victoria (2005): "Patrones lingüísticos para la búsqueda de información conceptual en el corpus textual especializado de la cerámica TXTCERAM".
- SOLER, Victoria, Amparo ALCINA y Anna ESTELLÉS (2006): "La digitalización de textos para la elaboración de un corpus lingüístico electrónico: una experiencia de trabajo en equipo con estudiantes". *X Jornadas de Traducción e Interpretación: tecnologías asequibles*. Vic: Universitat de Vic.
- SOWA, John F. (1984): *Conceptual Structures: Information Processing in Mind and Machine*. Reading, Mass.: Addison-Wesley.
- SOWA, John F. (2000): *Knowledge Representation. Logical, Philosophical, and Computational Foundations*. Pacific Grove et al: Brooks/Cole.
- SPYNS, Peter y Jan DE BO (2004): "Ontologies: A Revamped Cross-Disciplinary Buzzword or a Truly Promising Interdisciplinary Research Topic?". *Linguistica Antverpiensia*, NS, nº 3: pp. 279-292.
- STAAB, Steffen *et al* (2001): "Knowledge Processes and Ontologies". *IEEE Intelligent Systems* 16.1: 26-34.
- STAAB, Steffen y Rudi STUDER (2004): *Handbook on Ontologies*. Berlín: Springer.
- STANFORD MEDICAL INFORMATICS (2000) [en línea]: *Protégé User's Guide*. Disponible en: <http://protege.stanford.edu>. [Consulta: junio de 2006].

- STUDER, Rudi, V. Richard BENJAMINS y Dieter FENSEL (1998): "Knowledge Engineering: Principles and Methods". *IEEE Transactions on Data and Knowledge Engineering*, 25.1-2: pp. 161-97.
- SUBIRATS, Carlos y Miriam R. L. PETRUCK (2003): "Surprise: Spanish Framenet!". En: HAJICOVA, Eva, Anna KOTESOVCOVA y Jiri MIROVSKY. *Proceedings of CIL 17*. Prague: Matfyzpress.
- SWARTOUT, William R. *et al* (1997): "Toward Distributed Use of Large-Scale Ontologies". En: FARQUHAR, Adam *et al* (1997): *AAAI '97 Spring Symposium on Ontological Engineering*. California: Stanford University, pp. 138-48.
- TEBÉ, Carles (2002): "Los conceptos revisitados: una perspectiva cognitiva". En: GUERRERO RAMOS, Gloria y Manuel Fernando PÉREZ LAGOS (2002): *Panorama actual de la terminología*. Interlengua. Granada: Comares, pp. 359-71.
- TEMMERMAN, Rita (2000): *Towards New Ways of Terminology Description: The Sociocognitive Approach*. Amsterdam: John Benjamins.
- (2000): "Une théorie réaliste de la terminologie: le sociocognitivisme". *Terminologies Nouvelles*, n°21: pp. 58-64.
- TEMMERMAN, Rita y Koen KERREMANS (2003): "Termonography: Ontology Building and the Sociocognitive Approach to Terminology Description". En: HAJICOVA, Eva, Anna KOTESOVCOVA y Jiri MIROVSKY. *Proceedings of CIL 17*. Prague: Matfyzpress.
- TEMMERMAN, Rita (2003): "Innovative Methods in Specialised Lexicography". *Terminology* 9.1: pp. 117-35.
- TERMCAT (1990): *Metodologia del treball terminològic*. Barcelona: Generalitat de Catalunya.
- TIMPF, Sabine (2002): *The Need for Task Ontologies in Interoperable GIS*. Zürich: University of Zürich, Department of Geology.
- TOFT, Bertha (1994): "Conceptual Relations in Terminology and Knowledge Engineering". En: DRASKAU, Jennifer K. y Heribert PICT (eds.) (1994): *International Terminological Conference Terminology Science and Terminology Planning in Commemoration of E. Drezén (1892-1992)*. Viena: TermNet, pp. 259-282.
- (2002): "Systems of Concepts at a Crossroads?" *Terminology Science and Research IITF Journal*, 13, n°1-2: pp. 77-92.
- UNGERER, Friedrich (1997): *An Introduction to Cognitive Linguistics*. Harlow: Longman.

- USCHOLD, M y M KING (1995): "Towards a Methodology for Building Ontologies". En: SKUCE, Douglas (ed.) (1995): *IJCAI'95. Workshop on Basic Ontological Issues in Knowledge Sharing*. Montréal. 6.1-6.10.
- VAN CAMPENHOUDT, Marc (1996): "Recherche d'équivalences et structuration des réseaux notionnels: le cas des relations méronymiques". *Terminology* 3.1: pp. 53-83.
- VOSSEN, Piek (1999): *EuroWordNet: General Documentation*. Amsterdam: University of Amsterdam.
- WEISSENHOFER, Peter (1995): *Conceptology in Terminology Theory, Semantics and Word-Formation*. Viena: TermNet.
- WILKS, Yorick, Brian SLATOR y Louise GUTHRIE (1996): *Electric Words: Dictionaries, Computers and Meaning*. Cambridge, Mass.: MIT Press.
- WINSTON, Morton E. , Roger CHAFFIN y Douglas J. HERRMANN (1987): "A Taxonomy of Part-Whole Relations". *Cognitive Science* 11: pp. 417-444.
- WRIGHT, Sue Ellen y Gerhard BUDIN (eds.) (1997): *Handbook of Terminology Management*. Vol. 1. Amsterdam: John Benjamins.
- WRIGHT, Sue Ellen (1997): "Representation of Concept Systems". En: WRIGHT, Sue Ellen y Gerhard BUDIN (eds.) (1997): *Handbook of Terminology Management*. 2 vols. Amsterdam: John Benjamins. Vol. 1. Amsterdam: John Benjamins, pp. 89-97.
- WRIGHT, Sue Ellen y Gerhard BUDIN (eds.) (2001): *Handbook of Terminology Management*. Vol. 2. Amsterdam: John Benjamins.
- WÜSTER, Eugen (1996a): "Classificacions de conceptes i temes: diferències en la seva essència i en la seva aplicació". En: CABRÉ, María Teresa (1996): *Terminologia: Selecció de textos d'Eugen Wüster*. Barcelona: Universitat de Barcelona, pp. 113-151.
- (1996b): "La inversió d'una relació conceptual i la seva representació en els diccionaris". En: CABRÉ, María Teresa (1996): *Terminologia: Selecció de textos d'Eugen Wüster*. Barcelona: Universitat de Barcelona, pp. 205-223.
- (1998): *Introducción a la Teoría General de la Terminología y a la lexicografía terminológica*. Barcelona: IULA.
- ZAWADA, Britta y Piet SWANEPOEL (1994): "On the Empirical Adequacy of Terminological Concept Theories: The Case for Prototype Theory". *Terminology* 1.2: pp. 253-275.

ZINGLE, Henri (1996): "Outils et méthode d'extraction automatique de la terminologie".
Réflexions méthodologiques sur le travail en terminologie et en terminotique dans les langues latines. Niza: Realiter.

ANEXOS

Anexos

Anexo 1: Referencias bibliográficas del corpus TXTCERAM y de los recursos textuales sobre la industria cerámica

ALARCÓN, J., J. CARDA *et al* (1992): *Nuevos productos y tecnologías de esmaltes y pigmentos cerámicos. Su fabricación y utilización*. Castellón: Faenza Editrice Ibérica.

AMORÓS ALBERO, J. L., V. BARBA *et al* (1994): *Estructuras cristalinas de los silicatos y óxidos de las materias primas cerámicas*. Castellón: Instituto de Tecnología Cerámica-Asociación de Investigación de las Industrias Cerámicas (AICE-ITC).

AMORÓS ALBERO, J. L., V. BELTRÁN PORCAR *et al* (1991): *Defectos de fabricación de pavimentos y revestimientos*. Castellón: Asociación de investigación de las industrias cerámicas.

ASCER *et al* (2003): *Guía de la baldosa cerámica*. Valencia: Instituto Valenciano de la Edificación.

ATC (1990): *Tecnología de la fabricación de azulejos*. Asociación de Técnicos Cerámicos (ATC).

BARBA, A., V. BELTRÁN *et al* (1997): *Materias primas para la fabricación de soportes de baldosas cerámicas*. Castellón: Instituto de Tecnología Cerámica (ITC).

CLAUSELL SEBASTIÀ, A. y R. GALINDO RENAU (1991): *Apuntes de operaciones básicas en la fabricación de pavimentos y revestimientos cerámicos*. Castellón: Conselleria de Cultura, Educació i Ciència, Generalitat Valenciana.

ENRIQUE NAVARRO, J. E., F. NEGRE *et al* (1980): *Mayólica, gres y porcelana artística*. Castellón: Asociación de investigación de las industrias cerámicas (AICE).

ESCARDINO BENLLOCH, A. y M. GONZÁLEZ CUDILLEIRO (1991): *Azulejos y pavimentos cerámicos españoles*. Madrid: Ministerio de Industria, Comercio y Turismo.

ESCRIBANO LÓPEZ, P., J. B. CARDA CASTELLÓ *et al* (2001): *Esmaltes y pigmentos cerámicos*. Castellón: Faenza Editrice Iberica.

- ESTEBAN CHAPAPRÍA, V. (1998): *Naútica de recreo y turismo en el Mediterráneo: La Comunidad Valenciana*. Madrid: Síntesis.
- FERNÁNDEZ NAVARRO, J. M. (1991): *El vidrio*. Madrid: Consejo Superior de Investigaciones Científicas.
- GALINDO RENAU, R. (1994): *Pastas y vidriados*. Castellón: Faenza Editrice Iberica.
- GARCÍA VERDUCH, A. (1993): *Colocación de pavimentos y revestimientos cerámicos*. Castellón: Instituto de Tecnología Cerámica.
- GIOVANNINI, R. (1989): *La serigrafía en la cerámica*. Barcelona: Ediciones Omega.
- GUILLEM MONZONÍS, C. y M. C. GUILLEM VILLAR (1987): *Diccionario cerámico científico-práctico (español-inglés-alemán-francés)*. Castellón: Sociedad Española de Cerámica y Vidrio.
- IBERICA, F. E. (2001): *Tecnología cerámica en España. Ceramic Technology in Spain*. Castellón: Faenza Editrice Iberica.
- LUCAS, F. (2004): “Tensión superficial en suspensión de esmaltes”. *Frittainnova*. 1: 4-10.
- MCCOLM, I. J. (1994): *Dictionary of ceramic science and engineering*. 2ª ed. Nueva York: Plenum Press.
- MATTHES, W. E. (1990): *Vidriados cerámicos. Fundamentos, propiedades, recetas, métodos*. Barcelona: Omega.
- MONRÓS, G., J. A. BADENES *et al* (2003): *El color de la cerámica: Nuevos mecanismos en pigmentos para los nuevos procesados de la industria cerámica*. Castellón: Publicacions de la Universitat Jaume I. Servei de Comunicació i Publicació.
- NEBOT-DÍAZ, I., M. MARCHAL *et al* (2000): *Nuevas tecnologías para el sector cerámico de Castellón. Desarrollo de esmaltes vitrocrystalinos y vitrocerámicos*. Castellón: Publicacions de la Universitat Jaume I.
- NORTON, F. H. (1988): *Cerámica fina. Tecnología y aplicaciones*. Barcelona: Ediciones Omega.
- PADOA, L. (1990): *La cocción de productos cerámicos. Con especial referencia a los materiales de revestimiento y de pavimento*. Barcelona: Ediciones Omega.

- PORCAR, J. L., A. PLEGUEZUELO *et al* (1987): *Manual-guía técnica de los revestimientos y pavimentos cerámicos*. Castellón: Instituto de Tecnología Cerámica.
- QUALICER (1998): *Conferencias generales y ponencias. V Congreso mundial de la calidad del azulejo y del pavimento cerámico*, Castellón.
- QUALICER (2000): *Conferencias Generales. Ponencias. Posters. Mesa Redonda. VI Congreso Mundial de la Calidad del Azulejo y del Pavimento Cerámico*, Castellón.
- QUALICER (2002): *Conferencias generales. Ponencias. Posters. Mesa redonda. VII Congreso mundial de la calidad del azulejo y del pavimento cerámico*, Castellón.
- QUALICER (2004): *Conferencias generales. Ponencias. Posters. Mesa debate. VIII Congreso mundial de la calidad del azulejo y del pavimento cerámico*, Castellón.
- QUIMICER (2004): *Quimicer Spherical*, Castellón.
- RADO, P. (1990): *Introducción a la tecnología de la cerámica*. Barcelona: Ediciones Omega.
- RHODES, D. (1999): *Hornos para ceramistas*. Barcelona: CEAC.
- SACMI (2004): *Tecnología cerámica aplicada*. Castellón: Faenza Editrice Iberica.
- SANCHO, A. (1998): *Introducción al Turismo*. Madrid: Organización Mundial del Turismo.
- SUCH CLIMENT, M. P. (1996): *Turismo y medio ambiente en el litoral alicantino*. Alicante: Insitut de Cultura Juan Gil-Albert.
- TORRELLA ALCARAZ, E. (1996): *La producción de frío*. Valencia: Universidad Politécnica de Valencia, Servicio de Publicaciones.

Anexo 2: Referencias sobre las herramientas informáticas empleadas

SCOTT, M. [en línea]: *WordSmith Tools*. <http://lexically.net/wordsmith>.

[Consulta: julio de 2006].

STANFORD MEDICAL INFORMATICS. [en línea]: *The Protégé Ontology Editor and Knowledge Acquisiton System*. <http://protege.stanford.edu/>. [Consulta:

julio de 2006].

TRADOS IRELAND, LTD. (2001): *Ayuda en pantalla de Multiterm*. Trados: Dublín.