

UNIVERSIDAD DE MÁLAGA

FACULTAD DE FILOSOFÍA Y LETRAS

DEPARTAMENTO DE TRADUCCIÓN E INTERPRETACIÓN



TESIS DOCTORAL

El trabajo ontoterminográfico aplicado a la traducción de textos del turismo de aventura (español-inglés-alemán): fases de preparación, elaboración y edición

ISABEL DURÁN MUÑOZ

DIRIGIDA POR LA DRA. D.^a GLORIA CORPAS PASTOR

MÁLAGA, 2011



AUTOR: Isabel Durán Muñoz

Edita: Servicio de Publicaciones de la Universidad de Málaga

ISBN: 978-84-9747-641-6



Esta obra está sujeta a una licencia Creative Commons:

Reconocimiento - No comercial - SinObraDerivada (cc-by-nc-nd):

[Http://creativecommons.org/licenses/by-nc-nd/3.0/es](http://creativecommons.org/licenses/by-nc-nd/3.0/es)

Cualquier parte de esta obra se puede reproducir sin autorización
pero con el reconocimiento y atribución de los autores.

No se puede hacer uso comercial de la obra y no se puede alterar, transformar o
hacer obras derivadas.



Departamento de Traducción e Interpretación
Universidad de Málaga

Gloria Corpas Pastor, catedrática del Departamento de Traducción e Interpretación de la Universidad de Málaga, certifica que la tesis de D.^a Isabel Durán Muñoz, titulada *El trabajo ontoterminográfico aplicado a la traducción de textos del turismo de aventura (español-inglés-alemán): fases de preparación, elaboración y edición*, ha sido realizada bajo su dirección y cumple todos los requisitos legales para proceder a su defensa.

Málaga, a 13 de abril de 2011

Fdo. Dra. D.^a Gloria Corpas Pastor

A mi madre

*La única posibilidad de descubrir los límites de lo posible
es aventurarse un poco más allá de ellos,
hacia lo imposible.*

Arthur C. Clarke

AGRADECIMIENTOS

En primer lugar, quisiera mostrar todo mi agradecimiento a la Dra. D. ^a Gloria Corpas Pastor, directora de esta tesis doctoral, por todos sus sabios consejos, sus palabras de ánimo y por la confianza que ha depositado en mí desde el primer día. Durante estos cuatro años he tenido la gran suerte de trabajar a su lado y de recibir su apoyo en todo momento, lo que me ha permitido aprender y crecer como persona y como investigadora.

Quisiera hacer extensible este agradecimiento a mis tutores durante mis dos estancias de investigación: el Dr. D. Ruslan Mitkov, director del Research Group in Computational Linguistics, de la Universidad de Wolverhampton, y la Dra. D. ^a Rita Temmerman, directora del Centrum voor Vaktaal en Communicatie (CVC), de la Erasmushogeschool de Bruselas, por su maravillosa acogida así como por sus consejos y apoyo constante tanto en persona como desde la distancia. Asimismo, me gustaría agradecer al equipo de ambos centros de investigación su generosa acogida y apoyo durante el tiempo de mi estancia y, en especial al Dr. D. Le An Ha por haber trabajado incansablemente en la herramienta *ProTermino*, así como por su disponibilidad a ejercer como evaluador externo para esta tesis doctoral. De la misma manera, agradezco a Catelina Hallett por actuar también como evaluadora externa.

Me gustaría continuar los agradecimientos con D. Javier Alejandro Fernández Sola por haberme ayudado a hacer realidad la herramienta *Trandix* y, en especial, al Dr. D. Francisco Javier Durán Muñoz por haberme escuchado, entendido y allanado el camino para conseguir diseñar e implementar *Trandix*. Sin ellos esto se habría quedado en un sueño.

Le agradezco también el apoyo ofrecido por parte de todos los miembros del grupo de investigación HUM-106, en particular a los profesores Dra. D. ^a Miriam Seghiri Domínguez, Dr. D. Jorge J. Leiva Rojo y D. ^a M. ^a del Carmen Amaya Galván, y sobre todo de mis compañeras Dra. D. ^a Cristina M. Castillo Rodríguez, D. ^a María Muñoz Ramos y D. ^a Elena García Izquierdo, por apoyarme en los momentos más difíciles. Asimismo, agradezco de verdad a las profesoras del Departamento de Traducción e Interpretación de la Universidad de Málaga que me han permitido realizar los experimentos para el estudio empírico en sus clases, de nuevo la Dra. D. ^a Gloria Corpas Pastor y la Dra. D. ^a Miriam Seghiri Domínguez, que se ofrecieron a ayudarme de forma totalmente desinteresada.

Asimismo, no puedo olvidar a los profesionales y expertos en educación física del deporte, en turismo, en geografía y en traducción de diferentes puntos de España y del extranjero que me han recibido personalmente o a distancia y que me han aportado conocimientos y opiniones de forma absolutamente altruista que me han permitido llevar a cabo la parte de la investigación relacionada con los segmentos turísticos y las actividades del turismo de aventura.

Finalmente, en un terreno más personal, comienzo agradeciendo a mis amigos, que ajenos a esta investigación y a todo lo que esto significaba han estado siempre preocupados en persona y en la distancia por este trabajo.

A toda mi familia, que cada sábado me animaba y me apoyaba de forma incansable y me hacía olvidar las páginas en blanco que me esperaban a la vuelta. Y en especial a mi madre, que me ha enseñado que no hay nada imposible en la vida. Todo lo que he conseguido hasta ahora y todo lo que consiga en el futuro se lo debo a ella.

Y a Pepe, por mantenerse a mi lado siempre fuera donde fuere y por compartir conmigo buenos y malos momentos, por indicarme dónde estaba la luz cuando solo veía oscuridad y por andar conmigo el camino hacia la reinención.

ÍNDICE

ABSTRACT

| | |
|---|----|
| 1. Introduction | 5 |
| 2. Aims and goals | 7 |
| 3. Materials | 10 |
| 3.1. <i>ProTermino</i> | 10 |
| 3.2. <i>Trandix</i> | 11 |
| 3.3. Multilingual corpus | 12 |
| 3.4. Template of Assessment | 12 |
| 3.5. Electronic survey to professional translators | 13 |
| 3.6. <i>Markin</i> | 13 |
| 3.7. <i>ReCor</i> | 14 |
| 3.8. Electronic survey | 14 |
| 4. Methods | 14 |
| 4.1. Research on terminology and terminography field | 14 |
| 4.2. Study of professional translators' needs and expectations | 15 |
| 4.3. Steps to elaborate a multilingual ontoterminography-based database | 16 |
| 4.4. Empirical experiments with semi-professional translators | 17 |
| 5. Results | 18 |
| 5.1. Results of the research on terminology and terminography field | 18 |
| 5.2. Results of the study of professional translators' needs and expectations | 19 |
| 5.3. Results of the steps to elaborate a multilingual ontoterminography-based database | 20 |
| 5.4. Results of the empirical experiments with semi-professional translators | 21 |
| 5.4.1. Execution of tasks | 21 |
| 5.4.2. Time of execution | 22 |
| 5.4.3. Quality criteria | 23 |
| 5.4.4. Users' satisfaction | 24 |
| 6. Conclusions and future work | 28 |
| 6.1. Conclusions | 28 |
| 6.2. Future work | 30 |

| | |
|--|-----|
| INTRODUCCIÓN | 35 |
| | |
| CAPÍTULO 1. LA TERMINOLOGÍA MODERNA Y LOS NUEVOS ENFOQUES DEL TRABAJO TERMINOGRÁFICO | |
| 1.1. Introducción | 51 |
| 1.2. La dicotomía entre lengua general y lenguaje de especialidad | 53 |
| 1.2.1. Lenguaje de especialidad | 55 |
| 1.2.2. Caracterización de los lenguajes de especialidad | 61 |
| 1.2.2.1. Plano lingüístico | 62 |
| 1.2.2.2. Plano pragmático | 63 |
| 1.2.2.3. Plano funcional | 67 |
| 1.3. La dicotomía terminología - lexicología | 68 |
| 1.3.1. El concepto <término> | 74 |
| 1.3.2. Corrientes teóricas de la terminología | 78 |
| 1.3.2.1. La corriente clásica de la terminología (TGT) | 80 |
| 1.3.2.2. Las corrientes modernas de la terminología | 84 |
| 1.3.2.2.1. La Socioterminología | 88 |
| 1.3.2.2.2. La Teoría Comunicativa de la Terminología (TCT) | 89 |
| 1.3.2.2.3. La Teoría Sociocognitiva de la Terminología (TST) | 92 |
| 1.3.2.3. Corrientes teóricas de la lingüística aplicables a la terminología | 97 |
| 1.3.2.3.1. La Lingüística de Corpus | 98 |
| 1.3.2.3.2. La Lingüística Computacional | 101 |
| 1.3.2.3.3. La Lingüística Cognitiva | 104 |
| 1.3.2.3.3.1. La Semántica de Marcos | 111 |
| 1.4. La dicotomía terminografía – lexicografía | 115 |
| 1.4.1. Aproximación de la terminografía a la lexicografía especializada | 119 |
| 1.4.2. La terminografía basada en corpus | 127 |
| 1.4.2.1. La Lexicografía basada en corpus | 132 |
| 1.4.2.2. Algunas consideraciones en torno al corpus: concepto y definición | 136 |
| 1.4.2.2.1. Criterios generales para la compilación de corpus | 140 |
| 1.4.2.2.2. Tipos de corpus | 144 |
| 1.4.2.3. Criterios específicos para la elaboración y diseño de los corpus especializados para la terminografía | 150 |
| 1.4.2.4. Herramientas de análisis de corpus: necesidades específicas de los terminógrafos | 156 |
| 1.4.2.5. Ventajas e inconvenientes del empleo de corpus en terminografía | 171 |
| 1.4.3. La terminografía basada en la representación de conocimiento | 175 |
| 1.4.3.1. La representación de conocimiento especializado en la TGT y otros modelos de representación | 179 |
| 1.4.3.2. Algunas consideraciones en torno a las ontologías: definición y componentes | 190 |
| 1.4.3.2.1. Criterios generales para la creación de ontologías | 197 |
| 1.4.3.2.2. Clasificaciones de ontologías | 199 |
| 1.4.3.3. Criterios específicos para el diseño de ontologías en terminografía | 202 |
| 1.4.3.4. Herramientas para la elaboración y gestión de ontologías | 206 |
| 1.4.3.5. Ventajas e inconvenientes del empleo de ontologías en terminografía | 213 |

| | |
|--|-----|
| 1.5. Proyectos basados en la representación del conocimiento especializado mediante corpus dirigidos a traductores | 217 |
| 1.5.1. Proyecto COGNITERM | 220 |
| 1.5.2. Proyecto ONCOTERM | 221 |
| 1.5.3. Proyecto GENOMA-KB | 224 |
| 1.5.4. Proyecto PROLEX | 225 |
| | |
| CAPÍTULO 2. EL TRABAJO ONTOTERMINOGRÁFICO BASADO EN CORPUS DIRIGIDO A TRADUCTORES | |
| 2.1. Introducción | 231 |
| 2.2. La traducción de textos especializados y la competencia traductora | 234 |
| 2.3. La competencia documental en traducción | 243 |
| 2.3.1. Las fuentes de información para traductores | 246 |
| 2.4. Las fuentes de información terminológicas | 261 |
| 2.4.1. Clasificación de los recursos terminológicos | 265 |
| 2.4.1.1. Según la presentación | 266 |
| 2.4.1.2. Según el número de lenguas | 272 |
| 2.4.1.3. Según el tipo de soporte | 273 |
| 2.4.1.4. Según la ordenación de sus contenidos | 275 |
| 2.4.1.5. Según la normalización | 277 |
| 2.4.1.6. Según el grupo de destinatarios | 279 |
| 2.4.1.7. Según el tipo de unidades terminológicas | 279 |
| 2.4.1.8. Según la extensión en la especialidad | 281 |
| 2.4.1.9. Según la fuente de información utilizada | 282 |
| 2.4.1.10. Según la finalidad..... | 284 |
| 2.4.2. Evaluación de fuentes de información terminológica | 285 |
| 2.4.2.1. Plantilla de evaluación. Parámetros e indicadores | 290 |
| 2.5. Expectativas y necesidades de consulta de traductores profesionales | 301 |
| 2.5.1. Estudios previos realizados | 304 |
| 2.5.2. Presentación de la encuesta y su elaboración | 310 |
| 2.5.3. Resultados de la encuesta y conclusiones | 315 |
| 2.5.4. Algunos trabajos terminográficos dirigidos a traductores | 327 |
| 2.6. Metodologías para el trabajo ontoterminográfico sistemático basado en corpus | 339 |
| 2.6.1. Propuesta de metodología para el trabajo ontoterminográfico sistemático basado en corpus dirigido a traductores | 349 |
| 2.6.1.1. Fase I: Delimitación del trabajo y análisis del dominio | 352 |
| 2.6.1.2. Fase II: Compilación del corpus especializado | 356 |
| 2.6.1.2.1. Diseño del corpus | 357 |
| 2.6.1.2.2. Búsqueda documental | 361 |
| 2.6.1.2.3. Organización de textos | 365 |
| 2.6.1.3. Fase III: Gestión ontoterminográfica | 366 |
| 2.6.1.3.1. Extracción de candidatos a término y validación | 367 |
| 2.6.1.3.2. Búsqueda de equivalentes de traducción | 369 |
| 2.6.1.3.3. Extracción de relaciones semánticas | 371 |
| 2.6.1.3.4. Elaboración de la ontología de dominio | 377 |
| 2.6.1.4. Fase IV: Elaboración de una base de datos ontoterminográfica | 382 |
| 2.6.1.4.1. Selección de campos | 384 |
| 2.6.1.4.2. Elaboración de definiciones | 390 |

| | |
|---|-----|
| 2.6.1.4.3. Selección de ejemplos contextuales | 399 |
| 2.6.1.5. Fase V: Revisión y validación | 401 |
| 2.6.1.6. Fase VI: Exportación y edición | 402 |
| | |
| CAPÍTULO 3. PREPARACIÓN DEL TRABAJO ONTOTERMINOGRÁFICO | |
| 3.1. Introducción | 413 |
| 3.2. Delimitación de trabajo | 415 |
| 3.2.1. Selección de las herramientas de trabajo | 418 |
| 3.3. Análisis del dominio | 447 |
| 3.3.1. El concepto <turismo> | 447 |
| 3.3.2. El turismo como fenómeno económico en España | 450 |
| 3.3.3. Los segmentos turísticos en auge | 457 |
| 3.3.3.1. El turismo de aventura como turismo alternativo | 461 |
| 3.3.3.1.1. Vínculos entre turismo y deporte | 465 |
| 3.3.3.1.2. El turismo de aventura en España | 468 |
| 3.3.3.1.3. Actividades del turismo de aventura y propuestas para su clasificación | 480 |
| 3.3.4. El discurso turístico como lenguaje de especialidad | 484 |
| 3.3.4.1. Características del lenguaje del turismo como lenguaje de especialidad | 486 |
| 3.3.4.1.1. Plano lingüístico | 491 |
| 3.3.4.1.2. Plano pragmático | 501 |
| 3.3.4.1.3. Plano funcional | 504 |
| 3.3.4.2. Características formales del texto turístico | 506 |
| 3.3.5. La traducción de textos turísticos en España | 509 |
| 3.3.5.1. Características de la traducción de textos turísticos promocionales: problemas, dificultades y errores frecuentes | 522 |
| 3.3.6. Recursos terminológicos monolingües, bilingües y multilingües para las lenguas inglés, alemán y español en el segmento de turismo de aventura | 531 |
| 3.3.6.1. Búsqueda y selección de los recursos terminológicos | 535 |
| 3.3.6.2. Clasificación de los recursos terminológicos | 549 |
| 3.3.6.3. Evaluación de los recursos terminológicos | 554 |
| 3.3.7. Esquema categorial del dominio | 564 |
| 3.4. Compilación del corpus AVENTURIS | 567 |
| 3.4.1. El subcorpus comparable AVENTURIS 1 | 573 |
| 3.4.2. El subcorpus paralelo AVENTURIS 2 | 578 |
| | |
| CAPÍTULO 4. REALIZACIÓN DEL TRABAJO ONTOTERMINOGRÁFICO Y SU EDICIÓN | |
| 4.1. Introducción | 585 |
| 4.2. Realización del trabajo ontoterminográfico con <i>ProTermino</i> | 588 |
| 4.2.1. Gestión del trabajo ontoterminográfico | 589 |
| 4.2.1.1. Aspectos previos | 589 |
| 4.2.1.2. Extracción de candidatos a término | 593 |
| 4.2.1.2.1. Validación y eliminación de candidatos a término | 596 |
| 4.2.1.3. Extracción de relaciones semánticas | 602 |
| 4.2.1.4. Elaboración de la ontología de dominio | 613 |
| 4.2.2. Elaboración de la base de datos ontoterminográfica | 621 |
| 4.2.2.1. Campos terminológicos para el proyecto <i>EcoTerm</i> | 621 |
| 4.2.2.1.1. Campo <i>Definition</i> | 629 |
| 4.2.2.1.2. Campo <i>Equivalence</i> | 633 |

| | |
|---|-----|
| 4.2.2.1.3. Campo <i>Context</i> y <i>Collocations</i> . Módulo de concordancias | 639 |
| 4.2.2.2. Resultado de entrada terminológica y validación | 642 |
| 4.3. Exportación y edición del trabajo ontoterminográfico | 643 |
| 4.3.1. <i>Trandix</i> : herramienta para la búsqueda terminológica en contexto | 645 |
| 4.3.1.1. Antecedentes | 645 |
| 4.3.1.1.1. <i>Locolex</i> y <i>COMPASS</i> | 646 |
| 4.3.1.1.2. <i>MobiMouse Plus</i> | 649 |
| 4.3.1.1.3. <i>Smarty</i> | 651 |
| 4.3.1.1.4. <i>Evroterm</i> y <i>TERMINATOR</i> | 652 |
| 4.3.1.2. Objetivos | 657 |
| 4.3.1.3. Descripción | 660 |
| 4.3.1.3.1. Descripción de funcionalidades | 661 |
| 4.3.1.3.1.1. Formato TBX (Term Base eXchange) | 675 |
| 4.3.1.3.1.2. Vinculación entre archivos | 684 |
| 4.3.1.3.2. Comparativa entre <i>Trandix</i> y sus antecedentes | 688 |
| 4.3.1.4. Evaluación de <i>Trandix</i> | 694 |
| 4.3.1.4.1. Fase preparatoria | 696 |
| 4.3.1.4.1.1. Finalidad del experimento | 696 |
| 4.3.1.4.1.2. Planificación del experimento | 697 |
| 4.3.1.4.1.3. Muestra del experimento | 698 |
| 4.3.1.4.1.4. Material para el experimento | 699 |
| 4.3.1.4.1.5. Preparación del entorno web | 706 |
| 4.3.1.4.2. Fase de ejecución | 709 |
| 4.3.1.4.3. Fase de evaluación | 712 |
| 4.3.1.4.3.1. Evaluación de traducciones y conceptos relacionados | 713 |
| 4.3.1.4.3.2. Métodos de evaluación | 717 |
| 4.3.1.4.3.3. Propuesta de plantilla mixta de evaluación | 724 |
| 4.3.1.4.3.4. Resultados del experimento | 726 |
| 4.3.1.4.3.4.1. Criterio de finalización de la tarea | 728 |
| 4.3.1.4.3.4.2. Criterio de tiempo de realización | 730 |
| 4.3.1.4.3.4.3. Criterio de calidad | 731 |
| 4.3.1.4.3.4.4. Criterio de satisfacción | 739 |
| 4.3.1.4.3.4. Conclusiones del experimento y trabajo futuro con <i>Trandix</i> | 748 |
| CONCLUSIONS | 751 |
| BIBLIOGRAFÍA | 767 |
| ANEXOS | |
| Anexo I | 811 |
| Anexo II | 819 |
| Anexo III | 823 |
| Anexo IV | 827 |
| Anexo V | 833 |
| Anexo VI | 841 |

INDICE DE TABLAS, FIGURAS Y GRÁFICOS

FIGURAS

| | |
|--|-----|
| Figura 1. Relación entre lengua general y lenguajes de especialidad | 58 |
| Figura 2. Relación entre lengua general, lengua común y lenguajes de especialidad..... | 60 |
| Figura 3. Elementos constitutivos de una situación comunicativa | 67 |
| Figura 4. A la derecha, triángulo semántico de Gomperz. A la izquierda, triángulo adaptado por Wüster | 93 |
| Figura 5. Ejemplo de etiquetado sintáctico de la frase «The dog was playing with a ball on the yard» con el analizador sintáctico <i>Link Parser</i> | 159 |
| Figura 6. Ejemplo de anotación semántica del treebank <i>TIGER</i> | 160 |
| Figura 7. Ejemplo de alineación bilingüe a nivel de párrafo con <i>YouAlign</i> | 161 |
| Figura 8. Ejemplo de líneas de concordancia monolingüe con el programa <i>MonoConc Pro</i> | 166 |
| Figura 9. Ejemplo de líneas de concordancia bilingüe con el programa <i>ParaConc</i> | 167 |
| Figura 10. Representación de sistema genérico de conceptos en forma de árbol con algunos deportes de invierno | 184 |
| Figura 11. Representación alfanumérica de sistema genérico de conceptos con algunos deportes de invierno | 184 |
| Figura 12. Representación de sistema partitivo de conceptos con corchetes con las partes de una canoa | 184 |
| Figura 13. Representación de sistema de conceptos mixto con algunas de las actividades acuáticas de río (sistema de conceptos genérico) y las partes de una canoa (sistema de conceptos partitivo) | 185 |
| Figura 14. Representación de sistema de conceptos asociativo (relación instrumental) con la actividad de buceo y algunos de sus instrumentos básicos realizada con el programa <i>GeneSis</i> | 186 |
| Figura 15. Ejemplo de fragmento de mapa conceptual creado con <i>MindMapper</i> | 208 |
| Figura 16. The Protégé Ontologie (http://protege.semanticweb.org) | 209 |
| Figura 17. Metabuscador <i>Tradudidática</i> | 334 |
| Figura 18. Resultado de búsqueda en el motor de búsqueda <i>2lingual</i> para el término «Turismo» en español y alemán | 335 |
| Figura 19. Fragmento de la ventana con los resultados de la búsqueda “buceo con botella” en el par de lenguas español-inglés en <i>Linguee</i> | 338 |

| | |
|--|-----|
| Figura 20. Metodología del trabajo terminográfico sistemático y genérico propuesto por Cabré Castellví (1993: 290 y ss.) | 342 |
| Figura 21. Flujo de trabajo de la metodología termontográfica (Temmerman y Kerremans, 2003) | 345 |
| Figura 22. Interfaz del programa <i>ReCor</i> | 363 |
| Figura 23. Ejemplos de resultados gráficos de <i>ReCor</i> después de analizar un corpus | 364 |
| Figura 24. Representación lineal y en grafos con el programa <i>Protégé</i> | 378 |
| Figura 25. Representación del esquema categorial prototípico en el que se basan los marcos semánticos específicos. Ejemplo extraído del proyecto de investigación Oncoterm (Faber Benítez, 2002: 7) | 379 |
| Figura 26. Fases de la propuesta metodológica para la elaboración de recursos ontoterminográficos para traductores | 408 |
| Figura 27. Resultado para la búsqueda del término «Myeloyd-Leukemia» en la base de datos Oncoterm. Fuente: < http://www.ugr.es/~oncoterm/alpha-index.html > | 428 |
| Figura 28. Pantalla inicial del programa <i>Categorisation Framework Editor</i> | 430 |
| Figura 29. Pantalla del programa <i>Termonography Workbench</i> con texto (izquierda) y categorización (derecha) cargados | 432 |
| Figura 30. Pantalla del programa <i>Dictionary</i> | 433 |
| Figura 31. Pantalla del programa <i>Corpógrafo</i> | 435 |
| Figura 32. Pantalla del programa <i>iMap</i> | 437 |
| Figura 33. Pantalla del programa <i>iConcept</i> | 438 |
| Figura 34. Pantalla del diccionario especializado como resultado de las fases anteriores | 438 |
| Figura 35. Interfaz de <i>Spaterm</i> con los tres módulos del primer nivel y los módulos del segundo nivel del gestor de corpus | 440 |
| Figura 36. Módulo y submódulos del editor de base de datos de <i>Spaterm</i> | 441 |
| Figura 37. Ejemplo de selección de campos para la exportación | 442 |
| Figura 38. Ejemplo de representación gráfica de relaciones ontológicas de una UT. Fuente: Manual de usuario de <i>Spaterm</i> (2009) | 443 |
| Figura 39. Ventana para la creación, eliminación o modificación de relaciones semánticas ofrecidas de forma automática por <i>Spaterm</i> . Fuente: Manual de usuario de <i>Spaterm</i> (2009) | 443 |
| Figura 40. Marco conceptual del turismo de aventura | 464 |
| Figura 41. Anuncio de Destination Loch Ness Ltd (DLN) para promocionar el Lago Ness, en Escocia | 506 |
| Figura 42. Página principal del directorio <i>Yourdictionary.com</i> | 539 |
| Figura 43. Pantalla de la búsqueda «Aerial Sports» en el diccionario visual de <i>Merrien Webster</i> (sección <i>Games and Sports</i>) | 552 |
| Figura 44. Marco semántico del segmento del turismo de aventura | 565 |
| Figura 45. Modelo A | 566 |
| Figura 46. Modelo B | 566 |
| Figura 47. Marco semántico del segmento del turismo de aventura con relaciones inversas | 567 |
| Figura 48. Fases de realización de la base de datos ontoterminográfica | 588 |
| Figura 49. Ventana de creación de proyectos en <i>ProTermino</i> | 590 |
| Figura 50. Creación del proyecto <i>Ecoterm</i> | 590 |
| Figura 51. Tabla de proyectos creados en <i>ProTermino</i> | 590 |
| Figura 52. Subida de documentos en <i>ProTermino</i> | 591 |
| Figura 53. Ventana para añadir uno o más documentos | 592 |
| Figura 54. Recuento de documentos y de palabras del proyecto <i>Ecoterm</i> | 592 |
| Figura 55. Opción <i>Recalculate term score</i> para la extracción de candidatos a término ... | 594 |
| Figura 56. Módulo Base de datos con el número de candidatos a término extraídos, de términos validados, de unidades eliminadas y de términos incluidos de forma manual .. | 595 |
| Figura 57. Ejemplo de introducción de término a través de la opción <i>Manually add a</i> | |

| | |
|---|-----|
| <i>term</i> | 595 |
| Figura 58. Resultado de búsqueda de concordancias con el término «barranquismo» ... | 606 |
| Figura 59. Marco semántico con especificaciones en categorías semánticas principales | 614 |
| Figura 60. Visualización de las relaciones semánticas confirmadas | 618 |
| Figura 61. Ejemplo de marco semántico de la actividad de barranquismo | 620 |
| Figura 62. Opciones disponibles para el campo <i>part-of-speech</i> | 623 |
| Figura 63. Opciones disponibles para el campo <i>gender</i> | 623 |
| Figura 64. Opciones disponibles para el campo <i>gramatical number</i> | 623 |
| Figura 65. Ejemplo de los cuatro primeros campos rellenos de la entrada terminológica relativa a la unidad «barranquismo» | 624 |
| Figura 66. Opciones disponibles para el campo <i>term status</i> | 625 |
| Figura 67. Campo <i>Standard body</i> | 625 |
| Figura 68. Opciones disponibles para el campo <i>Geographical Usage</i> | 626 |
| Figura 69. Opciones disponibles para el campo <i>Subdomain</i> | 626 |
| Figura 70. Ejemplo de los siguientes cuatro campos rellenos de la entrada terminológica relativa a la unidad «barranquismo» | 627 |
| Figura 71. Opciones disponibles para el campo <i>Term type</i> | 627 |
| Figura 72. Ejemplo de tipos de términos en la entrada de «barranquismo» | 628 |
| Figura 73. Ejemplo de información añadida al campo <i>Linguistic remarks</i> | 628 |
| Figura 74. Campo <i>definition</i> en la aplicación <i>ProTermino</i> | 629 |
| Figura 75. Ejemplo de equivalentes de traducción en el campo <i>equivalence</i> para el término «barranquismo» | 633 |
| Figura 76. Opciones de candidatos a equivalentes arrojados por la aplicación de forma automática para la unidad alemana « <i>Orientierung</i> » | 636 |
| Figura 77. Ejemplo de equivalentes validados para la unidad alemana « <i>Orientierung</i> » . | 637 |
| Figura 78. Candidatos a equivalente ofrecidos por la aplicación para la unidad inglesa « <i>helmet</i> » | 640 |
| Figura 79. Ejemplo de concordancias para el término «barranquismo» con el módulo <i>Term concordance</i> | 641 |
| Figura 80. Campo <i>Context</i> con la inclusión automática del contexto seleccionado en el módulo <i>Term concordance</i> | 641 |
| Figura 81. Ejemplo de entrada terminológica para la unidad «barranquismo» según la información guardada | 642 |
| Figura 82. Parte de la tabla de exportación de campos terminológicos | 643 |
| Figura 83. Búsqueda contextual con COMPASS. Fuente: Feldweg y Breidt (1997: 160) | 648 |
| Figura 84. Información que ofrece el asistente, utilizando un diccionario inglés-polaco, al seleccionar una palabra del texto. Fuente: Prószéky y Földes (2005: 2) | 650 |
| Figura 85. Pantalla de <i>Smarty</i> con información en sus diferentes apartados. Fuente: Arnaudov y Mitkov (2009: 3292) | 652 |
| Figura 86. Interfaz de <i>Terminator</i> | 654 |
| Figura 87. Análisis terminológico de <i>Terminator</i> | 654 |
| Figura 88. Resultados de <i>Terminator</i> (en <i>Evroterm</i>) | 655 |
| Figura 89. Ventana inicial de <i>Trandix</i> | 661 |
| Figura 90. Ventana principal de <i>Trandix</i> con el texto cargado (a la izquierda) | 663 |
| Figura 91. Opción Buscar en la ventana principal de <i>Trandix</i> | 664 |
| Figura 92. Menú <i>Archivo</i> de la ventana principal de <i>Trandix</i> | 664 |
| Figura 93. Menú <i>Edición</i> de la ventana principal de <i>Trandix</i> | 665 |
| Figura 94. Menú <i>Ayuda</i> de la ventana principal de <i>Trandix</i> | 665 |
| Figura 95. Ventana de <i>Preferencias</i> , en el menú <i>Edición</i> | 666 |
| Figura 96. Ventana de <i>Preferencias</i> en inglés | 667 |
| Figura 97. Ventana de <i>Preferencias</i> en alemán | 668 |
| Figura 98. Ventana principal de <i>Trandix</i> con el texto cargado (a la izquierda) y la entrada terminológica del término seleccionado (a la derecha) | 670 |

| | |
|--|-----|
| Figura 99. Proceso para realizar búsquedas externas | 672 |
| Figura 100. Resultado obtenido tras una búsqueda externa con el término «traje de neopreno» | 672 |
| Figura 101. Ventana para búsquedas convencionales | 673 |
| Figura 102. Ventana con el mensaje de petición de colaboración | 674 |
| Figura 103. Ejemplo de mensaje de correo electrónico con la información buscada de forma externa | 674 |
| Figura 104. Jerarquía de etiquetas en TBX | 679 |
| Figura 105. Ejemplo de etiquetas TBX. Fuente: <i>Colección de terminología de Microsoft</i> . URL: < http://www.microsoft.com/Language/es-es/Terminology.aspx > | 681 |
| Figura 106. Ejemplo de etiquetas TBX-Basic exportadas de <i>ProTermino</i> | 683 |
| Figura 107. Texto promocional seleccionado para la traducción 1 (sin <i>Trandix</i>) | 701 |
| Figura 108. Texto promocional seleccionado para la traducción 2 (con <i>Trandix</i>) | 702 |
| Figura 109. Fragmento del cuestionario 1 elaborado en el CV | 708 |
| Figura 110. Espacio virtual destinado al seminario teórico-práctico | 709 |
| Figura 111. Juego de botones creado para la evaluación mixta | 732 |
| Figura 112. Evaluación y corrección con <i>Markin</i> de una traducción sin la herramienta <i>Trandix</i> | 734 |
| Figura 113. Evaluación y corrección con <i>Markin</i> de una traducción con la herramienta <i>Trandix</i> | 736 |
| Figura 114. Ejemplo de resultados del cuestionario 2 en forma de gráfico | 739 |
| Figura 115. Resultados de la pregunta 18 del cuestionario 2 | 745 |
| Figura 116. Ejemplos de comentarios obtenidos en la pregunta 19 del cuestionario 2 ... | 747 |

TABLAS

| | |
|--|-----|
| Tabla 1. Niveles de especialización de un texto y características | 66 |
| Tabla 2. Fuente: Relación entre Terminología y Lexicología | 73 |
| Tabla 3. Diferencias entre la Terminología clásica y la TST. Extraído de Temmerman (1997: 56) | 94 |
| Tabla 4. Diferencias entre la Lexicografía y la Terminografía según Rondeau (1983: 62-63) | 118 |
| Tabla 5. Comparativa de la Terminografía con la Lexicografía especializada, según Bergenholtz y Tarp (1995) y adaptación propia | 124 |
| Tabla 6. Ejemplo de etiquetado gramatical (<i>POS-tagging</i>) y lematización con <i>TreeTager</i> | 158 |
| Tabla 7. Indicadores del primer bloque de la plantilla de evaluación | 294 |
| Tabla 8. Indicadores del segundo parámetro de la plantilla de evaluación | 296 |
| Tabla 9. Indicadores del tercer parámetro de la plantilla de evaluación | 299 |
| Tabla 10. Porcentajes obtenidos de la pregunta «¿Qué tipo de recurso terminológicos prefiere cuando traduce?» | 320 |
| Tabla 11. Porcentajes obtenidos en la pregunta «¿Qué considera que debería incluir un recurso terminológico dirigido a traductores?» de la encuesta | 323 |
| Tabla 12. Categorías y secciones de los campos recogidos en la norma ISO 12620: 1999 según la <i>CLS Framework</i> | 385 |
| Tabla 13. Perfil aproximativo de los practicantes de deportes de aventura en España. Fuente: Rivera Mateos (2007: 24) | 472 |
| Tabla 14. Análisis DAFO del segmento del turismo de aventura. Fuente: Garau <i>et al.</i> (2000: 602) | 473 |
| Tabla 15. Lista de actividades de turismo activo. Elaboración propia a partir de diversos decretos autonómicos | 478 |
| Tabla 16. Elaboración propia a partir de la clasificación de Peñalver Torres (2004: 185-198) | 483 |
| Tabla 17. Uso de Internet por parte de los turistas que visitaron España en 2009 | |

| | |
|--|-----|
| (Ministerio de Industria, Turismo y Comercio, 2010b: 53) | 489 |
| Tabla 18. Recursos terminológicos del sector turístico en general para las lenguas inglés, alemán y español | 534 |
| Tabla 19. Ejemplo de palabras clave utilizadas como descriptores para la búsqueda por palabras clave | 541 |
| Tabla 20. Recursos terminológicos del sector turístico de aventura para las lenguas inglés, alemán y español | 542 |
| Tabla 21. Ejemplo de palabras clave utilizadas como descriptores para la búsqueda por palabras clave ampliada | 544 |
| Tabla 22. Recursos terminológicos encontrados para la actividad de montañismo y escalada | 545 |
| Tabla 23. Recursos terminológicos encontrados para la actividad de escalada | 545 |
| Tabla 24. Recursos terminológicos encontrados para la actividad de espeleología | 546 |
| Tabla 25. Recursos terminológicos encontrados para la actividad de bicicleta | 546 |
| Tabla 26. Recursos terminológicos encontrados para la actividad del mundo ecuestre ... | 547 |
| Tabla 27. Recursos terminológicos encontrados para la actividad de vuelo sin motor ... | 547 |
| Tabla 28. Recursos terminológicos encontrados para la actividad de buceo | 547 |
| Tabla 29. Recursos terminológicos encontrados para la actividad de piragüismo, windsurf y actividades acuáticas en general | 548 |
| Tabla 30. Recursos seleccionados en formato papel | 551 |
| Tabla 31. Valoración del primer parámetro de <i>Glosario sobre deportes de aventura</i> | 555 |
| Tabla 32. Valoración del segundo parámetro de <i>Glosario sobre deportes de aventura</i> .. | 556 |
| Tabla 33. Valoración del tercer parámetro de <i>Glosario sobre deportes de aventura</i> | 557 |
| Tabla 34. Valoración del primer parámetro de <i>Bicycle Dictionary</i> | 559 |
| Tabla 35. Valoración del segundo parámetro de <i>Bicycle Dictionary</i> | 560 |
| Tabla 36. Valoración del tercer parámetro de <i>Bicycle Dictionary</i> | 561 |
| Tabla 37. Actividades del turismo de aventura | 570 |
| Tabla 38. Ejemplos de páginas web obtenidas con la búsqueda de palabras clave para el segmento de turismo de aventura | 572 |
| Tabla 39. Ejemplo de tabla de registro | 573 |
| Tabla 40. Recuento del subcorpus Aventuris 1 | 574 |
| Tabla 41. Recuento del subcorpus Aventuris 2 | 578 |
| Tabla 42. Ejemplo de conceptos seleccionados referentes a la terminología de la actividad de barranquismo | 600 |
| Tabla 43. Relaciones semánticas directas e inversas del dominio del turismo de aventura con respecto a las principales categorías semánticas | 604 |
| Tabla 44. Patrones lingüísticos recurrentes de las relaciones jerárquicas en español, inglés y alemán | 605 |
| Tabla 45. Ejemplo de recursos terminológicos utilizados para el análisis <i>in vitro</i> de patrones lingüísticos..... | 607 |
| Tabla 46. Lista de patrones lingüísticos detectados en las tres lenguas | 611 |
| Tabla 46 <i>Cont.</i> Lista de patrones lingüísticos detectados en las tres lenguas | 612 |
| Tabla 47. Ejemplo de patrones semánticos incluidos en <i>ProTermino</i> a través del módulo <i>Knowledge patterns</i> | 617 |
| Tabla 48. Ejemplo de relaciones semánticas validadas a partir de las relaciones propuestas automáticamente por <i>ProTermino</i> | 618 |
| Tabla 49. Categorías conceptuales y sus correspondientes plantillas definatorias | 631 |
| Tabla 50. Definiciones basadas en las plantillas definatorias correspondientes | 632 |
| Tabla 51. Ejemplo de texto etiquetado con <i>TreeTagger</i> | 685 |
| Tabla 52: Tabla comparativa entre <i>Trandix</i> y sus antecedentes | 690 |
| Tabla 53. Plantilla analítica en Toledo Báez (2008) | 721 |
| Tabla 54. Plantilla analítica del proyecto de innovación educativa <i>Diseño de una plantilla electrónica para la evaluación y revisión de proyectos de Traducción</i> | |

| | |
|--|-----|
| <i>especializada según las normas del sector</i> | 722 |
| Tabla 55. Plantilla holística de De Rooze (2003) | 723 |
| Tabla 56. Plantilla holística de Toledo Báez (2008) | 723 |
| Tabla 57. Plantilla analítica propia | 725 |
| Tabla 58. Plantilla holística propia | 726 |
| Tabla 59. Conjunto de errores encontrados en la evaluación mixta para la tarea 1 | 735 |
| Tabla 60. Conjunto de errores encontrados en la evaluación mixta para la tarea 2 | 737 |
| Tabla 61. Resultados obtenidos de las preguntas repetidas incluidas en los dos cuestionarios | 741 |
| Tabla 62. Resultados obtenidos de las preguntas cuantitativas del segundo cuestionario | 744 |

GRÁFICOS

| | |
|--|-----|
| Gráfico 1. Datos profesionales de los encuestados | 316 |
| Gráfico 2. Volumen de mercado de traducción en diferentes campos de especialidad ... | 317 |
| Gráfico 3. El sector turístico en las economías de los países de la OCDE. Fuente: OECD (2010: 8) | 451 |
| Gráfico 4. Llegada de turistas internacionales a España durante los años 2005-2010 (elaboración propia a partir de los datos extraídos del IET | 453 |
| Gráfico 5. Porcentaje de turistas internacionales según el IET. Movimiento de estudios en Fronteras (Frontur) | 455 |
| Gráfico 6. Situación actual y estrategia de actuación para fomentar el turismo de montaña en España. Fuente: TURESPAÑA (2008: 148) | 470 |
| Gráfico 7. Escala de fiabilidad de los recursos evaluados en porcentajes | 563 |
| Gráfico 8. Resultados de <i>Recor</i> para el corpus comparable de inglés | 575 |
| Gráfico 9. Resultados de <i>Recor</i> para el corpus comparable de español | 575 |
| Gráfico 10. Resultados de <i>Recor</i> para el corpus comparable de alemán | 575 |
| Gráfico 11. Traducciones terminadas y no terminadas para la tarea 1 | 729 |
| Gráfico 12. Traducciones terminadas y no terminadas para la tarea 2 | 730 |
| Gráfico 13. Tiempo de realización de traducciones de la tarea 2 | 731 |
| Gráfico 14. Total de traducciones Aptas y No aptas para las tareas 1 y 2 | 738 |

ABSTRACT

INDEX

1. Introduction
2. Aims and goals
3. Materials
 - 3.1. *ProTermino*
 - 3.2. *Trandix*
 - 3.3. Multilingual corpus
 - 3.4. Template of Assessment
 - 3.5. Electronic survey to professional translators
 - 3.6. *Markin*
 - 3.7. *ReCor*
 - 3.8. Electronic survey
4. Methods
 - 4.1. Research on terminology and terminography field
 - 4.2. Study of professional translators' needs and expectations
 - 4.3. Steps to elaborate a multilingual ontoterminography-based database
 - 4.4. Empirical experiments with semi-professional translators
5. Results
 - 5.1. Results of the research on terminology and terminography field
 - 5.2. Results of the study of professional translators' needs and expectations
 - 5.3. Results of the steps to elaborate a multilingual ontoterminography-based database
 - 5.4. Results of the empirical experiments with semi-professional translators
 - 5.4.1. Execution of tasks
 - 5.4.2. Time of execution
 - 5.4.3. Quality criteria
 - 5.4.4. Users' satisfaction
 - 5.4. Results from the electronic survey
6. Conclusions and future work
 - 6.1. Conclusions
 - 6.2. Future work

We do not know what thoughts went through the minds of those Indian and Chinese priests who several thousand years ago produced the first dictionaries. But we know that even if their individual steps may seem small, in reality they represented a gigantic step for humanity, probably more important than the famous step taken by Neil Armstrong in 1969, although far more anonymously (Tarp, 2008: 117).

1. INTRODUCTION

During the last decades, Terminology and its applied branch, Terminography, have been subjected to several changes which have provoked a radical turn in their traditional assumptions and methodologies. This has been due to a number of factors: first, the electronic revolution has affected terminology management (Bouringault et al 2001) and the terminographers' working methods are now based on computers and new technologies; second, the principles and methodologies of General Theory of Terminology have been seriously questioned by many authors (Gaudin 1993, Cabré Castellví 1999; Temmerman 2000, among others) and the terminology approach is turning into specialised lexicography; third, the influence of some linguistic disciplines such as Corpus linguistics, Cognitive linguistics and Computational linguistics on Terminology has incorporated new methodologies and different approaches to theory and practice, and fourth, parallel working methods and mutual interests of ontology engineers and terminologists have led to the inclusion of ontologies in Terminography.

These changes have produced a number of principles and methodologies which focus on multidimensionality, dynamism, description and variability of terms and also are based on new technologies and corpora.

The development in the discipline has been accompanied by the changes of linguistic technologies and resources in general, which have evolved over time (Cabré Castellví 2006: 93) and have been adapted to users' needs. However, it is an unfortunate reality that the majority of these resources are of little use to translators, and therefore many are obliged to resort to the creation of their own terminological resources from comparable corpora, from existing translations or just surfing the Web without ending. These inadequate resources often pose a problem for translators since it is well known that they usually work under time pressure and do not always have the opportunity to create their own resources. As it said in Durán Muñoz (2010a) this reality is due to the fact that professional translators have always been pushed into the background and their needs for this kind of resource have hardly ever been seriously taken into consideration. This is the reason why most terminological resources do not fulfill their expectations, both regarding quality and quantity. In this context, with the mushrooming of the quantity of online information and with the lack of adequate terminological resources, it is urged to investigate on new resources addressed to translators and their needs and thus, to reduce their time when searching for information and reliable resources.

Our research is in line with this necessity of providing adequate terminological resources to translators and also with the current approaches to modern terminology. In this sense, on the one hand, it takes into account translators' needs regarding content and format of terminological resources and on the other hand, its methodology combines the use of corpora and ontologies (called Ontoterminography) and includes aspects from the Function Theory (Tarp, 2008a). In order to apply this methodology, we choose the tourism domain, since tourism has always been the cornerstone of Spanish economy, and more exactly the adventure tourism domain, as it is now one of the new growing tourist segments and one of the sectors where there is a greater lack of resources and confusion in its terminology.

To sum up, this thesis aims to investigate new methods to develop terminological resources for translators and then design and implement a new resource which fulfills translators' needs and expectations.

2. AIMS AND GOALS

Our main research hypothesis states the following: Professional translators require terminological resources created specifically to them as user group, since they present concrete needs and requirements that are different from other user groups. This main hypothesis is presented together with other sub-hypotheses related on the one hand to the methodology and on the other hand to the applied work on the adventure tourism domain:

1. Ontoterminography combined with the Semantic Frame (Fillmore 1977, 1982, 1985) is an adequate methodology to elaborate terminological databases for translators.
2. Multilingual corpora (comparable and parallel), virtually compiled by means of electronic resources, are valuable terminological and phraseological sources which enable us to make up for deficiencies found in traditional terminological resources (dictionaries, databases, etc.).
3. The use of definitional templates based on semantic frames for a specialised domain provides the possibility to write definitions in a systematic homogeneous and complete way.
4. There is a lack of high-quality, complete and adequate terminological resources for the tourist sector in general and for the adventure tourism in particular to carry out translation of specialised texts in this domain.
5. The different current types of searches and consultation of terminological resources do not fulfill translators' needs as they require faster, easier and complemented ways of accessing information.

6. A terminological resource which provides a flexible terminological management by translators and which includes different types of searches (including contextual searches) will enhance translator's tasks and will help them achieve better results regarding time and quality.

The aims listed above are achieved by the following goals, being the first exposed the main goal of the research:

1. The design of a systematic and descriptive methodology with a linguistic-cognitive approach to elaborate adequate terminological resources for translators taking into account the criteria of multidimensionality and dynamism in specialised languages.
2. A thorough review of major work in terminology field by presenting the differences and similarities between traditional and modern theories in terminology but also the linguistic disciplines which have significantly influenced the current terminological lines of research.
3. A revision of the different information sources which translators may use while they are translating and searching for information.
4. A review of the most relevant taxonomies to classify terminological resources up to this moment and an updated and complete own proposal.
5. A protocol to evaluate electronic and paper-based terminological resources based on several parameters and a scale of reliability to quantitatively indicate their quality.
6. Define and specify professional translators' needs as real users of a terminological resource.
7. Specify adequate criteria to design and compile specialised corpora to be used in a terminography work.
8. Elaborate a multilingual terminological database (Spanish, English and German) based on an ontoterminographical work to assist translators

while doing translations of promotional texts of the adventure tourism domain.

9. Collect the terminology of the adventure tourism domain and organise it regarding its diverse denominative variables, neologisms, foreign words, etc.
10. Employ semantic frames, in diverse forms and levels, as a means of conceptually representing the information, both at a macrostructural and microstructural level.
11. Propose a new way of searching and consulting terminological information by means of the design and implementation of an innovative and efficient tool in line with translators' needs and expectations.
12. Empirically assess the designed tool and its adequacy to semi-professional translators' needs.
13. Analyse translators' impressions and opinions regarding the use of the designed tool by means of a survey and quantitative data.

Apart from these goals, there are also other aims related to point 8:

- 8.1. Carry out an analysis about the tourist sector in Spain from different viewpoints (economy, translation, and linguistics) and, more specifically a study about the adventure tourism domain and its main features.
- 8.2. Study different tools to carry out the ontoterminography work and select the most adequate one for the proposed methodology.
- 8.3. Search for electronic and paper-based terminological resources about the adventure tourism and assess them by using the template for assessment presented above.
- 8.4. The compilation of the corpus AVENTURIS, a representative multilingual corpus (Spanish, English, and German) of promotional texts consisting of webpages in the subdomain of adventure tourism.

8.5. The conceptualisation of the adventure tourism domain and its ontology in the shape of a semantic frame.

8.6. Elaboration of the multilingual terminological database (Spanish, English, and German) according to the norm ISO 12620 (1999).

All these goals were achieved in this thesis by means of the following materials and methods.

3. MATERIALS

In this section, we present all the materials that we have used to carry out this research.

3.1. *ProTermino*

One of the main materials used for this thesis is the tool *ProTermino*, a modified version of the Spaterm prototype (Castillo Rodríguez, 2010), developed by the Research Group in Computational Linguistics from the University of Wolverhampton.¹ This application includes three main modules: corpus management, terminological database and knowledge patterns. These three modules are also divided in several sub-modules: the first one, corpus management includes a) corpus uploading and b) term extraction; the second module integrates a) terminological entry templates according to the ISO standard, b) cognate identification, c) term concordance, d) ontology building, and d) terminology exportation in several formats (.PDF, .RTF, .HTML and .TBX); and finally

¹ This tool has been designed by Dr. Mitkov, Dr. Corpas Pastor and Dr. Ha, and has been implemented by Dr. Ha, from the Research Group in Computational Linguistics (University of Wolverhampton, UK) in the framework of the following research projects: Project for Excellence *La contratación turística electrónica multilingüe como mediación intercultural: aspectos legales, traductológicos y terminológicos* (HUM-892. 2006-2009), directed by Dr. Corpas Pastor, and the I+D+I project *Ecosistema: espacio único de sistemas de información ontológica y tesauros sobre el medio ambiente* (FFI2008-06080-C03-03/FILO; 2008-2011), co-directed by Dr. Corpas Pastor and Dr. Faber Benítez.

the third module includes the option to upload semantic patterns and provide candidate semantic relationships and instances taken from the corpus.

3.2. *Trandix*

This tool is the other main material employed in this thesis. This one, on the contrary to *ProTermino*, is been designed and implemented in the framework of this research.² It is an innovative and flexible computer application which aim is to fulfill the gap that is found in terminological resources regarding searches and consultation and thus facilitate the search procedure by translators. This tool offers several functionalities, which try to satisfy their users' needs:

- It provides several types of searches: textual searches, which is the search done on the same working text; term-based searches, which refers to the traditional searches by term offered in most terminological resources, and external searches, which allow users to consult other resources without having to change the application.
- Users can select the data that they want to visualize in terminological entries as well as the external terminological resources that they want to use while carrying out external searches in the Preferences section, that is, users can select or de-select the resources they need or the fields (definition, equivalents, context, etc.) they want to see when they carry out a terminological search.
- It works with TBX files and hence, it promotes reusability and interchange among terminological sources.
- It gives the opportunity to send feedback from users to the administrator by offering the option to send emails with different information before closing the application.

² The design and assessment is been carried out in this research but the implementation is been done by J.A. Fernández Sola (2011) in his graduate research project, addressed by Dr. Francisco Javier Durán Muñoz.

3.3. Multilingual corpus

The multilingual corpus AVENTURIS compiled for this research consists of two subcorpora: one comparable subcorpus (named AVENTURIS 1), containing texts originally written in Spanish, German and English, and one parallel corpus (called AVENTURIS 2), with texts originally written in Spanish and their translations into German and English.

The texts included in both subcorpora are electronic promotional texts obtained from webpages on the Internet, and the distribution of texts is shown below:

- Aventuris 1

| LANGUAGE | DOCUMENTS | TOKENS |
|----------|-----------|---------|
| SPANISH | 455 | 252.878 |
| ENGLISH | 634 | 252.199 |
| GERMAN | 646 | 255.805 |

- Aventuris 2

| LANGUAGE | DOCUMENTS | TOKENS |
|----------|-----------|--------|
| SPANISH | 40 | 58.643 |
| ENGLISH | 40 | 51.573 |
| GERMAN | 40 | 51.076 |

Both corpora were used to carry out the ontoterminography work and thus, they served as a source for term and equivalent extraction, searches of term concordances and contexts, extraction of knowledge patterns, etc.

3.4. Template for assessment

In order to carry out the assessment of terminological resources about the adventure tourism segment and also check their quality regarding content, format and users' needs, we designed a template for assessment with several parameters and criteria, each of which had a grade assigned. We also established a scale of reliability based on four different levels:

- 0- 8 points: Low reliability
- 9-16 points: Low-middle reliability
- 17-22 points: Middle-high reliability
- 23-30 points: High reliability

This template for assessment was designed to assess electronic and paper-based terminological resources in general but it was used to check the quality of terminological resources in the adventure tourism sector.

3.5. Electronic survey to professional translators

As part of this research, we were driven by the necessity to obtain information about professional translators' needs and expectations regarding terminological resources and thus, we designed and carried out a survey among professional translators. This survey was translated into different versions and distributed worldwide among associations, enterprises, distributional lists, etc. so as to reach as many professional translators as possible. The number of answers obtained can be considered very high (around 400 responses), taking into account the respondents' and the fact that it was carried out during summer and many translators were on holidays.

3.6. *Markin*

The software *Markin* is an application which helps teachers mark and annotate texts. Once the evaluation parameters were established, *Markin* was used to evaluate the translations carried out by semi-professional translators. These evaluation parameters consist of both analytic error evaluation as well as global evaluation. The former pays attention to mistakes related to grammar, sense, style, coherence, function, or terminology and also to positive aspects such as correct translations of terms or expressions. The holistic evaluation evaluates the translation as a whole and it has four

different levels regarding the quality of the result. The levels range from 1 (very poor translation) to 4 (excellent translation).

3.7. ReCor

ReCor (version 3.1) is used to measure the representativeness of the corpus compiled in this thesis (AVENTURIS) and its results are shown in graphics and discussed in Chapter 3.

3.8. Electronic survey

Two electronic surveys were conducted so as to generate quantitative and qualitative data regarding the use of the application *Trandix* in direct translations by semi-professional translators. The results of these surveys are discussed in Chapter 4.

4. METHODS

Regarding the methodology followed in this research to achieve the aims and goals listed above, we can divide this section into four different sub-sections: 1. Research on terminology and terminography field, 2. Study of professional translators' needs and expectations, 3. Steps to elaborate a multilingual ontoterminography-based database, and 4. Empirical experiments with semi-professional translators.

4.1. Research on terminology and terminography field

We undertook a revision of the main terminology theories up to the present moment, from Wüster's traditional theory to the three main modern theories that have emerged in the last decade, that is, the Socioterminology approach (Gaudin 1993,

Boulanger 1995), the Socio-cognitive Theory of Terminology (Temmerman 2000), and the Communicative Theory of Terminology (Cabr  Castellv  2000).

Also, we present the different linguistic disciplines which have considerably influenced Terminology, which are: Corpus Linguistics, Cognitive Linguistics and Computational Linguistics, and the aspects that have been incorporated to Terminology thanks to them, highlighting the use of corpora and ontologies nowadays. At the same time, we present the current lines of research in Terminography and different projects and applications that follow them, both national and international projects.

4.2. Study of professional translators' needs and expectations

As it is said throughout this research, one of our main goals is to study and propose new terminological resources for translators. But, before we make any proposal it is necessary to approach their needs and expectations, as Bergenholts and Tarp (1995: 77) states in this quotation:

Lexicographical work often proceeds without any prior knowledge of the potential user group, and the dictionary may therefore be said to be the result of the lexicographer's own conjectures concerning user needs for lemmata, collocations, sentence examples, encyclopaedic and linguistic information, etc. To acquire more precise knowledge, the lexicographer may make a user survey before starting actual work on the dictionary, with the aim of uncovering the needs of potential users in relation to the information categories to be incorporated in the dictionary as well as the representation of this information (Bergenholtz y Tarp, 1995: 77).

Following their instructions, we designed and carried out a survey among professional translators in order to acquire knowledge about their needs and expectations regarding terminological resources. This survey was translated into different versions and distributed worldwide among associations, enterprises, distributional lists, etc. so as to reach as many translators as possible. By means of this survey, which was responded by approximately 400 professional translators, we were able to approach professional translators' needs and expectations and therefore to know more about their "ideal" terminological resource.

4.3. Steps to elaborate a multilingual ontoterminography-based database

We make a review about the most relevant methodologies proposed to carry out research on Ontoterminography and, afterwards, we proposed our own methodology mainly based on the use of corpora, new technologies and translators' needs.

These steps proposed as part of this methodology can be classified into the following phases:

1. WORK-TYPE DELIMITATION AND DOMAIN ANALYSIS

- 1.1. Specify pragmatic-linguistic variants (user group, working languages, etc.).
- 1.2. Select technological and human resources according to the project possibilities.
- 1.3. Thoroughly study the selected domain.
- 1.4. Collect and assess terminological resources in the selected domain.
- 1.5. Propose a provisional conceptualisation of the domain.

2. SPECIALISED CORPUS COMPILATION

- 2.1. Design the corpus
- 2.2. Search texts
- 2.3. Organise texts

3. ONTOTERMINOGRAPHY MANAGEMENT

- 3.1. Extract term candidates and validate
- 3.2. Search equivalent candidates
- 3.3. Extract semantic patterns and relationships
- 3.4. Create the domain ontology

4. ELABORATION OF AN ONTOTERMINOGRAPHY-BASED DATABASE

- 4.1. Select terminological fields
- 4.2. Write definitions
- 4.3. Select contextual examples

5. REVISION AND VALIDATION

6. EXPORTATION AND EDITION

This methodology proposed in the framework of this research was applied to the adventure tourism sector and thus, all the steps indicated above were followed to create a multilingual ontoterminography-based database for the adventure tourism domain and its exportation into an innovative application, *Trandix*, which allows users to satisfy their needs regarding terminological searches and consultations.

4.4. Empirical experiments with semi-professional translators

To prove the efficiency and adequacy of the tool *Trandix* we organised an empirical experiment with semi-professional translators (undergraduate students in their 4th year in Translation and Interpreting at the University of Málaga) attending the course *Traducción especializada inglés/español-español/inglés (b): textos científico-técnicos* during the academic year 2010/2011. These students were able to test the tool and give quantitative and qualitative data by means of a survey. This experiment was divided in three parts: the first one consisted of an introduction to new technologies and tools supporting the research phase of translation (information sources in general, comprehensive assistants including *Trandix*, etc.); the second one was devoted to a translation of an original promotional text in English about an adventure tourism activity without *Trandix*, but with the possibility to use any resource available on the Internet, and the last one was organised as another practical session in which the subjects had to carry out a translation with the same features as in the previous session but this time with the tool *Trandix*. The experiment took place in two different days: the first day was devoted to the theoretical session, and the second day was dedicated to carry out the practical session, which was temporary organised as follows:

1. The session started with the introduction of the experiment and a brief explanation about the material, requirements and the codification to be used. Time: 15 minutes.
2. Students translated Text 1 from English into Spanish (direct translation) with online information (dictionaries, databases, comparable corpora, etc.). They were not allowed to use the tool *Trandix* in this phase. Time: 20 minutes.
3. After finishing the first translation, they were required to answer the survey about the first part of the experiment. Time: 15 minutes.
4. Before starting the second part of the experiment, we proceeded to briefly explain the steps to be followed when using the tool *Trandix* and its functionalities. Time: 15 minutes.

5. Students translated Text 2 from English into Spanish (direct translation) with the application *Trandix*. Time: 20 minutes.
6. Students filled in the electronic survey about this second part of the experiment. This final part took 15 minutes.

In order to assess the results obtained in the practical sessions, we established four criteria, which measured different aspects of the tool and the procedure of translating with it. These criteria are discussed in Chapter 4 and the results obtained are also presented therein.

5. RESULTS

To present the results of this research, we will follow the same division as the previous section, that is: 1. Results of the research on terminology and terminography field, 2. Results of the study of professional translators' needs and expectations, 3. Results of the steps to elaborate a multilingual ontoterminography-based database, 4. Results of the empirical experiments with semi-professional translators.

5.1. Results of the research on terminology and terminography field

The research on terminology field led us to review the modern terminology theories and their differences to traditional terminology, as well as a revision of the most frequent dichotomies in this field: general language Vs. specialised language, terminology Vs. lexicology, and terminography Vs. lexicography. Therefore, this study conducted us to provide a review of the state-of-the-art in terminology field at the present moment.

Within this study we were able to present one of the most recent lines of research in terminography: ontoterminography, which is based on corpora and ontologies to carry out a terminological work. Related to this, we presented a revision of

different projects based on this new line and also the advantages and disadvantages of using corpora and ontologies in terminology.

5.2. Results of the study of professional translators' needs and expectations

The results obtained in this research clarify the needs and expectations that professional translators as real users present. We now know more about their opinions regarding the current terminological resources and have given them the opportunity to describe their “ideal” resource.

The results obtained from this research differ from the conclusions drawn in previous studies (based on trainee translators and second language students), which defended the needs to include linguistic, pragmatic and semantic information in resources for translators. With our research we observed that professional translators consider semantic information (semantic relations, semantic frames or domains, etc.) as desirable data but not as essential data, *i.e.* they do not see this information as essential but only as complementary. Consequently, we observe how trainee translators need different information than professional translators and therefore, it is necessary to distinguish between each user group' needs in order to elaborate resources for these professionals.

Also, it is relevant to take professional translators' needs into consideration in order to provide quick and easy access to information in online resources and include good and concrete definitions together with pragmatic information (context, tips of use, information about false friends, etc.), which help them understand source terms and correctly translate them.

5.3. Results of the steps to elaborate a multilingual ontoterminography-based database

Before we proposed our own methodology to create multilingual ontoterminography-based databases, we carried out a review of the methodologies applied for terminography works in this context, and we mainly focused on the methodology proposed by Temmerman and Kerremans (2003) and Faber Benítez and Jiménez Hurtado (2002). These two methodologies, together with the one proposed by Cabré Castellví (1993) are the cornerstone of our own proposal.

As said above, our methodology was divided into six different phases, which were also sub-divided in more steps. The results of this point are mainly two:

- On the one hand, a methodology based on Ontoterminography for the elaboration of terminological resources specifically addressed to translators.
- And, on the other hand, the application of this methodology to the adventure tourism domain and thus, the elaboration of a multilingual database (Spanish, English and German) about adventure tourism for translators. Within this result, we can highlight the execution of several internal aims:
 - Exhaustive analysis of the tourist sector from different viewpoints: economic, considering tourism as an economic sector; linguistic, considering it as the language of tourism; translation, as a translation of specialised tourist texts, and legislative, as an object to be regulated by law.
 - Conceptualisation of the domain.
 - Compilation of an adventure tourism corpus.
 - Collection of the most representative terminology in this domain.
 - Extraction of candidate terms, equivalents and semantic relations.
 - Proposal of an innovative tool, *Trandix*, to complete the process of terminology work after exporting the terminological database and thus,

fulfilling translators' needs and expectations related to searches and consultations.

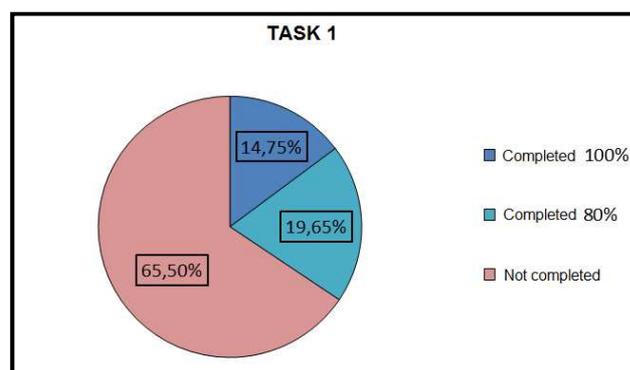
5.4. Results of the empirical experiments with semi-professional translators

In order to test the efficiency of the application *Trandix* in a quantitative and qualitative way, four criteria were established: execution of the tasks, time of execution, quality and user' satisfaction criteria. And, also, in order to assess translation quality we established several quality criteria related to analytical and holistic assessment, as we will see below.

5.4.1. Execution of tasks

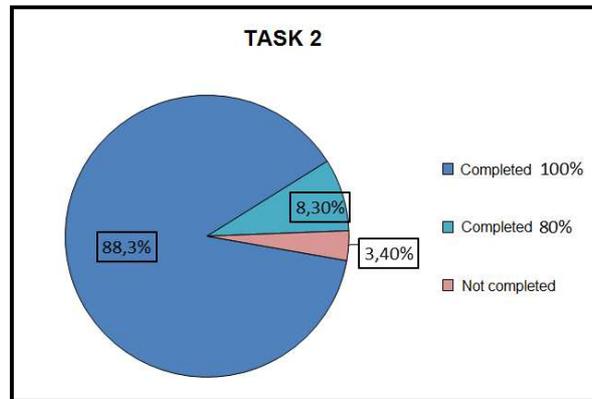
This first criterion takes into account whether the semi-professional translators managed to finish both tasks (translation of texts 1 and 2) in the time given (in this case, 20 minutes for each task). This criterion served as a filter for the next criteria, since all the unfinished translations were discarded, as we consider that a translation that has not been fully completed shows bad quality by itself and it is not worth assessing its quality.

In the following graphic we can observed the results obtained in the first task:



As we can see, most of the semi-professional translators did not have enough time to finish their translation and therefore they were considered as not completed and thus, discarded.

On the contrary, the results obtained in the second task, with *Trandix*, were radically different as we can observe in the next graphic:

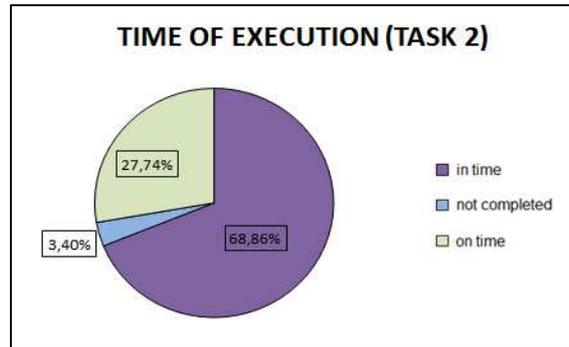


In this context, we can confirm the advantages that *Trandix* brings to translation regarding time and the number of words translated per minute.

5.4.2. Time of execution

The second criterion is about time of execution, that is, the time that semi-professional translators need to achieve the tasks required. This criterion is interesting, since it indicates whether translators finished their translation on time or before the given time (20 minutes).

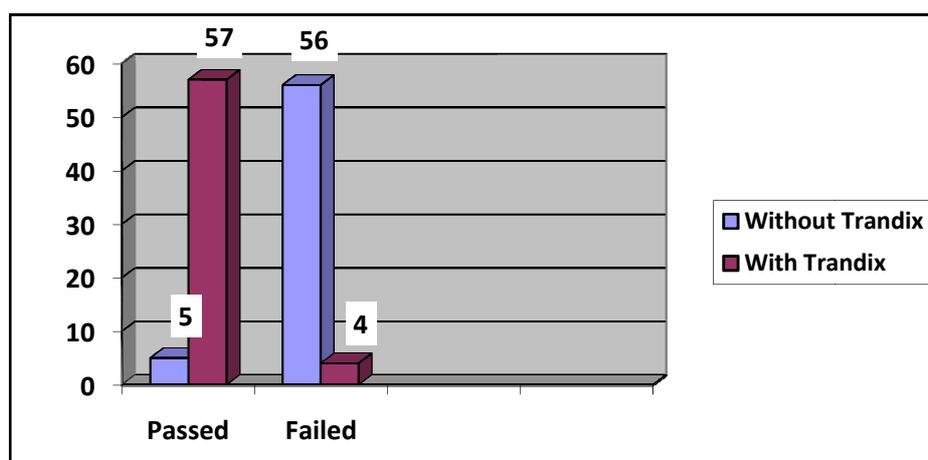
We can highlight that none of the students finished their translation in time during their first task (without *Trandix*) and very few completed it, as said above. On the contrary, there was a high number of students who finished the second task in time and very few who did not manage to. In the following graphic we can observe the proportion for the second task:



Again, we can confirm the advantages that *Trandix* offer regarding time of translation compared to other search protocols, since more than half of students finished their translations before the time had finished.

5.4.3. Quality criteria

Quality criteria are related to the evaluation parameters developed in this thesis and discussed in Chapter 4. Our main interest is to test whether the translations with *Trandix* as terminological source present a better quality (i.e. fewer errors) than translations carried out with usual resources (mainly online). We have already checked that *Trandix* reduces time of translating a text, a very important fact taking into account that translators are always working under time pressure. However, to the present moment, we have not measured whether *Trandix* also increases translation quality. In order to prove this, all the translations produced in the empirical experiment were evaluated with the software *Markin* according to the analytic and holistic evaluation parameters presented in Chapter 4. The results are summed up in the following tables:



According to these results, we can also confirm that the quality of translations carried out with the help of the application *Trandix* achieves better results than texts translated without it. To sum up, we can draw the conclusion that the aspects regarding quality and time improve when *Trandix* is used and decrease when other methods (traditional methods) are employed.

5.4.4. Users' satisfaction

A complete research must also offer qualitative results, and the surveys used in this thesis consisting of different types of questions (multiple choice, Yes/No questions, opinion questions, etc.) provide some indications about translators' impressions and opinions.

Firstly, we will present the questions that were repeated in both surveys, which make reference to respondents' opinions regarding the translated text, given time, most difficult aspects, etc. These questions conduct us to compare the methodology and results obtained in task 1 with the methodology with *Trandix* in task 2 and draw some conclusions.

| Repeated questions | SURVEY 1 (without <i>Trandix</i>) | SURVEY 2 (with <i>Trandix</i>) |
|--|--|--|
| 6. How did you find the translation? Please indicate from 1 to 5 the level of difficulty, being 1 very easy and 5 very difficult). | 2 → 21,9% 3 → 60,9% 4 → 9,4% 5 → 7,8% | 1 → 32,76% 2 → 44,83% 3 → 17,24% 4 → 5,17% |
| 7. Which are the most difficult aspects found in the translation? <input type="checkbox"/> Lack of information resources <input type="checkbox"/> Unknown terminology <input type="checkbox"/> Grammar and syntax <input type="checkbox"/> Format (ST and/or TT) <input type="checkbox"/> Cultural references <input type="checkbox"/> Other | Lack of information resources: 33% Unknown terminology: 98% Grammar and syntax: 11% Format (ST and/or TT): 5% Cultural references: 10 Other: 5% | Lack of information resources: 5% (3 personas) Unknown terminology: 68% Grammar and syntax: 29% Format (ST and/or TT): 11% Cultural references: 12% Other: 6% |
| 9. Do you think you had enough time to carry out the translation? <input type="checkbox"/> Yes <input type="checkbox"/> No | NO: 100% | YES: 84% NO: 16% |
| 10. Which level of difficulty have you found to find required information? Please indicate from 1 to 5 the level of difficulty, being 1 very easy and 5 very difficult). | 1 → 6,2% 2 → 18,5% 3 → 46,1% 4 → 16,9% 5 → 12,3% | 1 → 53,1% 2 → 31,0% 3 → 3,3% 4 → 8,3% 5 → 4,3% |
| 11. Which is your level of satisfaction with your translation? <input type="checkbox"/> Very satisfied <input type="checkbox"/> Satisfied <input type="checkbox"/> A little dissatisfied <input type="checkbox"/> Not satisfied | Satisfied: 13% A little dissatisfied: 63% Not satisfied: 19% | Very satisfied: 14% Satisfied: 78% A little dissatisfied: 9% |

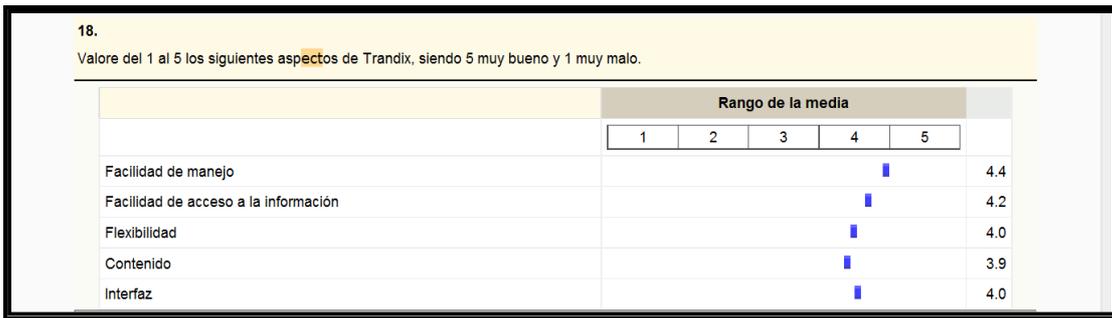
Once we have presented the results obtained from the questions repeated in both surveys, we can confirm that students' satisfaction with the second translation is much

higher than with the first translation as well as all the parameters indicate that the application *Trandix* makes easier and quicker the required translation. Now we continue exposing the answers received through the second survey related to the tool *Trandix*.

First, we show the answers to quantitative questions:

| Quantitative questions | SURVEY 2 (with <i>Trandix</i>) |
|--|---|
| 7. How did you find the use of <i>Trandix</i> ? Please indicate from 1 to 5 the level of difficulty, being 1 very easy and 5 very difficult). | 1 → 67,3% 2 → 31% 3 → 1,7% |
| 8. Do you think that <i>Trandix</i> has helped you during the translation? <input type="checkbox"/> Completely <input type="checkbox"/> Quite a lot <input type="checkbox"/> Enough <input type="checkbox"/> A little <input type="checkbox"/> Not at all | Completely → 50% Quite a lot → 38% Enough → 12% |
| 14. If you had to choose between <i>Trandix</i> and the previous methodology, which one would you prefer? <input type="checkbox"/> <i>Trandix</i> <input type="checkbox"/> Previous methodology | <i>Trandix</i> : 97% Previous methodology: 3% |
| 16. If <i>Trandix</i> would be available, would you use it? <input type="checkbox"/> Yes. <input type="checkbox"/> No | Yes: 100% |

Also, we find another quantitative question related to *Trandix* and different aspects. Respondents were asked to choose from 1 to 5 (being 1 very bad and 5 very good) to indicate their satisfaction with the following aspects: ease of use, ease of Access, flexibility, content and interface. As we can see in the next figure, the mean is around 4 in all cases:



With these results, we observe respondents' level of satisfaction, which is very high in all aspects. In general, we can again confirm the advantages that *Trandix* offer as a terminological resource.

Finally, respondents also gave their opinion in a qualitative way about positive and negative aspects of *Trandix*. As positive, all of them agreed with the fact that *Trandix* reduces time, improves translation quality, facilitates searches and consultations both within the tool and outside it by means of external searches, is flexible when selecting terminological information and is easy to use. On the other hand, they also indicated several negative aspects, which are related to its limitation to the adventure tourism sector, the working languages, the reduced number of linked resources and lack of the Undo option and other text processor options (color, orthography corrector, etc.).

Based on the results of the experiment and their assessment, we can confirm the benefits that our tool offers to translation regarding time of execution, efficiency, ease of use, among other aspects indicated in Chapter 4, and also highlight that 100% of respondents would use *Trandix* in their translations if it was available. Therefore, comparing it with a methodology without *Trandix* the advantages obtained become clear.

6. CONCLUSIONS AND FUTURE WORK

This section can be divided into two sub-sections: the first one is devoted to the main conclusions obtained in this thesis, regarding the goals proposed at the beginning, and the second one deals with the future lines of research.

6.1. Conclusions

The main contributions of this thesis are the methodological proposal to create terminological resources for translators based on Ontoterminography and also the achievement of designing and implementing an innovative tool which fulfills translators' needs and expectations. We will further discuss all the goals that have been achieved below:

- Goal 1 was devoted to the design of a systematic and descriptive methodology with a linguistic-cognitive approach to elaborate adequate terminological resources for translators. This goal was achieved in Chapter 2, where the proposed methodology was presented as well as the previous and close methodologies. Nevertheless, the application of this methodology was carried out in Chapters 3 and 4.
- Goal 2 was about carrying out a thorough review of major work in terminology field by presenting the differences and similarities between traditional and modern theories in terminology but also the linguistic disciplines which have significantly influenced the current terminological lines of research. This one was achieved in Chapter 1, where the state of the art in terminology was presented as well as the aspects incorporated due to the influence of other linguistic disciplines.
- Goal 3 was to assess the different information sources which translators may use while they are translating and searching for information, and it was achieved in Chapter 2, when a taxonomy of the different information resources was presented.

- Goal 4 and 5 were to study the most relevant taxonomies to classify terminological resources up to this moment, propose a new updated and complete classification of these resources and evaluate electronic and paper-based terminological resources following an assessment protocol. These goals were both achieved in Chapter 2.

- The goal 6 was to define and specify professional translators' needs as real users of a terminological resource, and it was managed by the design and distribution of a survey for professional translators, which is presented in Chapter 2.

- Goal 7 was based on the elaboration of a multilingual terminological database (Spanish, English and German) based on an ontoterminographical work to assist translators while doing translations of promotional texts in the adventure tourism domain. As said above, this goal is found in Chapters 3 and 4, where the proposed methodology is put into practice. This achievement includes specific goals, also managed in Chapters 3 and 4:

1. The goal to carry out an analysis about the tourist sector in Spain from different approaches (economy, translation, and linguistics) and, more specifically a study about the adventure tourism domain and its main features.
2. The goal to study different tools to carry out the ontoterminography work and select the most adequate one for the proposed methodology.
3. The goal to search for electronic and paper-based terminological resources about the adventure tourism and assess them by using the template for assessment presented above.
4. The goal to compile the corpus AVENTURIS, a representative multilingual corpus (Spanish, English, and German) of promotional texts from webpages in the subdomain of adventure tourism.
5. The goal to conceptualise the adventure tourism domain and its ontology in the shape of a semantic frame.

6. The goal to create the multilingual terminological database (Spanish, English, and German) according to the norm ISO 12620 (1999).
 - The goal 8, based on the collection of terminology of the adventure tourism domain and its organisation regarding its diverse denominative variants, neologisms, foreign words, etc., was also achieved in Chapter 4, when the term extraction was carried out and we were given the opportunity to analyse the terminology.
 - The goal 9 was to employ semantic frames, in its diverse forms and levels, as a means of conceptually representing the information, both at a macrostructural and microstructural level. This goal started in Chapter 1, when Frame Semantics was introduced in this research, but it was put into practice in Chapter 4, for the organisation of the adventure tourism sector.
 - Goal 10 was to propose a new way of searching and consulting terminological information by means of the design and implementation of an innovative and efficient tool in line with translators' needs and expectations. The achievement of this goal was possible thanks to the design and implementation of *Trandix*, the innovative tool that assist translators during their terminological searches and consultations.
 - Finally, goals 11 and 12 were to empirically assess the designed tool and its adequacy to semi-professional translators' needs and to analyse translators' impressions and opinions regarding the use of the designed tool by means of a survey and quantitative data, respectively. Both goals were achieved through a survey carried out during an empirical experiment organised within this research.

6.2. Future work

During this research a series of possible future directions have emerged. They are briefly discussed in this section.

1. The empirical study carried out with semi-professional translators should be repeated with professional translators. The result would serve as a more accurate indicator of the advantages of the application *Trandix* as a terminological resource in real translation markets.
2. Another line of future research related to our application is to extend the domain and working languages. To its current implementation, we have focused on a very specific domain (adventure tourism) and languages (Spanish, English, and German), because they were the relevant data for our terminological project. However, as any TBX file could be uploaded and used in the application, it would be of particular interest to update *Trandix* and include other information resources for external searches so that it could be used for other specialised texts.
3. Related to the previous point, it could be also useful to repeat the same empirical study, but with other domains and languages so as to prove whether the tool offers good results.
4. Also, it could be possible to test other methods for term and equivalent candidate extractions, like the methodology proposed by Le Serrec et al. 2010, Gamallo (2007, 2008) and Gamallo and Pichel (2005) or Ha et al. (2008), and compare results with the cognate methodology used in this research.
5. Regarding the difficulties found in *Trandix* due to the lemmatizer employed (*TreeTagger*), especially with the Spanish language, we consider it is necessary to test other options or even implement a new one.
6. Finally, another future line of research, also related to *Trandix*, is to study the possibilities of implementing this system into the most used translation memories, like *SLD Trados*, *Déjà Vu X*, *Across*, among others. Nowadays, most of these computer-assisted translation tools include terminological management systems, some of which are also

available as independent tools, like *SDL MultiTerm 7*, *TermStar XV* or *Uniterm*. These systems are helpful while translating but present several limitations: they cannot be interchanged between different systems, must be fed by translators, etc. In this sense, we consider it could be useful to integrate a textual search system like *Trandix*, which enables users to improve their terminological searches and consultation, provides them with a database developed in a systematic and coherent way through TBX files and saves their time by avoiding searches in numerous resources and by eliminating their needs to create terminological records in their terminological management systems.

Most of these future directions will be developed in the current research project *Ecosistema: espacio único de sistemas de información ontológica y tesauros sobre el medio ambiente* (FFI2008-06080-C03-03/FILO; 2008-2011), directed by Dr. Corpas Pastor and Dr. Faber, where the most innovative lines in terminology, ontology, and automatic summarization will emerge.

INTRODUCCIÓN

El comienzo de esta tesis doctoral se encuentra en el año 2007, momento en el que nos fue concedida una beca de Formación de Personal Universitario (FPU) por el Ministerio de Educación y Ciencia,³ conducente a la formación docente e investigadora. A raíz de esta concesión, entramos a formar parte del grupo de investigación "Lexicografía y Traducción" (HUM-106, Junta de Andalucía) de la Universidad de Málaga, dirigido por la catedrática Gloria Corpas Pastor y directora de nuestra tesis, y participamos como miembro investigador en el proyecto de excelencia denominado *La contratación turística electrónica multilingüe como mediación intercultural: aspectos legales, traductológicos y terminológicos*.⁴ En este marco de trabajo, defendimos nuestra memoria de máster, dentro del Máster Oficial en Traducción, Mediación Cultural e Interpretación de la Universidad de Málaga, titulada *Catalogación y evaluación de las fuentes terminológicas del sector turístico (alemán-español)*, que nos

³ Beca AP2006-03327. Resolución 31 de agosto de 2006. Ministerio de Educación y Ciencia.

⁴ Proyecto de Investigación de Excelencia en equipos de investigación andaluces de las Universidades públicas y organismos de investigación de Andalucía (Orden de 5 de julio de 2005; BOJA n.º 138, de 18 de julio). Dirección General de Investigación, Tecnología y Empresa, Secretaría General de Universidades, Investigación y Tecnología de la Junta de Andalucía (n.º de referencia HUM-892. Resolución de 26 de enero de 2006). Inv. Principal: Gloria Corpas Pastor. Duración: 2006-2009 (<http://www.uma.es/hum892>).

valió para la obtención del título de máster y la posibilidad de continuar con la carrera investigadora. Dicho trabajo, que constituye el punto de partida para la presente investigación, se centró en la catalogación y evaluación de las fuentes terminológicas del sector turístico en general en el par de lenguas español-alemán, con objeto de llevar a cabo una aproximación al sector del turismo y a la situación de la traducción de textos turísticos y, a la par, conocer los recursos terminológicos disponibles para las lenguas español y alemán y evaluarlas según las necesidades especificadas por traductores profesionales a través de los resultados obtenidos de una encuesta.

Anterior a este proyecto, también participamos con una Beca-Colaboración concedida por el Ministerio de Educación y Ciencia (BOE, de 17 de junio de 2005) en el proyecto interuniversitario e interdisciplinar denominado *TURICOR: Compilación de un corpus de contratos turísticos (alemán, español, inglés, italiano) para la generación textual multilingüe y la traducción jurídica*,⁵ cuyos estudios y trabajos realizados durante dicho proyecto, especialmente relacionados con la metodología para la compilación de corpus así como con la herramienta para la representación de corpus diseñada e implementada en dicho proyecto, *ReCor* (Seghiri Domínguez y Corpas Pastor, 2006), han sido de gran ayuda a la hora de llevar a cabo esta tesis doctoral.

En el año 2009 se nos concedió un proyecto interuniversitario e interdisciplinar en una línea similar, denominado *ECOSISTEMA: Espacio Único de Sistemas de Información Ontológica y Tesauro sobre el Medio Ambiente*,⁶ en el que también participamos como miembro investigador y que, por supuesto, nos siguió delimitando la línea de trabajo de esta tesis doctoral, en este caso con relación al empleo de las ontologías en la terminología y en la organización del conocimiento especializado.

Los tres proyectos nombrados anteriormente se encuentran en plena sintonía con las políticas nacionales y europeas en materia científica, las cuales se decantan

⁵ Proyecto nacional concedido por el Ministerio de Ciencia y Tecnología dentro del Plan Nacional de Investigación Científica, Desarrollo e Innovación Tecnológica y por la Unión Europea (Fondos FEDER). N.º de ref. BFF2003-04616. Inv. principal: Dra. D.ª Gloria Corpas Pastor. Duración: 2003-2006. (<http://www.turicor.com>).

⁶ Entidad financiadora: Ministerio de Ciencia y Tecnología. Dirección General de Investigación. VI Plan Nacional de Investigación Científica, Desarrollo e Innovación Tecnológica 2008-2011, a través del subprograma de Proyectos de Investigación Fundamental no Orientada. N.º de ref. FFI2008-06080-C03-03. Proyecto coordinado UGR- UMA. Inv. Principal UMA: Gloria Corpas Pastor. Inv principal UGR: Pamela Faber. Duración: 2008-2011

claramente por las aplicaciones de la ingeniería lingüística en las industrias de la lengua. Concretamente, los dos proyectos más recientes, y en los que se enmarca específicamente este trabajo de investigación, se encuentran en las líneas más innovadoras de este tipo de investigación:

Por un lado, el proyecto de excelencia denominado La contratación turística electrónica multilingüe como mediación intercultural: aspectos legales, traductológicos y terminológicos, se inscribe en las líneas de investigación más punteras de los programas europeos IST-Information Society Technologies y consiste en la localización y explotación de recursos electrónicos, tecnologías de la traducción y terminología especializada en español o alemán-inglés-italiano-francés, así como en la elaboración de repertorios terminológicos electrónicos, multilingües y modulares.

Por otro lado, el proyecto ECOSISTEMA: *Espacio Único de Sistemas de Información Ontológica y Tesauro sobre el Medio Ambiente* se trata de un proyecto multilingüe (inglés, español, alemán, francés e italiano), interuniversitario e interdisciplinar, que sigue las mismas líneas innovadoras que el proyecto anterior con relación a las tecnologías de la información. Su temática principal es el medio ambiente, y dentro de ella se engloban el ecoturismo, el turismo de naturaleza y el turismo de aventura, es decir, los tipos de turismo relacionados con la naturaleza y el medio ambiente. Este proyecto presta especial referencia a la terminología, la elaboración de ontologías y la ontoterminografía de los campos de especialidad citados con anterioridad.

Además de los conocimientos adquiridos durante las investigaciones llevadas a cabo con estos proyectos punteros, que nos ha permitido profundizar en diferentes aspectos importantes de nuestra tesis doctoral, a saber: la compilación de corpus, el cálculo de la representatividad de corpus, la búsqueda y evaluación de fuentes de información, la aplicación de sistemas de información ontológica, entre otros, hemos visto completada nuestra formación con diversos seminarios y cursos específicos en Terminología, relacionados tanto con la metodología terminológica como con las líneas de investigación más actuales y las herramientas disponibles para llevar a cabo trabajos terminológicos. Por último, cabe destacar también las dos estancias realizadas en diferentes centro de investigación durante los años de duración de esta tesis doctoral y

que han posibilitado que el presente trabajo pueda optar a la mención de «Doctorado europeo», según la Universidad de Málaga: por un lado, la estancia de tres meses de duración en el año 2008 con el Research Group in Computational Linguistics, dirigido por el Dr. Ruslan Mitkov y perteneciente al Research Institute in Information and Language Processing, de la Universidad de Wolverhampton (Reino Unido); y, por otro, la estancia de dos meses de duración durante el año 2009 en el Centrum voor Vaktaal en Communicatie (CVC), de la Erasmushogeschool Brussel, en Bruselas (Bélgica), dirigido por la Dra. Rita Temmerman. Ambos periodos de investigación nos han permitido profundizar en diferentes puntos de esta tesis doctoral a través del acceso a bibliografía, proyectos anteriores llevados a cabo en el seno de estos centros, herramientas para la realización de trabajos ontoterminológicos, entre otros, además de habernos posibilitado colaborar y entrar en contacto con otros investigadores expertos y de reconocida valía que nos han aportado conocimientos nuevos y diferentes y críticas muy enriquecedoras.

Con este marco de trabajo investigador como telón de fondo, podemos pasar a especificar los fundamentos teóricos de nuestra investigación, así como la hipótesis y objetivos planteados en ella.

Esta tesis doctoral constituye una investigación interdisciplinar, ya que en ella confluyen diferentes disciplinas interrelacionadas pero diferentes entre sí. No obstante, los fundamentos principales de esta investigación se encuentran, por un lado, la rama aplicada de la traductología que consiste principalmente en el estudio de herramientas que utiliza el traductor, sus necesidades concretas y las posibilidades de mejora (Holmes (1972 [1988])⁷ y, por otro, en la terminología, concretamente en las teorías modernas de la terminología y en la rama práctica más actual relacionada con la ingeniería ontológica: la ontoterminografía. Aquí podemos encontrar tanto el punto de partida como el eje central de esta investigación, lo que permite que este trabajo sea coherente y sistemático desde el principio hasta el final. Como decimos, también están presentes otra serie de disciplinas necesarias para llevar a cabo diferentes pasos de la investigación: a saber, la Lingüística de corpus, que nos guía a la hora de compilar el

⁷ Este trabajo, «The Name and Nature of Translation Studies», fue publicado por primera vez en el Third International Congress of Applied Linguistics, celebrado en Copenhague en 1972, pero aquí citamos la edición posterior recogida en Holmes (1988).

corpus de trabajo y de realizar la gestión del corpus de forma adecuada y eficiente; la Lingüística Computacional, que nos muestra los pasos necesarios para realizar procesamiento de textos y su gestión de forma (semi)automática; la Semántica de marcos, que nos muestra las posibilidades de realizar representaciones de dominios en forma de marcos y nos indica las pautas necesarias para llevarlas a cabo; la Documentación, que nos proporciona los conocimientos necesarios para realizar búsquedas y clasificación de fuentes de información y la *Metaterminografía*,⁸ que nos permite realizar una investigación empírica sobre los diferentes tipos de recursos terminográficos existentes y elaborar una plantilla de evaluación para dichos recursos.

Nuestra tesis doctoral comienza con la siguiente hipótesis principal: los traductores profesionales necesitan unos recursos terminológicos elaborados para ellos como usuarios reales y como grupo de destinatarios específico, ya que presentan unas necesidades y unos requisitos propios y distintos de otros grupos de usuarios. Junto a esta hipótesis principal, encontramos otras hipótesis específicas, relacionadas, por un lado, con la metodología de trabajo y, por otro lado, con el trabajo práctico sobre el turismo de aventura, a saber:

1. La Ontoterminografía combinada con la Semántica de marcos (Fillmore, 1977, 1982, 1985) es una metodología adecuada para la elaboración de bases de datos terminológicas dirigidas a traductores como grupo de destinatarios.
2. Los corpus multilingües (comparables y paralelos), compilados virtualmente a través de los recursos que ofrece la red, son una valiosa

⁸ El término «metaterminografía» está aún por acuñar por los expertos en Terminografía. Hasta el momento se ha empleado en algunos trabajos (Gómez González-Jóver, 2006: 691; Casas Gómez, 2006: 27), aunque siempre con referencia a las decisiones previas que determinar el modelo de recurso terminográfico que se va a compilar. En nuestro caso, hemos considerado necesario utilizar este término como equivalente en Terminografía al término «metalexigrafía», que hace referencia a una rama de la Lexicografía que tiene como objeto de estudio el diccionario y su objetivo es investigar sobre él y no producirlos. En la metalexigrafía, se diferencian cinco campos diferentes de actuación según Hartmann (2001:30), a partir de los propuestos por Wiegand (1998: 79-80), a saber: a) historia de los diccionarios (*Dictionary History*), b) crítica de los diccionarios (*Dictionary Criticism*), c) estructura de los diccionarios (*Dictionary Structure*), d) tipología de los diccionarios (*Dictionary Tipology*) y e) uso de los diccionarios (*Dictionary Use*). En nuestro trabajo, defendemos que esta clasificación establecida en Lexicografía se extrapole a la Terminografía y que, por tanto, se emplee el término *Metaterminografía* por analogía a *Metalexigrafía* y se determinen estos mismos campos con relación a los recursos terminográficos, es decir, a los diccionarios especializados en general, incluyendo vocabularios, léxicos, bases de datos, etc.

fuelle de información terminológica y fraseológica que permite subsanar las deficiencias de los repertorios terminológicos tradicionales (diccionarios, bases de datos, etc.).

3. La creación de plantillas definatorias basadas en marcos semánticos elaborados para un campo de especialidad en cuestión permite la elaboración de definiciones sistemáticas, homogéneas y completas.
4. Los recursos terminológicos para el sector turístico en general son escasos y poco fiables y, además, el turismo de aventura carece de recursos terminológicos completos y adecuados para realizar traducciones de textos de ese dominio.
5. Las formas de búsqueda y consulta de información terminológica disponibles en la actualidad no son adecuadas para los traductores profesionales, que necesitan métodos más rápidos y ágiles de acceso a la información.
6. Un recurso terminológico que proporcione una gestión flexible por parte del traductor y que integre diferentes tipos de búsquedas, incluida la búsqueda contextual, agilizará la labor del traductor y permitirá alcanzar mejores resultados en calidad y tiempo.

En este sentido, el objetivo principal de la tesis doctoral consistirá en la propuesta de una metodología sistemática y descriptiva, que presente un enfoque lingüístico-cognitivo y que atienda a los criterios de multidimensionalidad y de dinamismo para la elaboración de recursos terminológicos adecuados para los traductores e intérpretes. Para ello, se utilizará el segmento turístico de turismo de aventura como el ámbito de especialidad sobre el que se realizará dicha propuesta. Por su parte, como objetivos derivados de este objetivo principal, surgen los siguientes objetivos secundarios de este trabajo:

1. Presentar el estado de la cuestión de la terminología y de su rama aplicada, la terminografía, y de los lenguajes de especialidad, así como de las disciplinas lingüísticas aplicables a la terminología.

2. Revisar las fuentes de información y proyectos dirigidos a traductores, especialmente aquellos relacionados con la ontoterminografía.
3. Clasificar las fuentes de información terminológicas que se pueden encontrar en la actualidad.
4. Proponer un protocolo objetivo y cuantificable para la evaluación de recursos terminológicos en formato papel y electrónico.
5. Definir y concretar las necesidades profesionales del traductor como usuario real de un recurso terminológico.
6. Especificar criterios adecuados para el diseño de corpus especializados para el trabajo terminográfico.
7. Elaborar una base de datos multilingüe (español, inglés y alemán) para apoyar las necesidades de los traductores a la hora de realizar traducciones de material promocional para el segmento del turismo de aventura según la metodología propuesta como objetivo principal.
8. Recopilar la terminología propia del turismo de aventura y organizarla con respecto a las diferentes variantes denominativas, neologismos, extranjerismos, etc. propios de este campo.
9. Utilizar los marcos semánticos, en sus diferentes formas y distintos niveles, como medio de representación de base cognitiva más eficiente: desde la macroestructura del dominio hasta la microestructura.
10. Proponer una nueva forma de búsqueda y consulta de información terminológica a través del diseño de una herramienta innovadora, eficiente y acorde con las necesidades de los traductores.
11. Evaluar la herramienta diseñada para la búsqueda y consulta terminológica en la presente tesis doctoral a través de un experimento empírico con traductores semiprofesionales.
12. Analizar las opiniones de los sujetos del experimento a través de datos cuantitativos y cualitativos.

Además, el punto 7 anterior implica una serie de objetivos propios, que enumeramos a continuación:

- 7.1. Analizar el sector turístico en España desde diferentes puntos de vista (económico, traductológico y lingüístico) y, más específicamente, estudio del segmento del turismo de aventura y sus principales características.
- 7.2. Seleccionar la herramienta para llevar a cabo el trabajo ontoterminográfico según la metodología propuesta.
- 7.3. Analizar recursos terminológicos del segmento del turismo de aventura en papel y en formato electrónico según el protocolo de evaluación propuesto.
- 7.4. Compilar un corpus electrónico formado por documentos originales (comparable) y originales y sus traducciones (paralelo) de género promocional en español, inglés y alemán sobre turismo de aventura.
- 7.5. Crear un esquema categorial del turismo de aventura y, posteriormente, de una ontología de dominio en forma de marco semántico.
- 7.6. Elaborar una base de datos multilingüe (español, inglés y alemán) conforme a la norma ISO 12620 (1999).

Para la consecución de los objetivos anteriores, hemos dividido el presente trabajo en cuatro capítulos, a los cuales les precede un resumen inicial⁹ y les siguen tres apartados finales: conclusiones, referencias bibliográficas y anexos.

En el primer capítulo, que se titula «La terminología moderna y los nuevos enfoques del trabajo terminográfico», se lleva a cabo una revisión exhaustiva del estado de la cuestión en terminología con respecto a la terminología tradicional y las disciplinas lingüísticas afines, como son la Lingüística de Corpus, la Lingüística Computacional y la Lingüística Cognitiva, y, además, se establece un análisis de las

⁹ De acuerdo con la normativa establecida en la Universidad de Málaga para la obtención de la mención de «Doctorado Europeo» (aprobado en Consejo de Gobierno Provisional de 5 de mayo de 2004, con las modificaciones introducidas por acuerdo del Consejo de Gobierno del 11 de julio de 2005), es necesario que «parte de la Tesis Doctoral, al menos resumen y las conclusiones, se habrá de redactar y presentar en una de las lenguas oficiales de la Unión Europea distinta a alguna de las lenguas oficiales en España». Por ello, el resumen inicial, así como las conclusiones de este trabajo de investigación aparecen redactados en lengua inglesa.

dicotomías más frecuentes en este contexto lingüístico, a saber: lengua general/lenguaje de especialidad, terminología/lexicología y terminografía/lexicografía. A continuación, se exponen las líneas de trabajo más actuales en terminografía, junto con sus principales características, a saber: la terminografía basada en corpus y la terminografía basada en la representación de conocimiento. Por último, se presentan una serie de proyectos de investigación que combinan estas dos líneas de trabajo terminográfico y se comentan sus características y resultados.

El segundo capítulo, «El trabajo ontoterminográfico basado en corpus dirigido a traductores», se ocupa de la revisión de las competencias traductoras necesarias para llevar a cabo una traducción adecuada de textos especializados, centrándose principalmente en la competencia documental, además de abordar la evaluación de las fuentes de información disponibles para los traductores en la actualidad y las necesidades y requisitos por parte de estos a través de una encuesta *ad hoc* elaborada durante el verano de 2008. Por último, presenta una propuesta metodológica para la elaboración de trabajos ontoterminográficos que mejoren los recursos terminológicos disponibles en la actualidad y que, a la par, satisfagan las necesidades de los traductores, según los resultados obtenidos en la encuesta realizada.

Tras el establecimiento de una propuesta metodológica para llevar a cabo un trabajo ontoterminográfico, en el capítulo tres, «Preparación del trabajo ontoterminográfico», se presenta la puesta en práctica de dicha metodología en un dominio de especialidad concreto: el turismo de aventura. Concretamente, en este capítulo se llevan a cabo las dos primeras fases de la metodología, es decir, los pasos previos al inicio del trabajo ontoterminográfico *per se*. Estas dos fases son las siguientes: 1. Delimitación del trabajo y análisis del dominio y 2. Compilación del corpus especializado. A su vez, cada fase está dividida en diferentes subfases, cuya consecución va avanzando hacia las siguientes fases del trabajo. Con respecto a la primera fase, es necesario llevar a cabo una serie de pasos, a saber: se deben tomar una serie de decisiones acerca de las diferentes variables pragmático-lingüísticas del proyecto (destinatarios, lenguas de trabajo, etc.) y acerca de los recursos tecnológicos y humanos según capacidad económica y de tiempo, con objeto de delimitar el proyecto y establecer unos límites en el trabajo; a continuación, se lleva a cabo un análisis del

sector turístico en general y del segmento del turismo de aventura en particular según diferentes aspectos (económico, lingüístico y traductológico) y se realiza un estudio de los recursos terminológicos disponibles en el dominio seleccionado y en las lenguas de trabajo (español, inglés y alemán), a fin de aproximarse a las necesidades y características propias del segmento del turismo de aventura. Por último, se elabora una conceptualización inicial del dominio del turismo de aventura en forma de marco semántico, que servirá de base para las siguientes fases. La segunda fase de la metodología, la compilación de corpus de trabajo, se divide en tres subfases, según los pasos que hay que seguir para compilar un corpus, a saber: se debe comenzar con el diseño del corpus, es decir, se deben establecer unos criterios internos y externos para conseguir una homogeneización en los resultados de la búsqueda y, así, buscar la representatividad final del corpus. Por último, es necesario organizar los textos seleccionados, a fin de establecer un registro de los textos seleccionados y una correspondencia entre el documento físico y el registro.

El cuarto capítulo, «Realización del trabajo ontoterminográfico y su edición», completa la metodología propuesta en el segundo capítulo con el resto de las fases establecidas, a saber: 3. Gestión ontoterminográfica, 4. Elaboración de una base de datos ontoterminográfica, 5. Revisión y validación y 6. Exportación y edición. En estas fases se lleva a cabo el trabajo ontoterminográfico *per se*, puesto que las fases anteriores se consideran como fases preparatorias. Concretamente, las fases que presentan mayor complejidad y exigen gran esfuerzo son las tercera y cuarta fases, puesto que en ellas se llevan a cabo las tareas más exigentes, que van desde la extracción de candidatos a término, equivalentes y relaciones semánticas hasta la introducción de información en los campos terminológicos de las entradas, incluida la elaboración de definiciones basadas en plantillas definatorias según el marco semántico del dominio. Finalmente, después de haber llevado a cabo estas dos fases, se procede a una revisión y validación del trabajo realizado con objeto de detectar cualquier error cometido o falta de información y se pasa a la edición del trabajo. En nuestro caso, en la fase de edición, al contrario que ocurre en otros proyectos terminológicos, hemos pretendido ajustarnos a las necesidades de nuestros usuarios finales, los traductores, y hemos diseñado e implementado una aplicación informática integradora, flexible y novedosa a fin de satisfacer dichas necesidades. Así, partiendo de las herramientas comentadas en otros

apartados de esta tesis doctoral dirigidas a traductores y tras una sucinta revisión de aplicaciones informáticas anteriores a la herramienta que proponemos, presentamos una descripción completa de *Trandix*, la aplicación desarrollada en el seno de esta tesis doctoral, así como de sus diferentes funcionalidades y sus principales ventajas frente a las demás herramientas. Por último, con objeto de comprobar la validez de esta herramienta y conocer la opinión de sus futuros usuarios, hemos procedido a evaluarla a través de un experimento, cuya descripción y resultados se presentan al final de este capítulo.

Tras estos cuatro capítulos se incluye un apartado de conclusiones, donde se recogen las principales ideas desarrolladas y los resultados obtenidos a raíz de la investigación llevada a cabo durante el desarrollo de esta tesis doctoral, así como las posibles líneas de investigación futuras. A continuación, se ofrecen las referencias bibliográficas consultadas, ordenadas de forma alfabética, y un apartado formado por diferentes anexos, a saber: el Anexo I muestra la versión española de la encuesta llevada cabo para conocer las necesidades de los traductores profesionales durante el verano de 2008; el Anexo II hace referencia a la plantilla de evaluación elaborada de forma *ad hoc* para evaluar recursos terminológicos tanto en formato papel como electrónico; el Anexo III indica las asociaciones de traductores profesionales que se contactaron para distribuir la encuesta *ad hoc*; el Anexo IV presenta una versión breve del manual de usuario elaborado para la herramienta de búsqueda y consulta diseñada en el seno de esta tesis doctoral; el Anexo V ofrece la clasificación de los recursos terminológicos para el sector turístico según los diferentes criterios de clasificación aplicados para su catalogación, así como la evaluación de todas las fuentes, y, por último, el Anexo VI incluye la documentación elaborada para la puesta en práctica del experimento realizado a los traductores *semiprofesionales* para llevar a cabo la evaluación de la aplicación *Trandix*.

Por último, el volumen impreso se acompaña de un CD-ROM que incluye el presente trabajo en formato digital y un vídeo con la presentación de la herramienta *Trandix*.

CAPÍTULO 1.

LA TERMINOLOGÍA MODERNA Y LOS NUEVOS ENFOQUES

DEL TRABAJO TERMINOGRÁFICO

ÍNDICE DE CONTENIDOS

- 1.1. Introducción
- 1.2. La dicotomía entre lengua general y lenguaje de especialidad
 - 1.2.1. Lenguaje de especialidad
 - 1.2.2. Caracterización de los lenguajes de especialidad
 - 1.2.2.1. Plano lingüístico
 - 1.2.2.2. Plano pragmático
 - 1.2.2.3. Plano funcional
- 1.3. La dicotomía terminología - lexicología
 - 1.3.1. El concepto *término*
 - 1.3.2. Corrientes teóricas de la Terminología
 - 1.3.2.1. La corriente clásica de la Terminología (TGT)
 - 1.3.2.2. Las corrientes modernas de la Terminología
 - 1.3.2.2.1. La Socioterminología
 - 1.3.2.2.2. La Teoría Comunicativa de la Terminología (TCT)
 - 1.3.2.2.3. La Teoría Sociocognitiva de la Terminología (TST)
 - 1.3.2.3. Corrientes teóricas de la lingüística aplicables a la Terminología
 - 1.3.2.3.1. La Lingüística de Corpus
 - 1.3.2.3.2. La Lingüística Computacional
 - 1.3.2.3.3. La Lingüística Cognitiva
 - 1.3.2.3.3.1. La Semántica de Marcos
- 1.4. La dicotomía terminografía - lexicografía
 - 1.4.1. Aproximación de la terminografía a la lexicografía especializada
 - 1.4.2. La terminografía basada en corpus
 - 1.4.2.1. La Lexicografía basada en corpus
 - 1.4.2.2. Algunas consideraciones en torno al corpus: concepto y definición
 - 1.4.2.2.1. Criterios generales para la compilación de corpus
 - 1.4.2.2.2. Tipos de corpus
 - 1.4.2.3. Criterios específicos para la elaboración y diseño de los corpus especializados para la terminografía
 - 1.4.2.4. Herramientas de análisis de corpus: necesidades específicas de los terminógrafos
 - 1.4.2.5. Ventajas e inconvenientes del empleo de corpus en terminografía
 - 1.4.3. La terminografía basada en la representación de conocimiento
 - 1.4.3.1. La representación de conocimiento especializado en la TGT y otros modelos de representación
 - 1.4.3.2. Algunas consideraciones en torno a las ontologías: definición y componentes
 - 1.4.3.2.1. Criterios generales para la creación de ontologías
 - 1.4.3.2.2. Clasificaciones de ontologías
 - 1.4.3.3. Criterios específicos para el diseño de ontologías en terminografía
 - 1.4.3.4. Herramientas para la elaboración y gestión de ontologías
 - 1.4.3.5. Ventajas e inconvenientes del empleo de ontologías en terminografía
- 1.5. Proyectos basados en la representación del conocimiento especializado mediante corpus dirigidos a traductores
 - 1.5.1. Proyecto COGNITERM
 - 1.5.2. Proyecto ONCOTERM
 - 1.5.3. Proyecto GENOMA-KB

La Terminología, en cuanto materia que concierne a la recopilación, descripción, tratamiento y presentación de los términos propios de los campos especializados en una o más lenguas, no es una actividad práctica que se justifique por sí sola, sino que está destinada a resolver necesidades sociales vinculadas a la optimización de la comunicación entre especialistas y profesionales –ya sea directamente o mediante la traducción– [...] (Cabré Castellví: 1993: 35-36).

1.1. INTRODUCCIÓN

Este capítulo pretende establecer los fundamentos teóricos de los que parte nuestra investigación y, para ello, se ha llevado a cabo un estudio profundo sobre el estado actual en el que se encuentra la Terminología como disciplina, tanto en su rama teórica como en su rama aplicada, la Terminografía. Asimismo, se ofrece una visión general desde la terminología tradicional hasta la terminología moderna, con objeto de mostrar la evolución que ha sufrido esta disciplina desde sus comienzos.

En primer lugar, comenzamos con el tratamiento de la rama teórica, la Terminología, ofreciendo una revisión de los diferentes elementos que giran en torno a ella y que a menudo son objeto de ambigüedad, a saber: *lengua general-lenguaje de especialidad*, *palabra-término*, *Lexicología-Terminología*, etc., con objeto de plasmar el

estado actual de esta cuestión y de posicionarnos según nuestro punto de vista. Además de analizar la dicotomía general-especializado, y sus consecuentes binomios, realizaremos un breve repaso a la evolución de la Terminología como disciplina teórica en los últimos treinta años, desde la Teoría General de la Terminología hasta las teorías más modernas, presentando así sus postulados más actuales. La revisión de estos conceptos y postulados nos servirá para establecer la base teórica sobre las que se sustentan las premisas y líneas de actuación de este trabajo de investigación. Asimismo, para terminar esta parte teórica, nos parece oportuno recoger los aspectos más relevantes de las disciplinas lingüísticas que han influido en la Terminología tal y como la conocemos actualmente y, por ello, trataremos la Lingüística Cognitiva, la Lingüística de Corpus y, por último, la Lingüística Computacional.

Una vez expuestas las diferentes corrientes teóricas de la Terminología, así como sus disciplinas lingüísticas cercanas, continuaremos con la parte aplicada, es decir, la Terminografía, pero deteniéndonos en la metodología de las nuevas corrientes terminológicas: la terminografía basada en corpus y la terminografía basada en la representación del conocimiento, y algunos de los proyectos actuales llevados a cabo en estas líneas de investigación.

Como veremos, la combinación de la terminología descriptiva basada en el conocimiento (Meyer y Mackintosh, 1996) y la Lingüística de Corpus ofrece una base idónea para el trabajo terminográfico y permite alcanzar, además, resultados óptimos en los diferentes tipos de productos, ya que asiste al terminógrafo en todas las fases de su tarea y reduce el aspecto subjetivo de cualquier trabajo terminográfico. Además, la concepción introducida por la Teoría Comunicativa de la Terminología con respecto a la elaboración de aplicaciones terminográficas, mediante la cual se defiende una metodología de trabajo «configurable por el terminógrafo» (Gómez González-Jover, 2006: 113) según el usuario prototípico del recurso y los objetivos que se persigan, proporciona un valor añadido a cualquier producto terminográfico, ya que le permite adaptar su resultado a unos destinatarios potenciales con unas características determinadas.

Tanto la Terminología moderna como estas nuevas corrientes metodológicas basadas en la Lingüística del Corpus, la Lingüística Computacional y la Semántica de

Marcos ocuparán un lugar muy relevante en la consecución de los objetivos de este trabajo de investigación, como se podrá observar en la propuesta metodológica confeccionada dentro de esta investigación (cf. 2.6.1.).

1.2. LA DICOTOMÍA ENTRE LA LENGUA GENERAL Y EL LENGUAJE DE ESPECIALIDAD

Como señala Pérez Hernández (2000: 124), la dicotomía planteada entre lengua general y lenguaje de especialidad por un lado, y entre términos y palabras por otro, parece lógica si seguimos lo que Cabré Castellví denomina «planteamientos wüsterianos» (Cabré Castellví 1999b: 111), es decir, los fundamentos de la Teoría General de la Terminológica (TGT) (cf. 1.3.2.1.). Como veremos, en el seno de esta teoría los términos son considerados unidades específicas de un ámbito de especialidad concreto y su uso queda circunscrito a dicho ámbito, por lo que la Terminología solo debería interesarse por la comunicación entre expertos. Algunas de las críticas realizadas a esta teoría terminológica tradicional, recogidas en Cabré Castellví (1999b: 115) destacan la irrealidad de dichos planteamientos. El conocimiento especializado no puede guardarse en compartimentos estancos, cada uno con su terminología propia, ya que el saber es un continuo, y su segmentación en materias es puramente metodológica que en ningún momento responde a una segmentación en el conocimiento. La demarcación de diferencias especiales que definen los lenguajes de especialidad frente a la lengua general es una cuestión todavía hoy bastante polémica y difusa (Cabré Castellví, 1993: 132).

En esta misma línea, autores como Alcaraz Varó (2000) y Pérez Hernández (2002), entre otros, afirman que no es posible definir los lenguajes de especialidad en términos estrictamente lingüísticos, como pretendía Wüster, puesto que lo que los diferencia de la lengua general son las características pragmáticas y extralingüísticas que se dan en cada situación comunicativa, es decir, no es posible abstraer los lenguajes de especialidad del acto comunicativo, ya que de él dependen las características de este lenguaje en cada momento. En este sentido, la intención de cada situación comunicativa determinará el tipo de registro que se utilice en cada momento y, por tanto, habrá que

tenerla siempre muy presente. Esta intención vendrá determinada, a su vez, por las convenciones sociales que dictan las pautas de comportamiento para cada tipo de situación.

Según el DRAE, la lengua es el instrumento de comunicación (verbal y escrito) de una comunidad humana. Aunque no se trata del único instrumento para la comunicación (por ejemplo, la comunicación gestual), sí es considerado el más universal:

La lengua (las lenguas naturales) [es] el medio universal de la comunicación, cuyos signos pueden representar todos los contenidos cognitivos conceptuales del hombre y que abarca con esto todos los campos necesarios de la comunicación humana (Jäger, 1975: 13).¹⁰

En este sentido, la lengua es el medio que hace posible la comunicación entre una comunidad determinada de hablantes, ya que se trata de un sistema o conjunto de signos orales o escritos convencionales que sirven a una misma comunidad lingüística como medio de expresión y comunicación (Gómez González-Jover, 2006: 32). Como medio de comunicación, la lengua se utiliza en cualquier situación de la vida de los hablantes de la comunidad, tanto en situaciones cotidianas como en situaciones profesionales y en contextos especializados.

Al primer tipo de actuación se le ha denominado *lengua general*, mientras que al segundo se le ha identificado como *lenguaje de especialidad*, que se define como el «lenguaje específico que utilizan algunos profesionales y especialistas para transmitir información y para negociar los términos, los conceptos y los conocimientos de una determinada área de conocimientos, a saber, confirmar los existentes, matizar el ámbito de su aplicación y modificarlos total o parcialmente» (Alcaraz Varó, 2000: 15). A modo de ejemplo, al hablar del alemán del turismo o jurídico hacemos referencia al uso que hace ese idioma para exponer conocimientos específicos de esas áreas de especialidad y sus connotaciones.

Se debe poner de manifiesto que la lengua general es común a la mayoría de los hablantes de una comunidad y, por tanto, la emplean como medio en sus situaciones comunicativas. Sin embargo, los lenguajes especializados hacen referencia a la esfera de

¹⁰ Citado por Jung (2000:13).

especialización y, por tanto, no todos los hablantes de una misma comunidad conocen sus características lingüísticas y no lingüísticas particulares, que deben emplearse en la comunicación especializada. De esta forma, al hablar de lenguajes especializados en contraposición a la lengua común, se hace referencia a menudo también a la dicotomía conocimiento general-conocimiento especializado de sus usuarios.

En este sentido, podemos afirmar que la lengua general está constituida por un conjunto de subcódigos que los hablantes emplean, según las necesidades expresivas y las características particulares del contexto comunicativo en que se encuentran (Montero Martínez, 2002). Por tanto, para justificar la existencia de distintos tipos de actos comunicativos (lengua general vs. lenguajes de especialidad) tendremos que considerar aspectos como la función o el uso de los mismos en cada situación comunicativa, los interlocutores (emisor o receptor), la intención y el registro, que permite identificar la situación pragmática en la que se desarrolla un acto de habla, sin olvidar la función específica que los caracteriza.

1.2.1. LENGUAJE DE ESPECIALIDAD

Como veíamos en el apartado anterior, hablamos de lenguajes de especialidad cuando nos referimos al lenguaje que se utiliza en una situación comunicativa concreta en la que se transmite conocimiento especializado; por el contrario, nos referimos a lengua general para hablar de una comunicación de conocimientos generales. En esta línea, se utilizan nociones como ámbito o dominio especializado o de especialidad para hacer referencia a la parcela de saber determinada a la que pertenece un lenguaje de especialidad en concreto (en el ejemplo anterior, el ámbito de especialidad sería el derecho o el turismo), o se habla de lenguajes especializados para referirse a los recursos que se utilizan en situaciones comunicativas específicas y sobre temas específicos (Cabré Castellví, 2004: 20).

Sin embargo, no existe actualmente un consenso generalizado sobre la denominación de este tipo de lenguaje y, ni siquiera, sobre el uso del término *lenguaje* aplicado a los recursos utilizados en situaciones especializadas. Asimismo, existe hoy

en día una falta absoluta de consenso sobre la denominación que debería recibir esta realidad.

Así pues, podemos encontrar diversos términos para designar el concepto de «lenguaje de especialidad», como lenguas especializadas, lenguas de especialidad, lenguajes especializados, lenguas especiales, lenguajes especiales o lenguajes para fines (o propósitos específicos), cuyos significados se emplean de forma indistinta (Cabré Castellví, 1993: 135). Para algunos autores, como Kocourek (1982) o Rondeau (1983), los términos anteriores son todos sinónimos entre sí. Para otros autores, como Sager et ál. (1980), el término «lenguaje para fines o con propósitos específicos» (del inglés, language for special purposes o specific purposes, LSP), se circunscribe a la enseñanza de idiomas, mientras que el término «lenguajes especiales» (special languages) lo conciben con el sentido de «lenguaje especializados por la temática». Según estos autores, el uso de estos lenguajes especiales presupone una formación especializada y se limita a una comunicación entre especialistas que trabajan en el mismo ámbito o en ámbitos afines (Gómez González-Jover, 2006: 50).

Otros autores, como Lerat (1997), prefieren el término «lengua de especialidad», para evitar el uso del término «lenguaje» y así todo lo referente al estilo. Asimismo, Alcaraz Varó (2000) se ha referido a este concepto con una denominación nueva y más específica: «lenguajes profesionales o académicos». Con este término, pretende sustituir el término «lenguaje para fines específicos», ya que considera que cualquier uso que se haga del lenguaje siempre será «específico».

Otras variantes que se utilizan para denotar este concepto son «tecnolecto» (del alemán, *Technolekt*) (Realiter, 1996: 1) o «lengua en especialidad», aunque no han recibido una gran acogida en el ámbito académico y su utilización es escasa.

Además de todas estas posibilidades para conceptualizar la misma realidad, encontramos el término «discurso especializado» o «discurso de especialidad», usados como sinónimos (Rey, 1976; Kocourek, 1982; Condamines, 1995). En esta aproximación, se defiende el uso de «discurso» en vez de «lenguaje», por considerar este último «un uso abusivo del término ‘lenguaje’» (Cabré Castellví, 2004: 20). Un lenguaje, como sinónimo de una lengua, presupone un conjunto completo y autónomo de unidades y de reglas o condiciones que describen las producciones emitidas en una

situación de uso. Así pues, requiere siempre de una gramática propia, que incluye una fonología, una morfología, una sintaxis y un léxico. Por tanto, no sería correcto el empleo de «lenguaje», a no ser que se considerara en un sentido metafórico, «al igual que usamos lenguaje para referirnos al lenguaje de las flores o de las abejas» (ibíd.: 20), ya que las producciones lingüísticas en situaciones de especialidad proceden del conjunto de recursos que la lengua general ofrece para expresarse y comunicarse. En este sentido, solo la situación comunicativa determinará que se haga uso de una parte de estos recursos. Desde este punto de vista, los «lenguajes de especialidad» son registros funcionales caracterizados por una temática específica conceptualizada de forma determinada que se dan en situaciones comunicativas en las que el emisor cuenta con un conocimiento especializado y la función principal es la informativa o referencial. De esta manera, sería más adecuado hablar de usos o discursos especializados dentro de una lengua general, ya que, a pesar de tratarse de una comunicación especializada y diferente de la comunicación general, utiliza todos sus recursos.

Así, deberemos entender que el empleo de «lenguajes de especialidad» tiene un sentido metafórico, ya que según el significado real de «lenguaje» no sería adecuado su uso y que, por tanto, el término «discurso especializado» se adecua mejor a la realidad o, incluso, como se viene denominando últimamente «comunicación especializada». En nuestro caso, utilizaremos «lenguaje de especialidad» o «lenguaje especializado» como sinónimos y con sentido metafórico, por ser la forma más extendida entre los autores y expertos, y por considerarlo la forma más apropiada para esta realidad.

La problemática con la denominación de este concepto no es lo único que encontramos, sino que además se extiende a su definición y su delimitación, como veremos a continuación.

No existe una definición única con referencia a qué son los lenguajes de especialidad, ya que se han aportado diferentes definiciones de estos a lo largo de la historia de la Terminología,¹¹ pero sí que se pueden extraer ciertas características que comparten la mayoría de los lenguajes especializados y con los que están de acuerdo la gran parte de los autores. En la siguiente cita se resumen sus características principales:

¹¹ En Cabré Castellví (1993), se puede consultar un amplio estudio sobre la evolución de la Terminología y sus diferentes aproximaciones.

[Se consideran los lenguajes especializados como] productos predominantemente verbales, de registros comunicativos específicos, que tratan temas propios de un ámbito de especialidad, que respetan convenciones y tradiciones retórico-estilísticas y que dan lugar a clases textuales determinadas (Cabré Castellví, 2002: 22).

Además, continúa diciendo la autora (ibíd.: 23), utilizan recursos propios de la lengua común, a pesar de presentar características léxicas, sintácticas, morfológicas y gráficas determinadas que los diferencian de otros discursos especializados.

Aunque esta definición parece en principio delimitar con claridad el ámbito de la lengua general y el de los lenguajes de especialidad, la misma autora (ibíd.: 132) reconoce que la demarcación de esas peculiaridades especiales que definen los lenguajes de especialidad así como su relación con la lengua general es «una cuestión [...] todavía hoy bastante polémica y difusa». Asimismo, como nos indica Arribas Baño (2004), el énfasis en establecer una diferencia clara entre lo general y lo especializado reside en su indiscutible utilidad metodológica y pragmática. Sin embargo, hasta ahora no se han podido establecer unos criterios objetivos para limitar la frontera entre la lengua común y los lenguajes especializados, debido a los límites difusos que existen entre estas dos realidades.

Ahmad et ál. (1995) resumen las diferentes posturas existentes en lo que concierne a la relación entre lengua natural y lenguajes de especialidad en tres modelos diferentes representados en la figura siguiente:

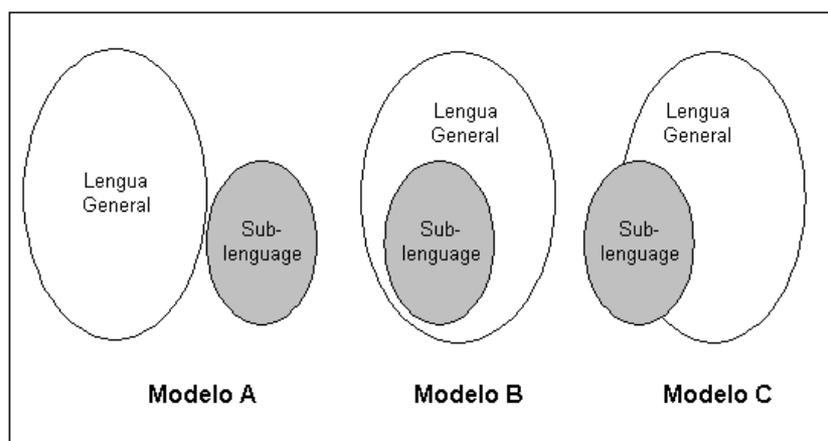


Figura 1. Relación entre lengua general y lenguajes de especialidad.

Los tres modelos propuestos se pueden considerar semejantes a las tres aproximaciones realizadas por Bergholtz y Tarp (1995: 17 y ss.), Cabré Castellví (1993: 132 y ss.) y Beaugrande (1987):

- a) La primera de las aproximaciones (Modelo A) considera los lenguajes especializados como códigos lingüístico completos diferenciados de la lengua general, ya que constan de características morfosintácticas y construcciones sintácticas propias, así como de una terminología transparente, monosémica y unívoca (Hoffman, 1979: 16). Sin embargo, no parece fácil delimitar la frontera entre la lengua general y los lenguajes especializados, ni parece existir consenso sobre cuáles son los fenómenos lingüísticos que los diferencian.
- b) La segunda (Modelo B) considera los lenguajes de especialidad como simples variantes del lenguaje general y reduce la especificidad del lenguaje de especialidad al léxico (Rondeau, 1983; Mounin, 1979; Rey, 1976).
- c) La tercera teoría (Modelo C) considera que la relación entre la lengua general y los lenguajes de especialidad deben considerarse como la unión de un subconjunto de la lengua general con un grupo de elementos específicos en sí mismos (Deville, 1989). Además, Deville y Herbigniaux (1995: 143-144) apuntan que los elementos que forman parte del sublenguaje se consideran como un conjunto de elementos ordenados según la situación comunicativa y el dominio de aplicación, y, que el carácter restrictivo de un sublenguaje no es una exclusión sino una jerarquía de preferencia.

Una cuarta postura, cercana a la tercera aproximación (Modelo C) es la que proponen aquellos lingüistas que consideran los lenguajes especializados como subconjuntos (esencialmente pragmáticos) de una lengua global,¹² que englobaría tanto la lengua común como todos los posibles lenguajes especializados. Según esta última posición teórica, resultaría imposible definir los lenguajes de especialidad en términos

¹² Término propuesto por Lehrberger (1986: 22) y también utilizado por Hoffman (1998:48).

estrictamente lingüísticos, ya que lo que los diferencia de la lengua general son características pragmáticas o extralingüísticas, es decir, estos lenguajes de especialidad se consideran variantes pragmáticas con una serie de peculiaridades “especiales”: temática, situación de comunicación, tipo de interlocutores, medio, nivel de formalidad, etc. que los diferencian entre sí y de la lengua general.

Autores como Sager et ál. (1980), Picht y Draskau (1985), Varantola (1986) y Cabré Castellví (1993) defienden esta última aproximación de dependencia entre los lenguajes de especialidad y la lengua común. Según Cabré Castellví (1993: 140):

Los lenguajes de especialidad están en relación de inclusión respecto del lenguaje general y en relación de intersección respecto de la lengua común, con la que comparten características y con la que mantienen una relación de trasvase constante de unidades y convenciones.

De la misma manera, Sager et ál. (1980) establecen lo siguiente:

The nature of language is such that general language and special languages can be accommodated within one natural language: the fundamental characteristics of language are manifested both in English and in the language of chemical engineering, both in French and in the language of physics. The difference between general and special languages is a difference of degree rather than kind: the degree to which the fundamental characteristics of language are maximized or minimized in special languages. Special languages are used more self-consciously than general language and the situations in which they are used intensify the user's concern with the language. It is therefore on the level of use that we look for more specific differentiating criteria.

La Figura 2 muestra esta aproximación teórica, adaptada de Cabré Castellví (1993: 140), que completa el modelo C de la Figura 1:

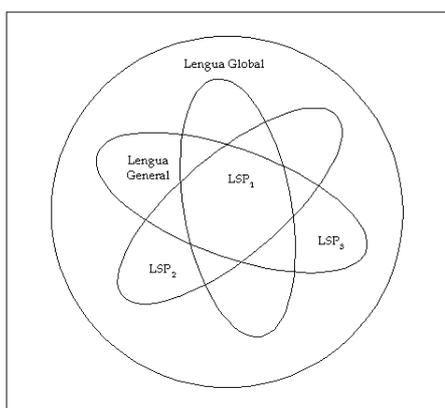


Figura 2. Relación entre lengua general, lengua común y lenguajes de especialidad.

En esta figura se observa cómo los lenguajes de especialidad están en relación de inclusión respecto a la lengua global y en relación de intersección tanto entre sí como con respecto a la lengua general, con la que comparten características y donde se produce un constante transvase de unidades. Este constante transvase de unidades es la causa fundamental de la dificultad que existe para separar las unidades léxicas de la lengua común de las unidades terminológicas de los lenguajes de especialidad.

Nuestra postura frente a la naturaleza de los lenguajes de especialidad es compartida con esta última aproximación teórica, ya que para nosotros, los lenguajes de especialidad son también subsistemas funcionales dentro de la lengua general. De este modo, y a modo de resumen de lo dicho anteriormente, la lengua global se establece como el sistema que engloba todos los lenguajes de especialidad, los cuales son considerados como subsistemas funcionales. Y, con respecto al lenguaje de especialidad objeto de estudio en esta tesis doctoral, el lenguaje del turismo, defenderemos, por tanto, su lugar como subsistema funcional dentro de la lengua general y al mismo nivel que cualquier otro lenguaje de especialidad.

1.2.2. CARACTERIZACIÓN DE LOS LENGUAJES DE ESPECIALIDAD

Como veíamos en el anterior apartado, no existe una definición acuñada sobre qué son los lenguajes de especialidad, sino una serie de características comunes a la mayoría de estos lenguajes que permiten su caracterización y, en cierta medida, también su definición. No obstante, como señala Cabré Castellví (2004b: 21):

Todos coinciden en que cuando se habla de lenguajes de especialidad se trata de producciones lingüísticas sobre temas relacionados con el saber descriptivo o procedural, sobre parcelas específicas del conocimiento. Desde este punto de vista, todas las áreas científicas, técnicas o profesionales serían ámbitos de comunicación especializada.

En este contexto, se confirma la naturaleza pragmática del concepto de lenguaje especialidad, que se distingue de la lengua común y de otros lenguajes de especialidad por producirse en un contexto comunicativo determinado. Además de las características

pragmáticas que determinan los lenguajes de especialidad según los elementos específicos del proceso de comunicación (tema, usuarios y situación), estos presentan una serie de peculiaridades lingüísticas y funcionales que le confieren cierta especificidad dentro los discursos especializados (Cabré Castellví et ál., 2001: 181).

A continuación, trataremos estos tres aspectos: plano lingüístico, plano pragmático y plano funcional, a fin de profundizar en la caracterización de los lenguajes de especialidad.¹³

1.2.2.1. Plano lingüístico

En el plano lingüístico, todos los expertos coinciden en afirmar que la terminología específica utilizada en cada lenguaje de especialidad es el elemento diferenciador entre los lenguajes de especialidad entre sí y con la lengua general, ya que es el encargado de la representación de la realidad (Alcaraz Varó, 2000: 24).

De esta manera, en el ámbito del léxico se presentan numerosos elementos específicos, muchas veces exclusivos, de las diferentes variedades de especialidad. Esos elementos específicos suelen manifestarse en forma de terminologías, es decir, unidades léxicas determinadas y específicas de un campo de conocimiento especializado particular. Estas terminologías son los elementos que le otorgan mayor especificidad y diferenciación al lenguaje especializado, por lo que en el plano lingüístico ocupan un lugar predominante con respecto a otros rasgos. Junto con estas unidades propias de cada lenguaje especializado, encontramos otras características léxicas importantes, como son la presencia de neologismos, extranjerismos (calcos o préstamos), cultismos, elementos no lingüísticos (símbolos, ecuaciones, etc.), entre otros.

Además de la terminología propia de cada lenguaje de especialidad, a veces se producen trasvases entre diferentes lenguajes y entre la lengua general y el lenguaje de especialidad en cuestión, lo que confirma la intersección de los lenguajes de especialidad con otros lenguajes próximos que veíamos con anterioridad. Por ejemplo, en el caso del turismo, el trasvase de terminología del campo del arte o la arquitectura al

¹³ En el capítulo 3, se puede consultar la caracterización realizada al discurso turístico según estos tres planos.

discurso turismo es muy frecuente, debido a que se tratan de disciplinas cercanas a algunos aspectos tan importantes en el turismo como son la descripción de monumentos, de cuadros, de edificios, etc.

Frente a lo que ocurre en el plano léxico, los lenguajes de especialidad se caracterizan por utilizar, en términos generales, la gramática de la lengua común, matizada por ciertos usos que pueden destacar cualitativa o cuantitativamente. Así pues, es frecuente que aparezcan formas de tratamiento específicas o habituales en ciertos ámbitos profesionales (como por ejemplo, para el ámbito jurídico, en formas como *señoría* o *letrado* o, para el ámbito académico, en formas como *profesor* o *doctor*); también es frecuente el uso específico de ciertas formas verbales (el futuro de subjuntivo o el gerundio, en la lengua jurídico-administrativo: *si no compareciere*, *alegación solicitando revisión de pruebas*), la presencia del plural mayestático en los artículos de investigación, la abundante nominalización (*La preparación industrial del monóxido de carbono se lleva a cabo mediante la oxidación del carbón por el aire o el agua*), la presencia de pasivas reflejas y oraciones impersonales, entre otros muchos recursos sintácticos y morfológicos. En resumen, los lenguajes de especialidad, al ser utilizados en contextos específicos, emplean rasgos propios de la lengua general de forma específica en sus contextos comunicativos, lo que les confiere cierta diferencia con el uso que se realiza de ellos en la lengua común o, incluso, en otros lenguajes de especialidad. Por lo tanto, la base gramatical de los lenguajes de especialidad es compartida con la lengua general, como ocurre también con la base fonético-fonológica y prosódica, aunque con ciertas precisiones pertinentes.

1.2.2.2. Plano pragmático

El plano pragmático es el plano que ha dado lugar a la caracterización de los lenguajes de especialidad más aceptada por los expertos en este tema. En este plano, se establecen dos ejes diferentes: por un lado, un eje horizontal, determinado por la temática; y por otro, un eje vertical, determinado por el grado de especialización o abstracción que puede presentar el contenido (Cabré Castellví, 2002: 30-31).

Por el criterio temático se han establecido tipologías textuales en función de la disciplina o campo de conocimiento sobre el que tratan. La mayoría de estas temáticas especializadas no forman parte del conocimiento general de los hablantes de una lengua y, por tanto, han sido objeto de un aprendizaje especializado, como ocurre con la medicina, la ingeniería, el derecho, etc. A veces se hace complicado realizar esta clasificación clara de la temática, debido a la naturaleza híbrida de muchos textos elaborados así como a su banalización para llegar a un público más general.¹⁴

La temática en sí misma no aporta el carácter de especializado de un discurso, sino que, más bien, se trataría del uso lingüístico, junto con el contexto comunicativo, lo que determina la perspectiva desde la que se conceptualiza y transmite el conocimiento especializado (Cabré Castellví, 1993: 136).

Por su parte, el eje vertical, es decir, el nivel de especialización que caracteriza los diferentes niveles, ha sido un criterio fundamental para la clasificación de los diferentes textos. Este plano tiene como variables dos factores determinados: por un lado, los usuarios de los textos, es decir, el emisor y el receptor o receptores; y por otro, la situación comunicativa donde se da. Así, los usuarios podrán ser desde expertos y doctos en la materia hasta estudiantes y el público interesado, pasando por los profesionales y técnicos, y las situaciones comunicativas variarán desde publicaciones en revistas y obras especializadas hasta periódicos diarios, normalmente reguladas por criterios profesionales o científicos.

La combinación de estos tres elementos, a saber: la temática, los usuarios y la situación comunicativa donde ocurra determinará el aspecto pragmático del mensaje y de la especialización del mismo, es decir, dependiendo del emisor y del receptor del mensaje, del lugar en el que se dé o se publique así como del tema que trate, su nivel de especialización será mayor o menor.

La comunicación especializada muestra una diversificación discursiva en función de las características de la situación comunicativa. En otras palabras, el nivel de especialización de los textos se ajusta a los conocimientos especializados de una

¹⁴No obstante, siempre es necesario tener en cuenta el eje vertical de la comunicación, es decir, el grado de especialización dentro de la temática.

comunidad lingüística en un tema determinado, con lo cual cuanto mayor sea la competencia de los interlocutores de un texto mayor será su especialización y mayor será la frecuencia de unidades terminológicas. Tanto es así que, como indica Pearson (1998: 36), los términos serán usados con un mayor o menor grado de especificidad dependiendo del grado de especialización del texto y los autores ofrecerán más o menos información sobre los términos y su significado dependiendo del conocimiento especializado que suponen a sus lectores —y, por supuesto, del suyo propio. Así, en los textos muy especializados, el número de unidades terminológicas será superior al de los textos divulgativos, aunque eso no impida que estos últimos se clasifiquen también como especializados. En cualquier caso, cuanto mayor sea la presencia de estas unidades terminológicas, mayor precisión y concisión tendrá el lenguaje de especialidad en cuestión.

Los emisores de los textos de especialidad solo pueden serlo si tienen el conocimiento especializado que se va a transmitir y, por ello, los denominamos “expertos”, ya sean los expertos propiamente dichos, es decir, que han adquirido el conocimiento en un proceso natural de aprendizaje (realización de estudios o inserción en un contexto profesional), ya sean aquellos que han adquirido este conocimiento únicamente para poder ejercer de mediadores entre expertos y legos. En cambio, los posibles receptores de conocimiento especializado pueden ser variados. Se habla en general de tres tipos de destinatarios, cada uno de los cuales justifica un nivel discursivo: el destinatario experto, el semiexperto y el lego. De esta manera, podemos distinguir tres situaciones diferentes:

- La comunicación entre expertos (*expert-expert communication*) se considera una comunicación especializada o altamente especializada y se suele utilizar un vocabulario altamente especializado y en la que los términos se usan de forma precisa y de acuerdo con un significado muy concreto, conocido y compartido por los participantes. Además, contienen abundantes rasgos característicos de la lengua de especialidad en cuestión. No es frecuente encontrar explicaciones sobre el significado de los términos en este tipo de textos, a no ser que el especialista esté redefiniendo un concepto existente o acuñando un término nuevo. Estos textos se presentan en revistas

muy especializadas, encuentros y reuniones de especialistas y libros altamente especializados.

- La comunicación experto-iniciado (*expert to initiates*) hace referencia a la comunicación que realizan los expertos cuando dirigen su mensaje a personas que poseen un conocimiento inferior al suyo sobre el área de especialidad, ya sean estudiantes, especialistas de otras materias o legos con amplia cultura. Es probable que en este caso, el especialista emplee los mismos términos que usaría para dirigirse a otro especialista (el caso anterior), aunque probablemente explique el significado de algunos de los términos que considere problemáticos o desconocidos para sus interlocutores. En el caso de la comunicación profesor-alumno, la comunicación tendrá un carácter eminentemente pedagógico y aparecerán gran cantidad de definiciones y explicaciones sobre el significado de los términos. Se suele dar en revistas de divulgación científica, manuales especializados, conferencias, debates y discusiones entre especialistas del mismo u otro campo.

- La comunicación entre (semi)experto y lego (*relative-expert to the uninitiated*) se da en algunas revistas de divulgación científica, las secciones de temática especializada de algunas publicaciones (secciones dedicadas a la medicina general, psicología, sexología o consultoría jurídica) o folletos informativos. En este tipo de textos no se presupone ningún conocimiento especializado por parte de los lectores, por lo que la densidad de términos especializados suele ser bastante baja y, cuando se usan, suele explicarse su significado de forma accesible y, en ocasiones, imprecisa.

| Nivel de especialización | Emisor | Receptor | Lugar de aparición |
|---|-----------|---|--|
| Especializado (altamente especializado) | Experto | Experto | Libro Revista especializada Conferencia (para especialistas) |
| Medio-especializado | Experto | - No experto con conocimientos previos (iniciado) - Estudiante | Revista de divulgación científica Manual Conferencia (para no especialistas) |
| Divulgación | - Experto | No experto sin | Revistas de divulgación |

| | | | |
|--|--------------|-----------------------|-----------------------|
| | - No experto | conocimientos previos | Folletos informativos |
|--|--------------|-----------------------|-----------------------|

Tabla 1. Niveles de especialización de un texto y características.

No obstante, y pese a esta clasificación en niveles de abstracción aceptados por la mayoría de los autores, no todos los autores están de acuerdo, ya que defienden la existencia de unos lenguajes marcadamente más especializados que otros, como es el caso de Rondeau (1983), que considera lenguajes con un alto grado de especialización a los lenguajes de la química, la física, la matemática, la arquitectura, etc.; mientras que los lenguajes de la banca, el derecho, la bolsa o la economía aplicada constituyen un terreno intermedio y los lenguajes de la peluquería, la restauración, los deportes presentarían un grado mucho menor de especialización. De esta manera, se podría incluir una cuarta variable en la caracterización de los lenguajes especializados dentro del plano pragmática para poder clasificar el lenguaje especializado en cuestión dentro de esta clasificación de Rondeau.

1.2.2.3. Plano funcional

A pesar de la variedad de situaciones comunicativas, los destinatarios del discurso especializado poseen todos una característica común: su expectativa de ser informados. Así, llegamos al tercer y último plano: el plano funcional.

Teniendo en cuenta las funciones que estableció Jakobson (1963), a partir de los elementos de la situación comunicativa, se establecen seis funciones del lenguaje diferentes, a saber: a) Función emotiva: centrada en el emisor; b) Función conativa: centrada en el receptor o destinatario; c) Función referencial: centrada en el contenido o referente; d) Función metalingüística: centrada en el código; e) Función fática: centrada en el canal, y f) Función poética: centrada en el mensaje.

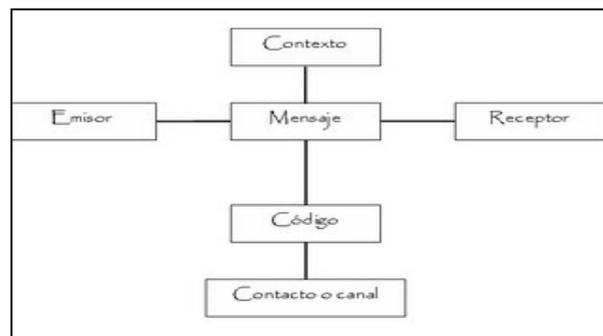


Figura 3. Elementos constitutivos de una situación comunicativa

Partiendo de esta clasificación, se determina que la función principal presente en todos los textos especializados es la función referencial, que consiste en la transmisión de la información de forma objetiva. Los lenguajes de especialidad se caracterizan por subordinar lo estético y lo expresivo a lo objetivo y a la eficacia comunicativa y, por ello, se ven favorecidos los usos lingüísticos capaces de expresar orden, claridad, concisión. No obstante, esta función siempre suele aparecer acompañada de otras funciones establecidas por Jakobson que ocuparían el lugar de funciones secundarias, como pueden ser la función apelativa, que se produce cuando el emisor exige al receptor una respuesta activa o intenta influir en su conducta; la función expresiva, que consiste en la transmisión de sentimientos del emisor, o la función poética, que presta atención a la forma en la que se elabora el mensaje.¹⁵

Por tanto, las funciones presentes en un mensaje variarán, más concretamente las funciones secundarias, según los factores pragmáticos citados anteriormente, es decir, dependerán del público al que vaya dirigido, del lugar de aparición y del tema del que traten.

A modo de conclusión y teniendo en cuenta la caracterización realizada anteriormente, consideraremos lenguaje de especialidad a los sublenguajes lingüísticos, funcionales y pragmáticos dentro de la lengua general, caracterizados por su función eminentemente referencial y su tipología textual y especializados según la temática, los interlocutores que intervienen en la situación comunicativa y el canal de comunicación utilizado.

1.3. LA DICOTOMÍA LEXICOLOGÍA - TERMINOLOGÍA

¹⁵ Estas funciones, junto con la función metalingüística y poética, están presentes en el discurso turístico, como veremos en el capítulo 3.

La diferenciación entre Lexicología y Terminología¹⁶ (y por extensión entre Lexicografía y Terminografía)¹⁷ es una cuestión polémica, ya que a menudo resulta imposible distinguir entre lo que se debe considerar unidad especializada de lo que no, es decir, trazar una división nítida y clara entre lo que se encuentra dentro del discurso especializado de lo que pertenece a la lengua general se vuelve confuso a menudo.

La Lexicología surgió con la finalidad de estudiar las unidades léxicas que constituyen el vocabulario de la lengua general, y su objetivo consiste en describir y explicar de la forma más adecuada posible el funcionamiento léxico del hablante, así como de prever las posibilidades reales que tienen los hablantes de cualquier lengua para formar unidades nuevas siguiendo pautas estructurales sistemáticas. Este estudio abarca, por tanto, la formación y el significado de las unidades léxicas, las combinaciones de estas entre sí, además de la relación de la Lexicología con otros niveles de la gramática, como la morfología, la fonología, la sintaxis y las informaciones extralingüísticas que rodean la unidad en cuestión y que pueden determinar que el hablante no emplee una unidad en ciertas circunstancias. Dentro de la lingüística aplicada también se manifiesta un interés por la praxis de esta materia, es decir, por la dimensión aplicada de los estudios teóricos. La Lexicografía es la disciplina que se ocupa de estos aspectos, así como de los principios que rigen la elaboración de los diccionarios de la lengua general.

Posteriormente, surgió la Terminología, que en vez de dedicarse a las unidades léxicas como la Lexicología, su interés se centró en las unidades especializadas. Wüster, considerado como el padre de la Terminología como disciplina científica, comenzó a estudiar estas unidades en los años 30 y, a partir de entonces, surge un interés mayor que llega hasta la actualidad. No obstante, la Terminología, cuyo objeto es el estudio y la recopilación de los términos especializados, no es una materia tan reciente, ya que la elaboración de vocabularios y glosarios terminológicos es bastante anterior a la época wüsteriana. No hay más que recordar por ejemplo los trabajos que en el siglo XVIII

¹⁶ «Terminología» designa tres conceptos diferentes: a) el conjunto de principios y de bases conceptuales que rigen el estudio de los términos; b) el conjunto de directrices que se utilizan en el trabajo terminográfico, es decir, la metodología; y c) el conjunto de términos de una determinada área temática de especialidad. En nuestro caso, nos referimos a la primera acepción, es decir, a la disciplina.

¹⁷ Cf 1.4.

realizaron Lavoisier y Berthold en química, o Linneo en botánica y zoología, con la intención de fijar las denominaciones de los conceptos científicos que usaban los especialistas. Durante los siglos XIX y XX, debido al desarrollo de la ciencia y la técnica, fue necesario no solo denominar los nuevos conceptos, sino sobre todo normalizar y fijar las nuevas denominaciones que permitieran la comunicación entre especialistas de forma precisa y sin ambigüedades (Tebé, 2003).

Al igual que la Lexicografía, la Terminología también tiene una rama aplicada que se encarga de la aplicación de los estudios teóricos: la Terminografía. Sin embargo, en este caso, la distinción entre una denominación y otra, es decir, entre Terminología y Terminografía no queda tan nítida como en el caso de Lexicología-Lxicografía, puesto que a veces estos términos se utilizan indistintamente. Según Bergholtz y Tarp (1995: 10), uno de los motivos de esta confusión es la reciente acuñación del término «Terminografía», ya que esta se acuñó oficialmente a partir de 1975 cuando la *International Organization for Standardization* (ISO) la empleó para hacer referencia a la parte de la Terminología dedicada a la producción de diccionarios terminológicos. Anteriormente, se habían utilizado los dos términos indistintamente, lo que ha dado lugar a que actualmente numerosos autores lo sigan haciendo así o, incluso, se hayan creado otras denominaciones para diferenciarlos, como es el caso de «*terminology processing*» (Wright y Budin, 1997) o «*terminology management*» (Sager, 1990) para hacer referencia a la Terminografía. No obstante, otros autores (Cabré Castellví, 1993, 2000; Meyer y Mackintosh, 1996; Rey, 1995; Sager, 1996) defienden la diferencia entre denominaciones para referirse a la rama teórica y la rama aplicada, respectivamente. En nuestro caso, también distinguiremos «Terminología» de «Terminografía» y, por tanto, no realizaremos un uso indistinto de dichos términos.

En este sentido, el término «Terminografía» hace referencia a la rama aplicada de la Terminología que se encarga de la elaboración de diccionarios especializados (Cabré Castellví, 1993: 263), teniendo en cuenta que el término «diccionario» hace referencia a su acepción más generalizada.

En esta misma línea, la Norma ISO 1087 (1990) *Vocabulary of Terminology* nos ofrece una definición de Terminología y de Terminografía muy concreta y clara. Por

Terminología¹⁸, entiende «Science studying the structure, formation, development, usage and management of terminologies in various subject fields»; en cambio, por Terminografía considera «a part of terminology work concerned with the recording and presentation of terminological data. NOTE: Terminological data may be presented in the form of term banks, glossaries, thesauri or other Publications».

Asimismo, con objeto de perfilar esta diferencia entre Terminología y Terminografía, la mayoría de los autores suelen establecer una correspondencia paralela entre, por un lado, la Lexicología, disciplina que se ocupa del estudio y descripción del léxico de una lengua, y la Lexicografía, concebida como la rama aplicada de la Lexicología centrada en la elaboración de diccionarios; y por otro, la Terminología, encargada del estudio teórico y metodológico, y la Terminografía, que hace referencia a la vertiente aplicada de la Terminología, encargada de la elaboración de diccionarios especializados. De esta forma, la Lexicología es a la Terminología lo que la Lexicografía a la Terminografía.

A pesar de que las líneas divisorias entre la Lexicología y la Terminología, y entre la Lexicografía y la Terminografía, parecen claras y consensuadas, la dicotomía no es tan precisa y sus límites parecen borrosos, al igual que ocurría con la distinción entre lenguaje de especialidad y lengua general. La razón de esta falta de claridad en su diferenciación se debe principalmente a la indiscutible interacción que existe entre ambas, ya que tanto una como otra se ocupan de las palabras, es decir, de la unidad léxica como referencia a la realidad; ambas presentan una vertiente teórica y una vertiente aplicada, y el objeto aplicado de ambas disciplinas es la elaboración de diccionarios (Cabré Castellví, 1993: 86). Asimismo, comparten principios en la metodología de trabajo (como los relacionados con la formulación definicional) (Pearson, 1998: 67), en las técnicas de recopilación y sistematización, en los programas informáticos actuales, entre otros.

No obstante, no todos los aspectos son coincidentes. De hecho, algunos autores consideran la Terminología y la Lexicología como dos disciplinas totalmente

¹⁸ En la Norma ISO 1087 (1990) *Vocabulary of Terminology* se ofrecen dos definiciones diferentes de Terminología: por un lado, la definición ofrecida arriba, que hace referencia a su acepción de disciplina; y por otro lado, la siguiente definición: «a set of designations belonging to one special language», que se refiere a su carácter de producto terminológico.

independientes mientras que otros consideran que la Terminología se encuentra dentro de la Lexicología. Las siguientes serían las dos posturas citadas:

Una primera postura defiende que se trata de una disciplina autónoma y autosuficiente, es decir, una materia absolutamente original, dotada de fundamentos propios, aunque evidentemente conectada históricamente a otras disciplinas. Esta sería la posición de los defensores de la Teoría General de la Terminología, de Wüster, para los que los términos no son lenguaje propiamente dicho, sino que incluyen el lenguaje.

Para aquellos que argumentan que la terminología, en tanto que conjunto de términos que se pueden describir desde una teoría del lenguaje, se puede explicar desde la lingüística y más concretamente desde la Lexicología, la Terminología sería igual a la Lexicología especializada. En este caso, nos encontraríamos con corrientes de la Terminología moderna.

Por lo que respecta a la primera propuesta, aquellos que defienden el carácter autónomo de la Terminología, se basan en la existencia de dos objetos de estudio diferentes: las palabras, para los lexicólogos y los términos, para la Terminología. Esto es lo que justifica que una disciplina como la Terminología pueda prescindir de la lingüística y desarrollarse sin ella. El principal foco impulsor de estas ideas es la Teoría General de la Terminología (TGT), propuesta por Wüster, el cual empezó a aplicar su método sistemático en los trabajos de terminología científica y técnica con objeto de normalizar la terminología empleada por los expertos en una especialidad determinada. Según este autor, la Terminología se distingue totalmente de la lingüística, y más concretamente de la Lexicología, por los siguientes motivos (Santamaría Pérez, 2006; Cabré Castellví, 1993: 86):

| Lexicología | Terminología tradicional |
|---|---|
| - Objeto de estudio: palabra. | - Objeto de estudio: término. |
| - Campo de trabajo: análisis y descripción de la competencia léxica del hablante. ¹⁹ | - Campo de trabajo: análisis y descripción de los términos de un campo de especialidad o área profesional. |
| - Finalidad: entender el comportamiento léxico. | - Finalidad: denominar y normalizar formas de referencia (forma y contenido), mediante elementos teóricos y prácticos capaces de regir la búsqueda, selección y ordenación de los términos propios de los campos de especialidad. |
| - Punto de partida: forma y significado | - Punto de partida: el concepto (significado), independiente de denominación (forma). |
| - Método de trabajo: semasiológico. | - Método de trabajo: onomasiológico. |
| - Incluye la gramática y la pragmática. | - Prescinde de la sintaxis y la morfología. |
| - Estudio sincrónico y diacrónico. | - Estudio sincrónico de los términos |
| - Evolución natural de la lengua. | - Formación consciente y consensuada (norma prescriptiva). |
| - Estudio de las lenguas particulares. | -Normalización internacional. |
| - Estudio de forma fónica y forma gráfica. | - Prioridad a la forma escrita. |

Tabla 2. Relación entre Terminología y Lexicología.

Como decíamos, desde este punto de vista, la Terminología se concibe como una disciplina autónoma con respecto a otras disciplinas, como la Lexicología.

¹⁹ Con respecto al campo de trabajo, en la Terminología moderna se consideraría esto como una relación de inclusión: el campo de la Lexicología es más amplio e incluye la Terminología y, por tanto, se hablaría de la Terminología como parte de la Lexicología. Sin embargo, en la Terminología tradicional, al realizar la distinción nítida entre palabra y término, no cabría tal relación de inclusión, sino que existiría una independencia.

En los últimos años se han criticado los supuestos teóricos de la TGT, lo que ha dado lugar al surgimiento de otras teorías terminológicas con concepciones diferentes y al surgimiento de lo que se conoce como la Terminología moderna, en contraposición a la Terminología tradicional, de Wüster, como veremos en el siguiente apartado.

1.3.1. EL CONCEPTO TÉRMINO

Como señala Pérez Hernández (2002), la dicotomía planteada entre Lexicología y Terminología, por un lado, y lengua general y lenguaje de especialidad, por otro, parece lógica si volvemos a lo que Cabré Castellví denomina planteamientos wüsterianos (Cabré Castellví, 1999b: 111), es decir, los fundamentos de la teoría terminológica tradicional (cf. 1.3.2.1.). Siguiendo esta misma línea de los planteamientos de la Terminología tradicional, no debemos olvidar la dicotomía palabra y término, siendo la palabra, la unidad léxica de la lengua general que se encarga de estudiar la Lexicología, y el término, la unidad especializada que se encarga de estudiar la Terminología.

De esta manera, en el seno de la TGT los términos se consideran como unidades específicas de un ámbito de especialidad y su uso queda circunscrito a ese ámbito, lo que hace que el interés para la Terminología se limite solo y exclusivamente a la comunicación entre especialistas. Dentro de esta concepción teórica, podemos distinguir la definición de «término» de la Norma ISO 1087 (1990: 5) *Vocabulary of Terminology*:

Term: Designation of a defined concept in a special language by a linguistic expression. Note: a term may consists of one or more words (i.e. simple term or complex term) or even contain symbols.

Donde la designación hace referencia a «any representation of a concept» (ibíd.: 5) y el concepto a «a unit of thought constituted through abstraction on the basis of properties common to a set of objects» (ibíd.: 1). Esta definición dada por un organismo de normalización internacional es tan amplia que resulta de poca utilidad para diferenciar una unidad léxica de una unidad especializada, además se ciñe a los fundamentos de la TGT y, por tanto, es demasiado restrictiva y no contempla todos los posibles aspectos, como son los aspectos pragmáticos y comunicativos.

Estos planteamientos tan estrictos y radicales que defienden una distinción absoluta entre las palabras y los términos han recibido numerosas críticas de la Terminología moderna, que destacan la irrealidad de dichos planteamientos debido a la imposibilidad de trazar objetivamente un límite entre palabra y término. Además, la Terminología moderna defiende el hecho de que el conocimiento especializado no puede restringirse a áreas limitadas y estancas, sino que se trata de un continuo, tanto con la lengua general como con los diferentes lenguajes especializados.

Siguiendo esta misma línea de crítica a la TGT, Ahmad et ál. (1995) nos define el concepto «término»:

A term is a linguistic expression of a concept belonging to one or more domains. It is pragmatically restricted, being used by domain experts. However, laypeople are not excluded from using the same terms or being confronted by them. Hence, terms are to be found in general-language dictionaries.

En esta cita, los autores nos presentan las principales características de los términos dentro de la Terminología moderna. En primer lugar, determinan que un término puede pertenecer a más de un campo de especialidad, es decir, la «expresión lingüística de un concepto» puede ser polisémica;²⁰ en segundo lugar, nos introducen el aspecto pragmático de la comunicación especializada, que normalmente se realiza entre expertos, es decir, la comunicación experto-experto, pero que no queda limitada a ellos, ya que contempla, además, la posibilidad de participación de hablantes con un limitado conocimiento especializado, y en tercer lugar, presentan la opción de la aparición de acepciones de palabras que se podrían clasificar como términos en diccionarios generales.

Esto último se debe, sobre todo, a los constantes trasvases que se realizan de los campos de especialidad a la lengua general, y viceversa, lo que provoca que la línea divisoria entre ambas sea todavía más borrosa. Por un lado, como señala Cabré Castellví (1993: 167), la extensión de los medios de comunicación de masas, la difusión comercial, el desarrollo acelerado de la tecnología, entre otros factores, ha favorecido la

²⁰ Como veremos en el apartado siguiente, esta es una de las modificaciones que se han incorporado a la Terminología moderna derivada de las críticas al principio de univocidad y monorreferencialidad de la TGT.

divulgación de las materias especializadas, lo que ha dado lugar a una gran difusión de las unidades especializadas y a la adquisición de conocimiento especializado, de forma indirecta, de los hablantes no especialistas. A este fenómeno se le ha venido a denominar «banalización» (1993: 168). Por otro lado, también se produce un transvase importante de términos de la lengua común a los lenguajes especializados. Este fenómeno ha sido denominado «terminologización» de palabras de la lengua general (ibíd.) y consiste en el trasvase de unidades que pasan a los distintos ámbitos de especialización con significados precisos en cada caso, y por lo tanto, diferenciados entre sí.

Asimismo, puede tener lugar un tercer transvase de unidades de una lengua de especialidad a otra, a veces con el consiguiente cambio de significado o bien el referente conceptual en ambas disciplinas.

The differentiation of terms from words is not straightforward, since the relationship between general language and sublanguages (or special languages - LSPs) is an interdependent one. Not only do terms and words draw on the same set of phonological, morphological and morpho-syntactic rules, as well as on the same inventory of word forms in many cases, but particular forms may also cross the boundary between the two language varieties (Ahmad et ál., 1985: 10).

Estos tres trasvases se pueden resumir de la siguiente manera (ibíd.: 10):²¹

- LE --> LG: un término de un lenguaje de especialidad extiende su uso a la lengua general, por ejemplo «parámetro» (matemáticas), «paranoia» (psicología).
- LG --> LE: una unidad léxica de la lengua general es adoptada por un lenguaje de especialidad, por ejemplo «ventana» o «ratón» (informática).
- LE1 --> LE2: un lenguaje de especialidad adopta un término perteneciente a otro lenguaje de especialidad, por ejemplo «virus» (microbiología/ tecnología de la información) o «morfología» (biología/ lingüística).

²¹ Siendo LE, lenguaje de especialidad, y LG, lengua general.

Vinculados a estos trasvases que ocurren entre la lengua general y los lenguajes especializados, y entre estos entre sí, algunos autores (Cabré Castellví, 2004a; Alcaraz Varó, 2000; Hoffman, 1985; Godman y Payne, 1981), han propuesto una división de los términos en tres grupos diferentes que nos muestra, de nuevo, la dificultad de distinguir entre el término y la palabra:

- En primer lugar, encontramos las unidades léxicas usadas exclusivamente en un ámbito especializado, como por ejemplo «sindeína», del discurso de la bioquímica o «sínfisis», de anatomía.
- En segundo lugar, las unidades léxicas usadas en un ámbito especializado y, además, en el discurso general aunque con un significado distinto, como es el caso de «síntoma», «síndrome» o «morboso» en el discurso biomédico. En este grupo, también encontramos las unidades polisémicas que adquieren un significado diferente según el discurso especializado en el que se encuentren, como es el caso de «sistema», que presenta un significado distinto si se emplea en un discurso matemático, geológico o informático.
- En tercer lugar, están las unidades léxicas que comparten el mismo significado tanto en el discurso especializado como en la lengua común, como son «ensayo» y «enfermedad».

En este contexto, y en contraposición a lo que defiende la TGT, la distinción entre la palabra y el término, y entre las diferentes acepciones de este, solo se puede realizar en contexto, es decir, mediante la observación del uso real de la unidad dentro de una situación comunicativa concreta.

[...] El carácter de término lo activan en función de su uso en un contexto y situación adecuados. Esta activación consiste en una selección de los módulos de rasgos apropiados, que incluyen los rasgos morfosintácticos generales de la unidad y una serie de rasgos semánticos y pragmáticos específicos que describen su carácter de términos dentro de un determinado ámbito (Cabré Castellví, 1998d: 123).

Esta concepción se deriva del hecho de que, desde el punto de vista de la Terminología moderna, no existe ninguna diferencia entre el término y la palabra, sino que se trata de unidades del léxico de una misma lengua que adquieren un valor de contenido especializado en cuanto «se sitúan» y son utilizados «situadamente» (Cabré Castellví,

2002: 29). Por lo tanto, son las condiciones pragmáticas las que activan la condición de término en una unidad léxica determinada y en función de ellas (el tema, el ámbito de especialidad, la perspectiva que toma el usuario, el tipo de texto, el receptor, etc.) se activa también el significado y el uso lingüístico específico de cada término.

1.3.2. CORRIENTES TEÓRICAS DE LA TERMINOLOGÍA

La Terminología, en tanto que disciplina científica, tiene un origen muy reciente; sin embargo, las actividades en esta área se remontan al siglo XVIII, con los trabajos de Lavoisier y Berthold en química o Linneo en botánica y zoología, como veíamos anteriormente.

Durante el siglo XX, comienza su evolución debido al desarrollo de la ciencia y la técnica y a la necesidad no solo de denominar nuevos conceptos, sino de normalizar y fijar las nuevas denominaciones que permitan la comunicación entre especialistas de forma precisa y sin ambigüedades (Tebé, 2003). La Terminología adquiere la categoría de disciplina científica de la mano de E. Wüster, que funda la Escuela de Viena en 1931, y años más tarde propone la Teoría General de la Terminología (TGT) (Wüster, 1998). Esta teoría estará en auge hasta la década de los 70, cuando aparecen nuevas teorías que critican los supuestos de la TGT y establecen cambios de base, y comienza lo que se conoce como la Terminología moderna.

Desde un punto de vista cronológico (Cabré Castellví, 1993: 28-29), la Terminología como disciplina científica puede dividirse en cuatro periodos básicos: los orígenes (de 1930 a 1960), la estructuración (de 1960 a 1975), la eclosión (de 1975 a 1985) y la ampliación (desde 1985):²²

- El primer periodo de la Terminología (1930-1960), donde se establecen los principios teóricos de la disciplina, se caracteriza por el inicio de métodos de trabajo terminológico que tienen en cuenta el carácter sistemático de los términos y por la creación de la Federación Internacional

²² En la actualidad, la autora ha ampliado estos periodos a seis, siendo el quinto periodo el de la apertura de la ampliación (1992-1997) y el sexto, la diversificación (1997 en adelante), según afirmó en una conferencia inédita pronunciada en Málaga en junio de 2002.

de Asociaciones Nacionales de Normalización (ISA), el precedente inmediato del actual Organismo Internacional de Normalización (ISO). En 1951, se establece, en el seno de ISO, el comité técnico para la normalización terminológica (Comité Técnico 37, ISO/CT 37, encabezado por E. Wüster).

- En la segunda etapa, de estructuración (1960-1975), se desarrollan sistemas informáticos y técnicas documentales. Aparecen los primeros bancos de datos y se publican las primeras normas sobre Terminología del ISO/CT 37.

- En el tercer periodo, de eclosión (1975-1985), destaca la proliferación de proyectos de planificación lingüística que incluyen la Terminología. En esta etapa se pone de relieve el papel que esta disciplina desempeña en el proceso de modernización de una lengua y de la sociedad que la utiliza.

- En el último periodo, de ampliación, que se inicia en 1985, el desarrollo y avance de la informática facilita la labor y el tratamiento de los datos y provoca un cambio en la elaboración del trabajo terminológico. Asimismo, los terminólogos disponen de instrumentos y recursos de trabajo más adaptados a sus necesidades, más sencillos en cuanto a su manejo y más eficaces. Finalmente, se consolida y amplía la cooperación internacional y se crean redes internacionales que sirven de enlace entre organismos y países.

Como decíamos anteriormente, la necesidad de denominar nuevos conceptos, de normalizar y fijar las nuevas denominaciones que van surgiendo en la industria y en la ciencia motivan el interés por la Terminología y, consecuentemente, por la sistematización de las unidades especializadas. Según Cabré Castellví (1993: 39-41), este interés por la sistematización terminológica surgió simultáneamente en varios países: Austria, la Unión Soviética y Checoslovaquia. En estos tres países, se crearon las primeras escuelas de Terminología a partir de los años treinta, es decir, en la primera fase de la Terminología, después de que se publicara la tesis doctoral del ingeniero Eugen Wüster sobre normalización internacional del lenguaje de la electrotécnica, *Internationale Sprachnormung in der Elektrotechnik* (1931), la cual se considera como un tratado de la TGT.

A partir de estas tres escuelas, el estudio y la práctica de la terminología se extendió hacia diferentes puntos del mundo: en una primera fase, se dirigió hacia el oeste (Francia, Canadá, Quebec) y hacia el norte (Bélgica, Dinamarca y países nórdicos); en una segunda etapa, se expandió hacia el sur (África, América Central y meridional, Portugal y España), y, recientemente, hacia el este (China y Japón). Cada uno de estos ejes se ha caracterizado por el contexto en que se desarrolla la Terminología y por las finalidades que debe cumplir, por lo que el objetivo de su sistematización varía también según su contexto y su finalidad. Asimismo, han dado lugar a la formación de diferentes teorías de la Terminología con enfoques y directrices particulares, como veremos a continuación.

1.3.2.1. La corriente clásica de la Terminología (TGT)

Como hemos visto con anterioridad, los puntos iniciales de donde surgió la Terminología y desde donde se extendió al resto del mundo conforman las tres escuelas principales de esta disciplina: i) la Escuela de Viena, en Austria; ii) la Escuela de Praga, en la antigua Checoslovaquia, y iii) la Escuela soviética, en la antigua Unión Soviética. En estas escuelas, se crearon las corrientes terminológicas tradicionales.

Además de estas escuelas, se deben destacar otros centros de investigación y otras iniciativas en los países a los que se extendieron los estudios y prácticas de las primeras escuelas, como los centros de investigación en Canadá (especialmente en Quebec), en Francia, en los países nórdicos y en Gran Bretaña, el antiguo UMIST (*University of Manchester Institute of Science and Technology*).²³

Las primeras escuelas de Terminología surgieron a partir de los años 30, después de la publicación de la tesis doctoral de E. Wüster sobre la normalización de un lenguaje de especialidad. En este trabajo, el ingeniero recoge los postulados de la Teoría General de la Terminología (TGT), que vendrían finalmente publicados en su obra póstuma en

²³ En 2004, el UMIST se fusionó con la Universidad Victoria de Mánchester, lo que ha dado lugar a la Universidad de Mánchester, una de las más grandes universidades de Europa.

Einführung in die Allgemeine Terminologielehre und terminologische Lexikographie,²⁴ la cual recoge y expone sistemáticamente su doctrina sobre el tratamiento de la terminología.

Wüster fue el fundador y máximo representante de la escuela de Viena y se ha venido a considerar el padre de la Terminología, ya que gracias a su trabajo esta disciplina alcanzó el estado de materia autónoma interdisciplinar. Su labor estuvo respaldada por la entonces ISA (*International Standard Association*),²⁵ que en 1947 se creó para desarrollar normas universales con objeto de mejorar la comunicación y cooperación internacionales y reducir las barreras de los intercambios comerciales en el mundo (Cabré Castellví, 1998c). Estos son los objetivos principales de la TGT y, por tanto, de la Escuela de Viena, es decir, la simplificación y normalización de los términos para mejorar las relaciones internacionales y evitar las ambigüedades y problemas debidos por falta de comprensión terminológica.

Por su parte, las otras dos escuelas tradicionales se dedican a otros aspectos, aunque relacionados: por un lado, la Escuela de Praga, con I. K. Weisenberg, se especializa en la lingüística funcional y realiza investigaciones sobre el lenguaje normalizado con fines comunicativos en todas las áreas de la vida social; y por otro lado, la Escuela soviética se centra, principalmente, en los problemas terminológicos desde un punto de vista lingüístico.

En general, las tres escuelas tradicionales cuentan con unos postulados muy coincidentes acerca de la noción de concepto, de término y la relación existente entre ambos; de los sistemas conceptual y terminológico, y de la relación entre Terminología, planificación lingüística y normalización.²⁶

²⁴ Véase Wüster (1979), trad. Cabré, 1998.

²⁵ En la actualidad, ISO (*International Organization for Standardization*). Actualmente, la ISO/TC 37 sigue trabajando con los mismos propósitos que los iniciales de la ISA.

²⁶ Véase Laurén y Picht (1993) y Felber (1984) para profundizar en los inicios de la Terminología como disciplina y en las escuelas tradicionales.

Temmerman (2000: 4-15) resume las características fundamentales que definen los principios de la TGT y, por tanto, las líneas principales de las escuelas tradicionales, en los cinco puntos siguientes:²⁷

- El análisis de los términos comienza a partir del concepto (enfoque onomasiológico) y, por tanto, se asume que el concepto precede y pre-existe a la denominación.
- Los conceptos de un mismo campo de especialidad deben ser estudiados como elementos pertenecientes a un mismo sistema conceptual y, por tanto, no se deben analizar de forma aislada. Mantienen entre sí relaciones de diferente tipo, las cuales constituyen la estructura conceptual de dicho campo²⁸ y su valor se establece por el lugar que ocupa en dicha estructura.
- Las definiciones terminológicas deben reflejar la posición de un concepto en un sistema conceptual (definición por intensión, *intentional definition*).
- Los términos son asignados permanentemente a un concepto, por lo que un concepto solo designará un término y un término solo se referirá a un concepto dentro de un campo de especialidad (principio de univocidad). En este sentido, se anula cualquier posible sinonimia y polisemia de los

²⁷ La semántica estructuralista de Saussure, teoría lingüística con gran importancia en el momento, ha influido en gran medida en la TGT. Debido a esto, se pueden reconocer supuestos saussureanos en algunos de los puntos principales de la teoría de Wüster, como se puede observar en los cinco puntos comentados a continuación. Entre ellos, podemos indicar los siguientes: uno, los significados (o conceptos) pueden ser delimitados de forma clara y precisa; dos, la mejor forma de describir los significados consiste en describir la delimitación mutua de los conceptos mediante relaciones semánticas (lógicas y ontológicas); tres, la mejor forma para describir significados es centrarse en el significado denotativo (opuesto a significado connotativo) y en el significado literal (por oposición al significado figurativo) de las unidades léxicas; y por último, el significado se debe describir de forma sincrónica y no atender a la evolución y desarrollo de las unidades.

²⁸ Estas relaciones entre conceptos se pueden dividir en dos tipos: por un lado, las relaciones jerárquicas y, por otro, las relaciones no jerárquicas. Las primeras pueden ser, a su vez, lógicas u ontológicas y se utilizan para la formalización de sistemas jerárquicos de conceptos. Las relaciones lógicas (o genéricas, hiperonimia-hiponimia) se basan en el grado de semejanza entre conceptos y se pueden expresar de la siguiente manera: “Y es igual a X”, por ejemplo “un barco es un tipo de medio de transporte”. Por su parte, las relaciones ontológicas (o partitivas, parte-todo) se crean entre el todo y sus partes y pueden ser del tipo “Y es parte de X” (“timón es parte de barco”), “Y sucede a X” (hijo sucede a padre), “Y está hecho de X” (“anillo está hecho de oro”), etc. Por último, se encuentran las relaciones no jerárquicas, es decir, las que no representan sistemas jerárquicos. Estas pueden ser causales (causa-efecto), genéticas (productor-producto), instrumentales (instrumento-uso), cronológicas (sucesión en el tiempo), de producción (material-producto), etc. En Sager (1990) se puede consultar una lista de relaciones semánticas de carácter general muy completa.

términos y se pretende realizar un trabajo de normalización conceptual y denominativa.

- Los términos se estudian desde un punto de vista sincrónico, ya que no se analiza la evolución o el desarrollo del lenguaje en el tiempo.

Mediante la explicación esquemática ofrecida por Temmerman (ibíd.), observamos los rasgos más característicos de la TGT, lo que nos permite conocer brevemente las principales ideas de las escuelas tradicionales. A modo de resumen, podemos indicar que esta teoría de la TGT defiende la univocidad absoluta de concepto y su denominación en una comunicación profesional formal y pretende, por tanto, establecer esquemas conceptuales estables, estáticos en tiempo y espacio y universales en los campos de especialidad, es decir, normalizar los conceptos de los campos de especialidad. Todo ello con objeto de evitar ambigüedades e incomprensiones en el ámbito internacional de la comunicación profesional. De esta manera, rechaza cualquier tipo de sinonimia o polisemia, defiende la diferencia entre lengua de especialidad y lengua común de forma clara e indudable, así como la independencia del término de su contexto, y la limitación a un registro formal profesional, entre otras cosas.

Actualmente, la Organización Internacional para la Estandarización (ISO), con su Comité Internacional de la Terminología (TC/37), así como otros organismos de normalización (CEN, AENOR), siguen trabajando y utilizando las directrices de la TGT para la realización de sus trabajos relacionados con la terminología. De esta forma, llevan a cabo una labor de prescripción, normalización y estabilización de la terminología de los campos de especialidad, al igual que pretendía Wüster.

No obstante, a pesar de la importancia de la prescripción en algunos casos y de la labor de estos organismos en la actualidad, se han puesto en duda determinados aspectos de la TGT a lo largo de las últimas décadas. Autores como Rey (1979), Sager (1990), Cabré Castellví (1993), Temmerman (1997, 2000), Condamines (1994, 1995), Pearson (1998), por citar algunos de los más representativos, han cuestionado los principales puntos defendidos por esta teoría tradicional y han comprobado que sus principios son insuficientes para dar cuenta de una realidad terminográfica inmersa en un proceso de comunicación en continuo cambio. En palabras de Cabré Castellví

(1998f: 112), el modelo propuesto por Wüster «silencia la diversidad denominativa y conceptual de la realidad».

A continuación, presentamos los principios básicos de las teorías terminológicas de la Terminología moderna, así como los cambios que han provocado este giro de visión.

1.3.2.2. Las corrientes modernas de la Terminología

It is impossible to name without names; it is impossible to name scientifically without terms; but names and terms belong to a language. There are no such things “ideal” or, as some wrongly claim, “international” terms (Rey, 1995: 59).

Los cambios sociales, políticos y económicos que han tenido lugar durante el siglo pasado, y más concretamente a partir de su segunda mitad, han dado lugar al surgimiento de unas necesidades nuevas que han puesto en cuestión muchos de los valores del periodo anterior. De esta manera, han forzado a replantear muchos aspectos de la organización social, política y económica tradicionales (Cabré Castellví, 2005). Algunos de estos cambios han afectado a la Terminología y han dado lugar al surgimiento de nuevas teorías terminológicas que discrepan con los supuestos teóricos de la TGT y que incorporan rasgos comunicativos y pragmáticos. Entre los aspectos más relevantes que han influido en el cambio de concepción teórica y práctica de la terminología cabe mencionar los siguientes (ibíd.):

- La diversificación de las situaciones de comunicación especializada debido a la diversidad de necesidades en la transmisión del conocimiento especializado.
- La multiplicación de la temática de intercambio internacional causada por el crecimiento acelerado del conocimiento y por la diversificación de los contenidos especializados, con la entrada de temas de las ciencias sociales y

humanidades y no solo de conocimientos propios de las ciencias y las técnicas.

- La difusión del conocimiento especializado a través de los medios de comunicación de masas y la extensión de la formación especializada en todos los niveles educativos.

- La aparición e implantación de las tecnologías en todos los ámbitos profesionales, así como la entrada de estas en los hogares. El aumento de la capacidad de almacenamiento de la información ha dado lugar a un interés prioritario por la adecuación de los recursos y la eficiencia de las estrategias de procesamiento, así como al surgimiento de herramientas, programas y técnicas que permiten agilizar la labor del terminólogo.

- La tendencia de las administraciones a la planificación lingüística de las lenguas oficiales, con objeto de generalizar el uso de la lengua propia en todas las situaciones de comunicación nacional. Este hecho ha dado lugar a la creación de gran número de recursos terminológicos, lo que ha influido, indirectamente, en el desarrollo de la Terminología como campo de conocimiento.

- La importancia y revalorización de todas las lenguas en los intercambios comunicativos en el ámbito nacional e internacional, con la reivindicación del plurilingüismo, especialmente motivado por la política europea actual.

Como decíamos, estos cambios han motivado el surgimiento de comentarios críticos con respecto a la TGT y han dejado entrever la necesidad de la Terminología de adaptarse a la evolución de la sociedad en todos sus aspectos. De esta manera, las teorías surgidas a raíz de las críticas a la TGT han puesto en tela de juicio las cinco líneas principales de la TGT mencionadas anteriormente.

En primer lugar, el enfoque onomasiológico utilizado por la TGT pretende reducir la conceptualización a una actividad mental que ocurre de forma independiente al lenguaje. De esta forma, no se tiene en cuenta el contexto comunicativo de un lenguaje especializado ni la aparición del mismo término en diferentes contextos lingüísticos ni el empleo de variantes que están motivadas por las condiciones

contextuales de la situación comunicativa (Sager, 1990: 59), lo que nos lleva a determinar la necesidad de incorporar el contexto de cualquier comunicación profesional en el análisis lingüístico que se realice de una unidad terminológica. En este sentido, Sager (1990) incorpora la dimensión comunicativa a las otras dos dimensiones tradicionalmente consideradas en la TGT, es decir, la dimensión cognoscitiva y la dimensión lingüística, con objeto de complementar la teoría tradicional de Wüster.

En segundo lugar, la división clara y exacta que realiza la TGT a la hora de tratar los términos como pertenecientes a un campo de especialidad y, por tanto, a un mismo sistema conceptual no es posible, ya que la clasificación de numerosos términos es confusa, tanto entre lenguajes de especialidad como con respecto a la lengua común, y es imposible clasificarlos a todos mediante reglas lógicas y ontológicas. En este sentido, es necesario tener en cuenta la multidimensionalidad de los conceptos, el trasvase que existe entre los lenguajes de especialidad y con la lengua común, así como la flexibilidad de los términos de adquirir significados nuevos, de reducir su frecuencia de uso, de la sinonimia existente, de los grados de representatividad, etc.

En tercer lugar, las definiciones que indican la posición del concepto dentro de un sistema conceptual (definiciones por intensión) son definiciones más sistemáticas que cualquier otro tipo; sin embargo, para la mayoría de los conceptos es imposible crear definiciones por intensión que sean significativas y, a veces, no son ni siquiera deseables. La TGT reconocía incluso que era imposible definir todos los conceptos mediante este tipo de definiciones y, a su vez, reconocían la necesidad de ofrecer definiciones explicativas, es decir, que no indican la posición del concepto en un sistema conceptual (Felber, 1984: 160), aunque de forma excepcional. Asimismo, los usuarios no especialistas necesitan disponer de definiciones enciclopédicas y de información semántica y léxica y, por tanto, las definiciones intensionales no son suficientes para satisfacer sus necesidades.

En cuarto lugar, el supuesto de la univocidad concepto-denominación hace referencia a la independencia de los conceptos con respecto al lenguaje y a la situación comunicativa, así como a la carencia de sinonimia y polisemia en el lenguaje de especialidad. Como ya veíamos en el primer punto de crítica, no es posible separar el lenguaje de la conceptualización, ya que ambos son complementarios y necesarios, así

como tampoco se puede obviar la situación comunicativa en la que se produce la comunicación profesional. Cualquier estudio terminológico sobre cualquier lenguaje de especialidad nos indicaría que el contexto comunicativo influye en la terminología utilizada, que se utiliza a menudo un lenguaje figurado, que existe gran variedad denominativa y que, por tanto, los términos no son independientes del lenguaje y es importante tener en cuenta los aspectos discursivos y comunicativos de cualquier lenguaje de especialidad. Roelcke²⁹ nos habla de «univocidad textual», en vez de univocidad concepto-denominación, de modo que se refiere a este fenómeno con relación al uso concreto de un término en un contexto determinado y no al significado aislado de este dentro de un sistema terminológico. De esta forma, el uso de un concepto se concreta dentro de una situación comunicativa y adquiere un significado unívoco en la mayoría de los casos.

Por último, el estudio sincrónico que defiende la TGT es muy cuestionable, ya que tanto los conceptos como las denominaciones evolucionan a lo largo del tiempo y, por tanto, sus significados también varían, ya sea con la incorporación, la reducción o la modificación de alguna acepción. En este sentido, el esquema conceptual de un lenguaje de especialidad no es invariable ni estable como defiende la TGT, sino que está en movimiento y sufre modificaciones con relación al desarrollo de sus conceptos y, por ello, es necesario realizar estudios diacrónicos de los términos y atender al dinamismo de las unidades terminológicas.

Estos cinco puntos comentados resumen brevemente los principales comentarios críticos que se han expuesto sobre los principios de la TGT y demuestran la visión simplista y el carácter reduccionista e idealista de esta teoría, así como la limitación de la materia a la normalización. Estas críticas no pretenden invalidar esta teoría, sino subrayar su limitación conceptual y funcional y su excesivo interés por la generalización para explicar las unidades terminológicas (Cabré Castellví, 1998f: 114). Así, las nuevas corrientes de la Terminología abogan por un enfoque menos restrictivo y menos normalizador que tenga en cuenta la complejidad, multidimensionalidad y variabilidad de las unidades terminológicas en los lenguajes de especialidad, así como por la apertura de la Terminología a otras disciplinas, como son la lingüística, la sociología, la

²⁹ Citado por Oster (2000: 91).

psicología cognitiva, entre otras. Como nos señala Díaz Rojo (2001), no se puede obviar la diversidad inherente del lenguaje humano natural que a menudo provoca incorrecciones, variaciones y alejamientos de la norma.

De entre estos trabajos críticos que han dado lugar a la Terminología moderna, cabe destacar la Socioterminología, desarrollada en Francia y en la parte francófona de Canadá por autores como Y. Gambier, F. Gaudin y J. C. Boulanger; la Teoría Sociocognitiva de la Terminología, de R. Temmerman, en Bélgica, y la Teoría Comunicativa de la Terminología, de M. T. Cabré Castellví, en España. Estas tres nuevas teorías comparten las críticas expuestas anteriormente y se alejan del trabajo prescriptivo del Wüster y de la lingüística estructuralista y ofrecen nuevas líneas de investigación.

En palabras de Sager, (1993: 93):

La teoría terminológica moderna acepta la presencia de expresiones sinonímicas y de variantes de términos y rechaza la actitud prescriptiva estricta del pasado, que asociaba un concepto con un solo término. Finalmente, se ha reconocido que un concepto puede tener tantas representaciones lingüísticas como situaciones comunicativas distintas existen y que requiere formas lingüísticas diferentes. La Terminología ahora adopta un planteamiento de la recopilación de datos léxicos a partir de un conjunto de documentos. Los términos, al ser estudiados en el contexto de las situaciones comunicativas, no se ven ya como elementos individuales en los diccionarios o como parte de un lenguaje semiartificial vaciado deliberadamente de cualquiera de las funciones de otros elementos léxicos. La tendencia creciente a analizar la Terminología en su contexto comunicativo, es decir, el contexto lingüístico, lleva a un número de suposiciones teóricas nuevas y también nuevos métodos de compilación y representación.

A continuación, mostraremos brevemente los aspectos característicos de cada una de estas tres teorías terminológicas.

1.3.2.2.1. La Socioterminología

La Socioterminología comienza a tomar fuerza como línea de investigación a principios de la década de los noventa en Francia y Canadá, y sus principales

precursores han sido Gambier (1991), Gaudin (1993) y Boulanger (1995). Las primeras publicaciones claramente explícitas sobre esta teoría aparecen en la revista *Cahiers de Linguistique Sociale* (1995), de la Universidad de Rouen, de la mano de L. Guespin, Y. Gambier y F. Gaudin.

Esta teoría apuesta por un enfoque descriptivo que contemple la Terminología dentro del contexto de uso real de la lengua, aunque defiende la prescripción en la elaboración de diccionarios. Siguiendo en la línea de los comentarios críticos anteriores, esta propuesta rechaza el principio de univocidad concepto-término de la TGT e incorpora estudios sobre la sinonimia y la polisemia; defiende la realización de estudios diacrónicos de los términos y, a su vez, aboga por la existencia de un *continuum* de los campos de especialidad y rechaza, así, la delimitación absoluta y homogénea de los dominios propuestos por la TGT.

Asimismo, esta corriente critica la falta de fundamentos sociales de la TGT sobre la base de los trabajos en análisis del discurso político y en sociolingüística y considera imprescindible incorporar el estudio de la producción oral en la investigación terminológica. En este contexto, se opone a la limitación a los textos escritos del trabajo terminológico. Esta defensa por el empleo de la producción oral surge como rechazo a la idea tradicional de que toda buena terminología ha de reducirse a la dimensión escrita, ya que se considera el único medio capaz de asegurar el buen uso planificado y controlado (Boulanger, 1995: 202). En definitiva, la Socioterminología defiende un equilibrio entre la dimensión escrita y la dimensión oral, ya que son los dos los canales a través de los cuales se produce comunicación profesional.

1.3.2.2.2. La Teoría Comunicativa de la Terminología (TCT)

Como alternativa a la Terminología tradicional, M. T. Cabré Castellví presenta una nueva teoría (TCT) para rechazar o completar algunos de los postulados presentados por la Terminología clásica. Su principal objetivo consiste en elaborar una teoría de la Terminología que atienda a la dimensión comunicativa de las unidades terminológicas. Según palabras de la propia autora, esta propuesta se define de la siguiente manera:

Esta propuesta pretende dar cuenta de los términos como unidades singulares y a la vez similares a otras unidades de comunicación, dentro de un esquema global de representación de la realidad, admitiendo la variación conceptual y denominativa, y teniendo en cuenta la dimensión textual y discursiva de los términos (Cabré Castellví, 1998e: 99).

En este sentido, nos encontramos de nuevo con comentarios críticos similares a los expuestos con anterioridad, a saber: el rechazo al principio de univocidad denominación-concepto, interés por la dimensión comunicativa de los términos y por el estudio de estos en un contexto real, entre otras cosas.

En su publicación «Una nueva teoría de la Terminología: de la denominación a la comunicación» (1998d: 122 y ss.), Cabré Castellví presenta sus comentarios críticos acerca de diferentes aspectos de la TGT, especialmente acerca de su reduccionismo, idealismo y afán normalizador, y recoge los principios y fundamentos básicos de su nueva teoría. A modo de resumen, indicaremos algunas de sus premisas principales a continuación:

- La Terminología se considera un campo interdisciplinar que integra aportaciones de tres teorías diferentes: la *teoría del conocimiento*, para explicar los tipos de conceptualización que se pueden dar y la relación de los conceptos entre sí y con sus posibles denominaciones; la *teoría de la comunicación*, para describir los tipos de situaciones que pueden producirse y explicar las características, posibilidades y límites de los diferentes sistemas de expresión de un concepto y de sus unidades; y la *teoría del lenguaje*, para dar cuenta de las unidades terminológicas dentro del lenguaje natural, teniendo en cuenta que participan de todas sus características, pero singularizando su carácter terminológico y explicando cómo se activa este en la comunicación.
- El objeto de estudio son las unidades terminológicas propiamente dichas, las cuales forman parte del lenguaje natural. Los términos no son unidades autónomas que forman un léxico especializado diferenciado, sino que se describen como unidades denominativo-conceptuales dotadas de capacidad de referencia, que pueden ejercer funciones distintas y que, integradas en el discurso, pueden constituir núcleos predicativos o argumentos de los

predicados. Estas unidades, que no son inicialmente ni palabras ni términos, sino solo candidatos a términos, pueden pertenecer a ámbitos distintos. El carácter de término se activa en función de las condiciones pragmáticas de adecuación a un tipo de comunicación, del tema, del ámbito, del tipo de texto y de los actores que intervienen en la situación comunicativa (emisor-destinatario).

- Los términos son unidades léxicas de forma (denominación) y contenido (significado) en las que el contenido es simultáneo a la forma. Un contenido puede ser expresado con mayor o menor rigor por otras denominaciones del sistema lingüístico (y constituiría así nuevas unidades lingüísticas de contenido especializado relacionadas semánticamente con la primera) o por denominaciones de otros sistemas simbólicos (que conformarían unidades no lingüísticas de contenido especializado).

- Los conceptos de un mismo ámbito especializado mantienen entre sí relaciones de diferente naturaleza. El conjunto de estas relaciones entre los conceptos constituye la estructura conceptual de la materia. Así, el valor de un término se establece por el lugar que ocupa en una estructuración conceptual determinada, ya que puede participar en más de una con el mismo o diferente valor.

- El objetivo de la terminología teórica es el de describir formal, semántica y funcionalmente las unidades que pueden adquirir valor terminológico, dar cuenta de cómo lo activan y explicar sus relaciones con otros tipos de signos del mismo o distinto sistema. El objetivo de la rama aplicada consiste en recopilar las unidades de valor terminológico en un tema y situación determinados y establecer sus características de acuerdo con esta situación. Esta labor puede aplicarse de diversas formas, entre las que se encuentra la normalización. En todas ellas se activa la doble función de los términos: la representación del conocimiento especializado y su transferencia, aunque en grados y modos distintos, y en situaciones también diversas.

En estos cinco puntos, Cabré Castellví nos señala los puntos más importantes de su teoría: su objeto de estudio, el objetivo y la finalidad, la aplicación y la

interdisciplinariedad de la Terminología. Con ello, observamos cómo se aleja del reduccionismo de la Terminología tradicional para aportar una Terminología más realista y completa, de la cual podemos destacar su interés por la variación denominativa y conceptual de los términos, por el carácter comunicativo de la Terminología, por la adecuación de los términos a la situación comunicativa, así como el aspecto social y psicolingüístico que presentan.

A modo de resumen, presentamos una descripción realizada por Cabré Castellví (1998e: 126):

Esta teoría, que hemos denominado Teoría Comunicativa de la Terminología, requiere que, tanto desde un punto de vista teórico como desde el metodológico, se contemple la variación lingüística en toda su dimensionalidad, se asuma la condición de adecuación de los términos y se integren los aspectos psicolingüísticos implicados (compartidos con la perspectiva social). [...] el conocimiento especializado como los textos especializados, como las unidades terminológicas pueden darse a diferentes niveles de especialización y describirse en distintos niveles de representación. Solo así, los términos pueden explicarse en toda su realidad comunicativa y representacional.

1.3.2.2.3. La Teoría Sociocognitiva de la Terminología (TST)

La Teoría Sociocognitiva de la Terminología (TST), cuya máxima exponente es R. Temmerman (2000), surge también como alternativa a los principios de la TGT, los cuales los considera irreales por concentrarse solo en la estandarización y por obviar la influencia comunicativa y sociocognitiva en el trabajo terminológico. A través de un estudio empírico sobre los procesos de categorización y lexicalización de un corpus de publicaciones sobre ciencias biológicas, la autora establece su teoría (TST) y propone una metodología inspiradas en la hermenéutica y en la semántica cognitiva.

Al igual que la Socioterminología y la TCT, la TST comparte las críticas expuestas con anterioridad, es decir, rechaza el principio de univocidad concepto-denominación, defiende el estudio de la polisemia y sinonimia, así como el estudio diacrónico de los términos, rechaza la idea de elaborar solo definiciones intensionales y extensionales y, además, mantiene la importancia de la multidimensionalidad y flexibilidad de los términos derivada de la influencia que ejerce la situación

comunicativa. Asimismo, defiende la necesidad de estudiar los términos en su uso real, por lo que el trabajo terminológico se debe basar en un estudio de documentos especializados reales.

Esta corriente también realiza novedosas aportaciones provenientes de la semántica cognitiva, que modifican totalmente los supuestos básicos de la TGT, como son la idea de concepto y el enfoque onomasiológico de la Terminología.

La principal renovación que realiza la TST (Temmerman, 1997), y donde podemos observar sus principales supuestos teóricos, está enfocada a la representación e interpretación que realiza la Terminología clásica del triángulo semántico de Gomperz (1908)³⁰ (Figura 4, izquierda). En este triángulo, se exploran básicamente las relaciones existentes entre la realidad (el mundo), el medio para comunicar esa realidad (el lenguaje) y el centro de razonamiento sobre la comprensión del mundo y del lenguaje (la mente humana).

En la propuesta de Wüster,³¹ se observa gran influencia del estructuralismo saussureano y del objetivismo imperante en su teoría. En la concepción de su triángulo (Figura 4, derecha), ignora gran parte de la interacción existente entre los elementos de cada vértice, ya que se centra simplemente en la actividad normalizadora de la Terminología.

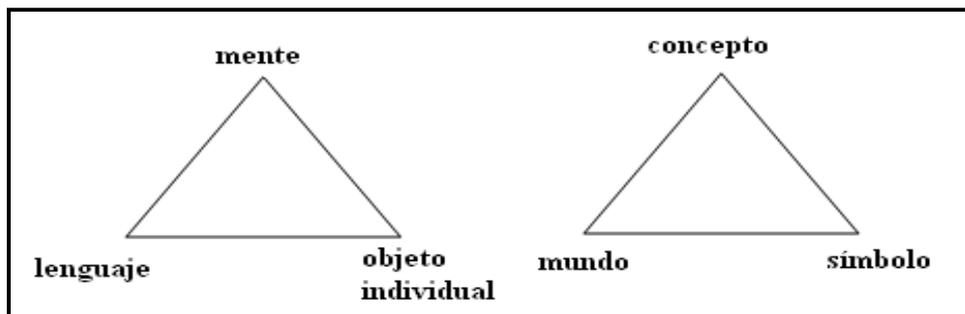


Figura 4. A la derecha, triángulo semántico de Gomperz. A la izquierda, triángulo adaptado por Wüster.

El mundo queda reducido a lo puramente objetivo; el lenguaje solo se considera como una capacidad de denominación, y la mente humana, como capaz de clasificar los

³⁰ Citado por Felber (1984: 100).

³¹ Citado y traducido por Felber (1984: 100).

objetos individuales según las características comunes de todos los miembros de la clase de un concepto.

Temmerman (ibíd.) presenta su triángulo basado en la semántica cognitiva, donde incorpora el aspecto comunicativo y social de la Terminología al triángulo de Wüster. Para ella, el mundo no existe de forma objetivista, sino que existe a través de las percepción sensoriales y el razonamiento del ser humano, y se transfiere a otros usuarios mediante el lenguaje, que es el medio a través del cual realizamos la categorización en un contexto social.

| | Terminología clásica | Terminología sociocognitiva |
|-------------------------------|--|--|
| El mundo y el lenguaje | El mundo existe de forma objetiva y se puede nombrar. | El lenguaje tiene un papel relevante en la comprensión del mundo. |
| El mundo y la mente | La mente humana puede entender el mundo gracias a su capacidad de clasificación. | El mundo se encuentra (en parte) en la mente humana. |
| El lenguaje y la mente | El potencial creativo del lenguaje no se debe tener en cuenta. | La comprensión del lenguaje no se puede separar de la comprensión del mundo. |

Tabla 3. Diferencias entre la Terminología clásica y la TST. Extraído de Temmerman (1997: 56).

Partiendo de estas ideas iniciales, la TST presenta sus principios nuevos que varían y amplían la visión de la TGT.

En primer lugar, esta propuesta rechaza la aproximación “de significado” (*‘meaning’ approach*) que propone la TGT, ya que denota el aspecto onomasiológico de la teoría, y la reemplaza por una aproximación “de comprensión” (*understanding approach*), que implica que el término se considera el punto inicial para descubrir la atribución de la categoría, es decir, se decanta por un enfoque semasiológico:

Traditional Terminology takes the concept, which is said to be the meaning of the term as its starting point. We prefer to replace it by an “understanding” approach which implies that the term is considered the starting point for discovering categorical attribution (Temmerman, 2000: 39).

En su opinión, la Terminología solo se puede estudiar en el discurso donde aparece y, por lo tanto, mantiene que el término debe ser el punto de partida en la descripción terminológica y no lo que tradicionalmente se ha venido a entender por concepto.

Por otro lado, la idea de los conceptos en la TGT son sustituidos en esta teoría por lo que se denomina «unidades de comprensión» (*units of understanding*). Según la autora, los conceptos no existen como entidades aisladas e independientes y, por tanto, objetivas, sino que existen gracias a los textos donde los distintos autores exponen su forma de entender unas categorías dentro de un determinado *Modelo Cognitivo Idealizado* (Lakoff, 1987)³² que puede diferir del de otro autor y de la interpretación realizada por el lector. En estos modelos idealizados, se relacionan entre sí diferentes unidades de comprensión estructuradas prototípicamente y no mediante reglas lógicas u ontológicas de un sistema conceptual. Estas unidades prototípicas son las que se consideran categorías.³³ Así, la categoría se presenta como la estructura más idónea para describir la unidad de comprensión. No obstante, algunas unidades de comprensión no presentan una estructura prototípica, por lo que se les consideran conceptos, como entendía la Terminología clásica. De este modo, las unidades de comprensión que presentan una estructura prototípica serán consideradas «categorías» y, por el contrario, las que no presenten esa estructura serán «conceptos».

Otra de las críticas que realiza esta corriente está relacionada con el empleo del lenguaje figurado en los lenguajes de especialidad. Wüster obviaba cualquier elemento o rasgo de este tipo de lenguaje por no considerarlo como parte de un lenguaje de especialidad. Sin embargo, Temmerman defiende la presencia de este tipo de lenguaje y, por tanto, sostiene la necesidad de prestarle atención. De hecho, proclama que los modelos metafóricos tienen una labor funcional de facilitar el pensamiento y la comprensión del mundo a través del lenguaje.

³² Una de las tesis fundamentales de Lakoff radica en lo que ha denominado *Idealized cognitive models* (ICM) o modelos cognitivos idealizados (MCI). Según este autor, estos modelos organizan nuestro conocimiento, nos permiten comprender el mundo, organizándolo en un conjunto de categorías que contienen prototipos y se relacionan entre sí mediante las diferencias y similitudes que presentan (Lakoff, 1987: 68). Estos modelos se patentizan a través del lenguaje. A partir de esta idea, han surgido diferentes teorías en la Lingüística Cognitiva que pretenden explicarlos: los campos semánticos de Fillmore (1982); la teoría de la metáfora y la metonimia de Lakoff y Johnson (1980), así como la gramática cognitiva de Langacker (1987). Estas teorías de la categorización están basadas en la teoría de los prototipos (*prototype theory*) de Rosch, una psicóloga especializada en psicología cognitiva. Esta teoría pretende explicar la forma en la que los seres humanos se enfrentan de forma cognitiva a las percepciones del mundo. Para la autora, el ser humano tiene un prototipo, un buen ejemplo de cada categoría, en su mente y la categorización la realiza con relación a las similitudes y diferencias con ese prototipo (cf. 1.3.2.3.3.).

³³ La categorización se diferencia de la clasificación lógica y ontológica de la Terminología clásica en el hecho de que estas consideran el sistema conceptual como algo preexistente en la mente, previo a la denominación e independiente del lenguaje.

Esta corriente también hace mención al tipo de definición necesaria en Terminología. Según esta autora, es necesario tener en cuenta el tipo de unidad de comprensión, así como el nivel y tipo de especialización del emisor y receptor en la comunicación para seleccionar la información que se incluirá en la definición. Por ello, rechaza la exclusividad de las definiciones intensionales y extensionales de la TGT, así como la distinción entre información definicional y enciclopédica, y se decanta por prestar atención a la variabilidad entre información más esencial o menos esencial. Según esto, lo que puede resultar esencial para una situación o un dominio, o desde un punto de vista particular o una intención, puede resultar menos esencial en otras circunstancias. Para ello, propone la existencia de plantillas de comprensión (*templates of understanding*) para la descripción de significados las cuales se componen de diferentes módulos de información (propiedades, definición básica, información histórica, información procedimental) más o menos esencial, dependiendo del tipo de unidad de comprensión y de factores como la perspectiva desde la que se concibe dicha unidad o la intención del emisor del mensaje.³⁴

Además de la parte teórica de esta corriente, Temmerman también presenta una metodología alternativa basada en sus postulados, que denomina Termontografía (*Termontography*). Esta metodología cuenta con elementos comunes al resto de las corrientes modernas, como son: a) el estudio de la ocurrencia de los términos y conceptos en un discurso especializado concreto y real, y nunca como entidades aisladas e independientes; y b) el estudio descriptivo a partir de dos parámetros determinados: por un lado, el contenido de los dominios de especialidad y, por otro, el perfil del usuario potencial del trabajo terminográfico.

No obstante, esta metodología cuenta con elementos novedosos que consisten en la combinación de la ingeniería ontológica con los principios de la TST para realizar el trabajo terminográfico. En este sentido, este modelo defiende la incorporación de las ontologías como elemento principal en los trabajos terminográficos, lo que la sitúa en las líneas de trabajo más actuales en consonancia con el auge de las nuevas tecnologías en la Terminología y de la Lingüística Computacional (cf. 2.6.).

³⁴ Véase Temmerman (2000: 226).

1.3.2.3. Corrientes teóricas de la lingüística aplicables a la Terminología

Además de los cambios sociales, económicos y tecnológicos de las últimas décadas que indicábamos con anterioridad, se han producido varios cambios científicos dentro de la lingüística, que han dado lugar también a la evolución de esta.

En primer lugar,³⁵ tenemos la expansión creciente de la Lingüística Cognitiva a partir de los años 80 y que da lugar a un cambio de perspectiva para muchas disciplinas, entre ellas la Terminología, cuya dimensión conceptual encuentra en la Lingüística Cognitiva un marco adecuado para desarrollar distintas estructuras de conocimiento y formas de representación.

En segundo lugar, encontramos el desarrollo espectacular de la lingüística basada en corpus, denominada Lingüística de Corpus,³⁶ que ha sido uno de los factores de apertura a la Terminología descriptiva. Los análisis de los datos basados en corpus permiten disponer de materiales adecuadamente seleccionados para la descripción de diferentes fenómenos lingüísticos, así como observar y formular generalizaciones basadas en estudios empíricos, entre otras cosas.

Por último, es imprescindible hablar de la Lingüística Computacional, ya que ha significado un gran avance para la lingüística en general, y para la Terminología en particular. Se trata de un campo multidisciplinar de la lingüística y la informática que se ocupa del procesamiento del lenguaje mediante ordenadores, con objeto de elaborar modelos computacionales y formales del lenguaje. Desde sus comienzos en la década de los 80, se han desarrollado numerosas aplicaciones informáticas para su cometido, que han permitido avanzar en las diferentes ramas de la lingüística de forma importante, desde la traducción automática hasta la enseñanza de lenguas y, por supuesto, la Terminología y su rama aplicada. La Lingüística Computacional está muy relacionada con la Lingüística de Corpus en su trabajo diario, ya que los corpus textuales tienen un papel primordial dentro de esta rama de la lingüística, tanto en su vertiente teórica como

³⁵ En este apartado, nos centraremos solo en los cambios que se han producido dentro de la Lingüística que están relacionados con nuestra tesis doctoral, por lo que trataremos únicamente la Lingüística de Corpus, la Lingüística Cognitiva y la Lingüística Computacional.

³⁶ También puede denominarse Lingüística de corpus, ya que en español no parece existir consenso sobre su denominación.

en su vertiente aplicada (Corpas Pastor, 2008: 10). Tanto es así que la Lingüística de Corpus se ha llegado a considerar como una rama de la Lingüística Computacional (Oostdijk, 1991: 2).

A continuación, trataremos estas ramas lingüísticas en mayor detalle.

1.3.2.3.1. La Lingüística de Corpus

La Lingüística de Corpus es una rama de la lingüística que centra sus investigaciones en ejemplos de uso de la vida real para el estudio de la lengua con la ayuda de herramientas de software diseñadas específicamente para el análisis textual. Más que una disciplina, debemos tratar la Lingüística de Corpus como una metodología de análisis empírico que ha sido, y es, de gran utilidad para otros campos de la lingüística, como son la Terminología (Budin y Wright, 1997), la Lexicografía bilingüe (Klavans y Tzoukermann, 1990), la Traductología (Baker, 1993, 1995; Corpas Pastor, 2001, 2002, 2008), la enseñanza de lenguas (Barlow, 1996), entre otros, y se ha convertido en uno de los focos principales de la lingüística actual, tanto en su vertiente teórica como práctica.

El punto de partida de la Lingüística de Corpus es siempre el corpus,³⁷ es decir, los datos reales del lenguaje, ya que su principal objetivo consiste en indicarnos cómo se utilizan las palabras realmente, como aparecen juntas con otras palabras y forman unidades de sentido con ellas; en hacernos entender cómo está construida la realidad social en el discurso, cómo cambia el sentido de las palabras, las diferencias entre el idioma hablado del idioma escrito, y en demostrarnos demuestra cómo varían los registros en los diferentes idiomas, cómo cambian de significado las palabras (Teubert y

³⁷ Existen numerosas definiciones de «corpus», pero la más aceptada hasta la fecha por los lingüistas es la propuesta por EAGLES (Expert Advisory Group on Language Engineering Standards, 1996: 4): «A corpus is a collection of pieces of language that are selected and ordered according to explicit linguistic criteria in order to be used as a sample of language» (Sinclair, 1996: 4). También hay numerosos corpus en la actualidad, pero se pueden considerar como los más conocidos los siguientes: de la primera generación de corpus, el Survey of English Usage Corpus (SEU), en 1959, de Quirk (University College, Londres); el Standard Corpus of Present-Day American English (Brown Corpus), en 1963-64, de Nelson y Kucêra; el Lancaster-Oslo/Bergen Corpus (LOB), en 1961, de Leech y Johansson (de las universidades de Lancaster y Oslo, respectivamente); y de las siguientes generaciones, el British National Corpus (BNC), The Bank of English (conocido como COBUILD corpus), del proyecto COBUILD (Collins-Birmingham University International Database), entre otros (Bowker, 1996: 28) (cf. 1.4.2.2.).

Krishnamurthy, 2007: 3), lo que se denomina el estudio de la lengua *in vivo*.³⁸ Actualmente, cuando hablamos de «corpus» nos referimos normalmente a los corpus electrónicos, es decir, los que se encuentran en formato electrónico y que permiten, por tanto, un tratamiento informático.

Aunque sus orígenes se remontan a la época estructuralista norteamericana (Bloomfield, 1933³⁹ y sus seguidores) y a la tradición anglosajona iniciada por Firth (1935, 1957a, 1957b), y continuada principalmente por sus discípulos Halliday y Sinclair,⁴⁰ no ha sido hasta la década de los ochenta cuando la Lingüística de Corpus ha adquirido su relevancia actual y el corpus se ha convertido en una herramienta indispensable para el estudio del lenguaje en todas sus vertientes. Solo entonces se ha generalizado el término de Lingüística de Corpus, especialmente a partir de 1984, cuando se publica el libro *Corpus Linguistics I: Recent Developments in the Use of Computer Corpora*, editado por Aarts y Meijs, y se empieza a tratar la Lingüística de Corpus en el sentido actual del término.

El resurgimiento de estudio empírico basado en corpus se ha debido principalmente a tres factores:

- A los comentarios críticos que se empiezan a realizar en contra de las restricciones del racionalismo y de sus principales métodos de trabajo, a saber: el uso de la introspección y de los datos intuitivos como único recurso para la formulación de principios lingüísticos sobre el conocimiento léxico, gramatical, semántico, la relación entre las palabras y su contexto, etc. (Sinclair, 1991).

³⁸ La Terminología *in vivo*, que hace referencia al estudio de los términos dentro de un contexto comunicativo real, se contrapone a la Terminología *in vitro*, que consiste en el estudio de los términos fuera de contexto, ya sea en una lista de palabras, un diccionario, un glosario, etc.

³⁹ Citado en Leech (2007: 3).

⁴⁰ Esta corriente se conoce como *Early Corpus Linguistics* (McEnery y Wilson, 1996), la cual se abandonó a finales de los años cincuenta por la llegada de los estructuralistas a la lingüística, especialmente de la teoría chomskiana. A partir de ese momento, el panorama lingüístico se inundó de los estudios lingüísticos basados en la intuición del lingüista (el racionalismo) y la Lingüística de Corpus recibió numerosas críticas, tanto teóricas (Chomsky, 1957,1965) como prácticas (Abercrombie, 1965). Para profundizar en esta época y en las críticas recibidas, véanse McEnery y Wilson (1996), Lüdeling *et ál.* (2008), entre otros.

- Al aumento de interés por una visión más descriptiva de los usos reales del lenguaje y, por tanto, al abandono de los estudios prescriptivos;
- A los avances informáticos actuales, que permiten almacenar y manejar gran cantidad de información en ordenadores, facilitando en gran medida la gestión y el análisis de los datos.

En la actualidad, parece existir un consenso generalizado sobre la utilidad del corpus electrónico como herramienta de apoyo para los estudios lingüísticos, y su utilización se ha convertido en indispensable a la hora de estudiar material real y de poner a prueba hipótesis sobre el funcionamiento y manejo de la lengua y para construir sistemas de procesamiento del lenguaje natural. Además de estas ventajas, los corpus electrónicos presentan otras que justifican también el interés actual, como son la objetividad de las comprobaciones de teorías e hipótesis; la rapidez, precisión y consistencia del procesamiento de datos; la facilidad de manejo y acceso a grandes cantidades de datos reales; la reutilización y multiplicidad de usos y aplicaciones, entre otros (Villagrande Llamazares, 2006).

No obstante, también debemos reconocer las limitaciones de utilización que presentan en algunos casos este tipo de recursos: cuando los corpus no aportan suficiente información es necesario realizar análisis manuales; no siempre es sencillo compilar un corpus para la finalidad determinada ni hacerlo de una forma totalmente objetiva, entre otras.

Pese a sus desventajas, no podemos dejar de valorar las ventajas que presenta esta herramienta para el trabajo sistemático y descriptivo de la lingüística actual. En palabras de Sinclair (1991: 4): «the ability to examine large text corpora in a systematic manner allows access to quality of evidence that has not been available before».

En el caso de la Terminografía, se observa claramente la repercusión que ha tenido el corpus electrónico en la disciplina, ya que ha permitido al terminógrafo disponer de gran cantidad de documentación especializada, que no podría disponer de otra manera, y, por tanto, basar su trabajo en un estudio empírico de usos reales de los lenguajes de especialidad de una forma objetiva y sistemática. Como veremos más adelante, el proceso de trabajo de construcción de recursos terminológicos ha cambiado completamente gracias a la aplicación de los corpus al trabajo terminográfico, lo que ha

dado lugar a una mejora sustancial de la calidad de los recursos terminológicos creados en la actualidad así como del trabajo realizado por el terminógrafo, que ha visto agilizadas y facilitadas sus tareas la mayoría de sus tareas.

1.3.2.3.2. La Lingüística Computacional

La Lingüística Computacional es un campo interdisciplinario vinculado a la lingüística y a la informática que se encarga del procesamiento del lenguaje mediante sistemas informáticos. Desde sus inicios con la traducción automática hace alrededor de cincuenta años,⁴¹ la Lingüística Computacional ha crecido y se ha desarrollado exponencialmente, expandiéndose en su parte teórica mediante el desarrollo de modelos formales e informáticos del lenguaje, a la par que ha incrementado el abanico de sus aplicaciones útiles.

Dentro de la Lingüística Computacional, se pueden distinguir principalmente dos ramas principalmente: por un lado, la Lingüística Computacional teórica, que engloba la parte teórica de la disciplina y, por tanto, sus fundamentos y principios teóricos y metodologías; y por otro, la Lingüística Computacional aplicada, conocida también como Ingeniería lingüística, que se trata de la aproximación más tecnológica, centrada en el diseño y el desarrollo de sistemas informáticos capaces de comprender, producir y traducir enunciados orales y escritos en cualquier lengua natural, en definitiva, de la aplicación de herramientas informáticas para el estudio lingüístico (Lavid, 2005: 76). De esta manera, nos la presenta también Kay (2005: 429), una de las figuras actuales más relevantes del campo de la Lingüística Computacional:

⁴¹ Los orígenes de la Lingüística Computacional se sitúan a finales de la Segunda Guerra Mundial, cuando los servicios de inteligencia y las fuerzas armadas de los Estados Unidos y de la Unión Soviética comenzaron a trabajar en diversos proyectos para elaborar programas de traducción automática entre el inglés y el ruso. Este interés por la traducción automática duró hasta mediados de la década de los sesenta, cuando la Automatic Language Processing Advisory Committee de la Academia nacional de las ciencias publicó un informe en el que se describían los mínimos resultados obtenidos hasta ese momento. Como consecuencia, disminuyeron drásticamente los fondos para las investigaciones en traducción automática y esta quedó limitada a unos pocos proyectos en Europa y Asia. A partir de ese momento, la Lingüística Computacional se centró fundamentalmente en investigaciones científicas relacionadas con el lenguaje y su procesamiento informático (Kay, 2003: XVII). No obstante, la traducción automática retomó cierta importancia en la década de los ochenta.

Computational linguistics is not a natural language processing. Computational linguistics is trying to do what linguistics do in a computational manner, not trying to process text, by whatever methods, for practical purposes. Natural language processing, on the other hand, is motivated by engineering concerns. I suspect that nobody would care about building probabilistic models of language unless it was thought that they would serve some practical end.

El término «Ingeniería lingüística» ha sufrido diversos cambios de denominación durante su evolución: «Industrias de la lengua», «Ingeniería Lingüística», «Tecnologías del Lenguaje Humano» y «Tecnologías lingüísticas», que ha tenido lugar dentro de los Programas Marco de la Unión Europea (Moreno Sandoval, 2009: 102).⁴² En esencia, estos cambios terminológicos no han venido acompañados por cambios sustanciales de objetivos y aplicaciones de esta disciplina, a excepción de la integración del estudio del texto y del habla en una misma perspectiva. En otras palabras, desde sus orígenes, la investigación y el desarrollo en tecnologías de habla y de texto han seguido caminos aislados hasta los años 90, en los que ambas líneas de investigación se encontraron auspiciados por el término «Tecnologías del lenguaje» o «Tecnologías lingüísticas».

Como nos indican Corpas Pastor (2008: 10) y Orăsan et ál. (2007), el corpus ha tenido un papel principal en prácticamente todos los ámbitos de las dos partes de la Lingüística Computacional y ha traído consigo consecuencias muy importantes para su desarrollo. Este recurso ha tenido una enorme repercusión en la creación de sistemas de análisis morfosintáctico, lexicones computacionales, correctores, diccionarios computerizados, entre otras muchas aplicaciones, como nos lo indican Atkins et ál. (1992: 1):

There has been over the past few years a tremendous growth in interest and activity in the area of corpus building and analysis. European, USA, and Japanese efforts in the development of NLP and IT are converging on the recognition of the importance of some sort of corpus-based research as part of the infrastructure for the development of advanced language processing applications.

⁴² La Unión Europea considera las Tecnologías del lenguaje como una prioridad política y se ha convertido en la principal institución financiadora de los proyectos en Europa. En concreto, considera que las tecnologías del lenguaje “son hoy en día un instrumento esencial de conservación del rico patrimonio cultural europeo y, al mismo tiempo, una fuente del crecimiento económico futuro” (<http://europa.eu/languages/es/chapter/17>). Todas las URLs mencionadas en el presente trabajo estaban operativas a fecha de 13 de abril de 2011.

También McEnery (2003: 460), nos destaca el trabajo colaborativo entre la Lingüística de Corpus y la Lingüística Computacional:

Corpora have played a useful role in the development of human language technology to date. In return, corpus linguistics has gained access to ever more sophisticated language processing systems. There is no reason to believe that this happy symbiosis will not continue –to the benefit of language engineers and corpus linguistics alike– in the future.

El uso de los corpus en esta disciplina comenzó en los años sesenta, aunque solo se empezaron a emplear a gran escala a partir de la década de los noventa. Para ello, se observan dos razones principalmente, que son muy próximas a las que motivó la recuperación de los corpus en la Lingüística de Corpus, a saber: por un lado, los avances en los ordenadores permitieron a los investigadores trabajar con corpus de grandes dimensiones sin necesidad de utilizar ordenadores muy especializados, lo que provocó a su vez que el número de corpus aumentara de forma importante. Por otro lado, el aumento de interés en estos recursos vino motivado por un cambio de paradigma en el ámbito de la inteligencia artificial a finales de los ochenta y principio de los noventa, cuando los investigadores vieron las limitaciones de los sistemas basados en el conocimiento y empezaron a utilizar métodos que estaban basados en la evidencia empírica.

No obstante, no solamente la Lingüística Computacional se ha beneficiado de los recursos de la Lingüística de Corpus, sino que se ha producido una simbiosis entre estas dos corrientes lingüísticas en las últimas décadas. Ambas disciplinas han evolucionado y se han desarrollado gracias a los recursos y aplicaciones de la otra, lo que ha llevado a considerarlas prácticamente como imprescindibles la una para la otra. La Lingüística de Corpus no hubiera podido crecer si no hubiera sido por las aplicaciones y recursos desarrollados en la Lingüística Computacional, puesto que, para poder trabajar correctamente con corpus textuales, es necesario que estos sean tratados de manera informática con objeto de poder extraer la información que se desee de ellos. Como afirma Leech (1991: 13):

[...] the availability of vastly increasing computer corpus resources has enabled syntactic and lexical phenomena of a language to be open to empirical investigation on a scale previously unimagined.

Y además:

It is widely acknowledged today that a corpus needs the support of a sophisticated computational environment, providing software tools both to retrieve data from the corpus and to process linguistically the corpus itself (ibid.: 22).

Este tratamiento informático dependerá de la finalidad del corpus en cuestión, así como de las hipótesis que se deseen comprobar con este recurso y, por tanto, el lingüista deberá decidir qué herramientas informáticas necesita aplicar al corpus en cada caso para así poder trabajar con él de la mejor manera.

En el caso de la Terminografía, las herramientas informáticas desarrolladas por la Lingüística Computacional también han significado un gran avance en el trabajo terminográfico, ya que, gracias a estos recursos, se ha vuelto más sencilla la labor de buscar y extraer términos, equivalentes, contextos, etc. en los corpus electrónicos, así como de elaborar definiciones, ejemplos, etc. En este contexto, se podrían utilizar herramientas como, por ejemplo, extractores de candidatos a término, anotadores sintácticos, lematizadores, extractores de relaciones semánticas, analizador de concordancias, entre otros, aunque la mayoría de estas herramientas se encuentran todavía en fase de desarrollo y mejora para su aplicación óptima en el campo de la Terminografía. En el siguiente capítulo, nos detendremos en algunas de estas herramientas y en su aplicación terminográfica.

1.3.2.3.3. La Lingüística Cognitiva

La perspectiva cognitiva surge en la década de 1970 y se desarrolla activamente a partir de los años 80 como respuesta al descontento de algunos generativistas⁴³ que defendían la necesidad de una semántica que tuviera presente al individuo y sus habilidades cognitivas. En esos años, se produce un cambio de orientación en las investigaciones psicolingüísticas sobre la cognición humana que dan lugar a una

⁴³ Se considera que la Lingüística Cognitiva nació en 1987, con la publicación de dos libros fundamentales en esta rama: en primer lugar, *Women, Fire and Dangerous Things*, de G. Lakoff, y en segundo lugar, *Foundations of Cognitive Grammar: Theoretical Prerequisites*, de R. Langacker. Ambos autores son considerados como los padres de esta corriente. No obstante, su gestación comienza en los años 1970 (Cuenca *et ál.*, 1999: 11).

evolución en la disciplina. Debido a este cambio, se pasa de considerar la comprensión como una simple descodificación secuencial de los elementos lingüísticos a tratarla como un componente predictivo (por el contexto y la información previamente procesada) de la reacción del hablante en una situación comunicativa determinada.

La Lingüística Cognitiva se trata de un área del saber interdisciplinar en la que confluyen los intereses de varias ramas del saber, fundamentalmente la lingüística, la psicología, la neurología y la inteligencia artificial. Esta disciplina pretende dar cuenta del conocimiento lingüístico poniéndolo en relación con otros procesos cognitivos de la mente humana, tales como la percepción, la memoria y la atención. Según la corriente actual, los códigos y estructuras de la lengua se aprenden, se almacenan en la memoria y se recuperan de manera muy similar a otros tipos de conocimientos que poseen las personas. En otras palabras, la adquisición de la lengua sigue los mismos patrones de aprendizaje que otras habilidades cognitivas complejas. Este planteamiento teórico se opone totalmente a la teoría de la gramática generativista, ya que, según se aduce, el conocimiento lingüístico no difiere en esencia de otros tipos de conocimientos que poseen los humanos y no es, por tanto, preciso postular que hay un módulo del cerebro especializado en el procesamiento de la lengua, ni que existe una gramática innata que el niño adapta a la lengua que aprende durante sus primeros años de vida. La mayoría de sus investigaciones han estado centradas en la semántica, aunque también se han dedicado a la sintaxis y la morfología. En nuestro caso, trataremos solo la parte relacionada con la semántica, ya que es el tema influyente en la Terminología.

El término Lingüística Cognitiva, sin embargo, no describe una única teoría bien delimitada y articulada, como veremos a continuación, sino que han adoptado una serie de compromisos y principios básicos que han dado lugar a un abanico de teorías complementarias y coincidentes que, en ocasiones, pueden resultar incluso contradictorias.⁴⁴

De entre estos aspectos compartidos por todas las teorías cognitivas, podemos destacar el tratamiento que se le da al «lenguaje» en la Lingüística Cognitiva. Esta lo

⁴⁴ En el volumen editado por V. Evans, B. K. Bergen y J. Zinken, *The Cognitive Linguistics Reader*, (2007), se pueden consultar las teorías más influyentes de la Lingüística Cognitiva actual y su desarrollo, así como los principios básicos comunes a todas las teorías.

considera como un instrumento para la organización, procesamiento y transmisión de información (Geeraerts, 2006: 3), lo que lo convierte en una herramienta básica durante el proceso cognitivo. Además, a partir de esta concepción, se pueden determinar cuáles son los principios fundamentales de la Lingüística Cognitiva (Langacker, 1987; Cuenca y Hilferty, 1999: 181):

- El lenguaje no es una facultad cognitiva autónoma, sino que se relaciona con otras facultades cognitivas humanas. Las estructuras lingüísticas están íntimamente conectadas con el conocimiento y el pensamiento y deben entenderse en relación con la función comunicativa del lenguaje. En este sentido, la Lingüística Cognitiva intenta dar cuenta de cómo interactúan el cuerpo, la mente y el lenguaje.
- El lenguaje tiene un carácter difuso y dinámico, que puede variar con el uso, es decir, dependiendo de la situación comunicativa, de los usuarios, de la época en la que se realice la comunicación, etc. el lenguaje tendrá una serie de particularidades semánticas y sintácticas que no se pueden distinguir de forma clara y nítida.
- No es posible establecer una división clara y exacta entre los aspectos lingüísticos y los no lingüísticos, entre lo denotativo y lo connotativo, es decir, entre la semántica y la pragmática. Ni tampoco entre los diferentes niveles del lenguaje (semántica y pragmática, semántica y gramática, gramática y léxico), ya que se impone una caracterización dinámica del lenguaje que difumina los límites.

Asimismo, se puede considerar como compartido la concepción de «significado lingüístico» que se defiende en esta disciplina. Según Geeraerts (2006: 4), el significado lingüístico:

- Es subjetivo, ya que el significado no es un reflejo objetivo del mundo, sino una forma de modelar ese mundo. En otras palabras, no existe una línea divisoria entre el conocimiento lingüístico y el conocimiento del mundo, ya que depende de la interpretación del sujeto.

- Es dinámico y flexible, ya que está relacionado con el modelado de nuestro mundo, que va cambiando.
- Es enciclopédico pero no autónomo, ya que el significado que construimos dentro y a través del lenguaje no es independiente del módulo de la mente, sino que refleja una experiencia general de las personas.
- Está basado en el uso y en la experiencia.

Más allá de los aspectos relacionados con la adquisición, la Lingüística Cognitiva se ha interesado también por la explicación de fenómenos lingüísticos en sí mismos. En especial, se ha preocupado por problemas relativos a la categorización conceptual y por los modos en los que se relaciona el conocimiento específicamente lingüístico con el conocimiento del mundo. Es aquí donde encontramos gran variedad de teorías cognitivas, que pretenden dar cuenta de las claves necesarias para la organización y distinción de conceptos.

Antes de continuar con niveles superiores de organización, es importante entender cuál es el primer paso de cualquier representación de conocimiento: la categorización, y cuáles son las principales teorías al respecto.

Según Croft y Cruse (2004: 74), la categorización es una de las actividades cognitivas más básicas y la definen de la siguiente manera:

Categorization involves the apprehension of some individual entity, some particular of experience, as an instance of something conceived more abstractly that also encompasses other actual and potential instantiations.

Otra posible definición que ofrece Cohen y Lefebvre (2005: 3) es la siguiente:

Categorization, the basic cognitive process of arranging objects into categories, is a fundamental process in human and machine intelligence and is central to investigations and research in cognitive science.

Pese a la similitud de estas definiciones y del acuerdo que existe a la hora de entender qué es «categorización», no existe aún consenso sobre el modo en el que categorizamos, es decir, en la forma en la que ordenamos el conocimiento en nuestra mente, ni sobre la estructura interna que presentan las categorías resultantes.

El modelo clásico, conocido como aristotélico, considera las categorías conceptuales como entidades cognitivas fijas que se asocian de forma estable con las expresiones lingüísticas. Esta corriente ha sido la opinión predominante en Lingüística Cognitiva y en la semántica estructural; sin embargo, recientemente están surgiendo opiniones diferentes que defienden la dinámica de los conceptos, según la cual se crearían en el momento de uso. Esta última corriente es la que seguiremos en nuestro trabajo de investigación, pero antes, nos parece adecuado proceder a revisar brevemente este modelo clásico, que ha tenido gran influencia en el desarrollo de la Lingüística Cognitiva.

El modelo clásico define las categorías conceptuales como un conjunto de características necesarias y suficientes y, por lo tanto, definidas y delimitadas. Las características son necesarias en el sentido de que las entidades que no posean el conjunto completo no se pueden considerar como pertenecientes a una categoría; y son suficientes en el sentido de que la posesión de todas las características garantiza la membresía a esa categoría. En otras palabras, se debe llevar a cabo un análisis de los atributos y características de un concepto para poder clasificarlo en una categoría u otra, como en el ejemplo de STALLION [EQUINE, MALE]; MARE [EQUINE, FEMALE] (Croft y Cruse, 2001: 7). Esta propuesta se encuentra muy próxima a los postulados de la Escuela de Viena y de la TGT, ya que considera que los conceptos solo pueden pertenecer a una única categoría y, por tanto, presentan un carácter unívoco e invariable.

Al igual que aparecieron comentarios críticos para la TGT por sus restricciones y su alejamiento de la realidad y de la situación comunicativa, se han realizado también críticas similares a este modelo clásico. En primer lugar, en muchos casos no se dispone de una definición adecuada en términos de características suficientes y necesarias para la adscripción de ciertos conceptos con sus correspondientes categorías, ni todo el mundo coincide en otorgar las mismas características a los conceptos. En segundo lugar, no se pueden considerar en el mismo nivel a todos los conceptos, ya que existen ciertos conceptos más representativos que otros dentro de una misma categoría. Y, por último, el modelo clásico no ofrece explicación alguna al hecho de que los límites entre categorías sean a menudo vagos y difusos.

Como oposición al modelo clásico, surgió en los años 1970 la teoría de los prototipos (*prototype theory*), de Rosch (1973, 1975, 1978), que se ha convertido en la teoría más influyente en la Lingüística Cognitiva con relación a las categorías conceptuales. Esta teoría se fundamenta en el hecho de que no todos los miembros de una misma categoría muestran el mismo *status* ni esta se puede definir a partir de condiciones necesarias y suficientes comunes a todos sus miembros. En otras palabras, no considera que la existencia de propiedades comunes compartidas por todos los miembros sea condición necesaria para el establecimiento de una categoría (Kleiber, 1995: 17). En este sentido, el proceso de categorización ya no se limita al descubrimiento de una regla de clasificación, sino a señalar diferencias y similitudes globales y la formación de prototipos, entendiéndose como tales los ejemplares que son más reconocidos, más representativos e idóneos de una categoría (Cuenca y Hilferty, 1999: 35). Así pues, el establecimiento de categorías por parte de los hablantes conlleva el establecimiento de prototipos, a partir de los cuales se clasifican jerárquicamente los demás elementos integrantes de las categorías. De esta manera, mientras que para el modelo clásico todos los miembros de una categoría presentan las mismas características (características necesarias y suficientes), es decir, todos los pájaros son igual a <pájaro>⁴⁵ una vez han sido incluidos en esa categoría, para la semántica de los prototipos unos elementos de la categoría son más «pájaros» que otros (Kleiber, 1995: 18).

Aunque el modelo de Rosch en líneas generales es aceptado y valorado como hipótesis de trabajo central en categorización, se han detectado algunas limitaciones en esta teoría. Una de ellas es el hecho de no tener en cuenta el contexto y de continuar basándose en listas de propiedades, aunque no tan estrictas como las del modelo clásico. Consecuentemente, se aleja de la realidad comunicativa y obvia información extralingüística del contexto relevante para la organización de conceptos en categorías.

A partir de esta teoría, han surgido otras, como la teoría de los ejemplares (Smith y Medin, 1981) o la teoría de “la teoría” (Murphy y Medin, 1985), que intentan solventar este problema de la existencia de un elemento ideal que cumple todas las

⁴⁵ En nuestro trabajo, utilizaremos las comillas simples para identificar los conceptos, siguiendo las indicaciones de la Norma ISO 704:2000 (2000: viii).

características de la categoría y tener más en cuenta el contexto donde aparecen los conceptos, pero hasta el día de hoy no han podido subsanar los problemas que presentan por su parte.

Además de estas teorías que tratan el primer nivel de la comprensión, es decir, la categorización, en la Lingüística Cognitiva encontramos otras teorías que pretenden explicar la forma de organizar el conocimiento, de comprender el mundo a través de relaciones establecidas entre diferentes categorías y conceptos. Como ejemplos, podemos citar la teoría de los *Idealized cognitive models* (ICM) o modelos cognitivos idealizados (MCI), de Lakoff (1987); la teoría de los esquemas (*Schema Theory*), de Schank y Abelson (1977); la Semántica de Marcos (*Frame Semantics*), de Fillmore (1977, 1982, 1985), entre otras. Estas teorías tienen en consideración el contexto como elemento que determina la organización de una situación determinada, así como la dinámica de las estructuras conceptuales creadas entre las categorías.

En general, estas teorías han tenido gran influencia en la Terminología moderna, especialmente en lo que respecta a la conceptualización de la realidad dentro de un contexto comunicativo determinado, tanto con respecto a la categorización de los conceptos (primera fase de comprensión) como a la organización del conocimiento (fases superiores). Faber destaca la importancia de la representación de conocimiento en Terminología en la siguiente cita (2009: 289):

La posibilidad de encontrar la estructura conceptual reflejada en la lengua ha hecho que la investigación en representación semántica adquiriera un protagonismo del que hasta ahora no había gozado.

Tanto es así que recientemente se han incorporado a la práctica de la Terminología métodos de representación del conocimiento basados en las teorías cognitivas, como son las ontologías,⁴⁶ inicialmente procedentes de la filosofía y de la ingeniería informática. Estos elementos han adquirido mucha relevancia en los últimos años, ya que se consideran puentes entre la representación del conocimiento en la mente y el lenguaje.

⁴⁶ La cuestión de las ontologías se verá en profundidad más adelante en este capítulo.

Como teoría que ha influido en la Terminología moderna para la representación del conocimiento, y especialmente en este trabajo de investigación, trataremos a continuación la teoría de la Semántica de Marcos, de Fillmore, con objeto de presentar sus características básicas y los elementos útiles que han servido para crear una versión más coherente de relaciones entre categorías y de la organización del mundo para la Terminología.

1.3.2.3.3.1. La Semántica de Marcos

La Semántica de Marcos (*Frame Semantics*) tiene como objetivo enfatizar las relaciones entre el lenguaje y la experiencia y proporciona un marco (*frame*) para presentar los resultados de dicha investigación, teniendo en cuenta que una palabra se considera como representante de una categoría conceptual (cf. Petruck, 1996: 1). Según esta autora (ibíd.):

A frame is any system of concepts related in such a way that to understand any one concept it is necessary to understand the entire system; introducing any one concept results in all of them becoming available.

Fillmore (1968)⁴⁷ fue uno de los primeros en introducir el concepto de marco (*frame*) en la lingüística, considerándolo como un instrumento lingüístico para representar eventos extralingüísticos. Este autor propone que un «marco semántico» (*semantic frame*) es un esquema de la experiencia, es decir, una estructura de conocimiento que se representa en el nivel conceptual y que se retiene en la memoria a largo plazo (Fillmore, 1977). Los autores Atkins y Rundell (2008: 145) indican de manera más explícita qué es un *marco semántico*:

A semantic frame is a schematic representation of a situation type (e.g. speaking, eating, judging, moving, comparing, etc. -activities and situation which make up our everyday life) together with a list of the typical participations, props, and concepts that are to be found in such a situation; these are the semantic roles, or "frame elements" (FEs).

⁴⁷ Citado en Petruck (1996: 2).

Tanto Fillmore y Atkins (1992) como Minsky (1977) coinciden en que se trata de un conocimiento estereotipado (situación prototípica) que es necesario para entender conceptos y significados. Asimismo, los marcos semánticos reflejan el modo en el que las personas almacenan la información que reciben según la cultura en la que se encuentran o su conocimiento del mundo, lo que hace que esta teoría tenga gran relevancia tanto en los estudios de traducción como en la Terminología y la Terminografía, como veremos en el siguiente capítulo. Por ejemplo, en el caso del término *breakfast*, en inglés, es necesario entender las prácticas comunes de la cultura donde se utiliza este término (o categoría) para poder comprender su significado en contexto correctamente y crear el marco prototípico correspondiente, ya que no hace referencia a la misma realidad cultural si se emplea en Reino Unido que en Estados Unidos o, con sus equivalentes, en España o Alemania. En cada situación, estará vinculada a una situación prototípica concreta y los sujetos esperarán un determinado contexto.

Por tanto, en la categorización y estructura conceptual de un dominio como es el del turismo de aventura, donde el aspecto cultural es tan importante (Durán Muñoz, 2008a) y en el que los conceptos se encuentran en constante movimiento, tanto por nueva creación como por cambios intercategoriales, una representación del conocimiento en forma de marcos, entendiéndolos como principios de organización del léxico, parece ser la más apropiada y completa, ya que permite un análisis de los datos del corpus más sistemático y menos subjetivo, a la par que muestra las características culturales y pragmáticas de las unidades objeto de estudio y permite plasmar las relaciones semánticas entre los conceptos y el contexto determinado en el que tiene lugar la comunicación.

Actualmente, se está poniendo en práctica esta teoría de marcos semánticos mediante el proyecto Berkeley FrameNet Project,⁴⁸ que se trata de un proyecto del International Computer Science Institute en Berkeley (Estados Unidos). El objetivo de este proyecto consiste en documentar el abanico de posibilidades combinatorias semánticas y sintácticas (valencias) de cada palabra en cada uno de sus acepciones,

⁴⁸ Para más información, véase Ruppenhofer *et ál.* (2006) y la web oficial en el siguiente enlace: <<http://framenet.icsi.berkeley.edu/>>.

mediante la anotación asistida por ordenador de ejemplos de oraciones y la tabulación automática, y mostrar los resultados de la anotación. En este proyecto se están llevando a cabo diferentes tareas, como por ejemplo la creación de una herramienta de visualización de los marcos semánticos que se crean (FrameGrapher), entre otras, pero el objetivo principal de este proyecto consiste en la elaboración de la base de datos léxica FrameNet, un recurso lexicográfico electrónico en lengua inglesa basado en marcos semánticos y en información extraída de corpus textuales. Esta base de datos, FrameNet, contiene actualmente más de 11.600 unidades léxicas, más de 6.800 de las cuales están totalmente anotadas, en más de 960 marcos semánticos, ejemplificados en más de 150.000 oraciones anotadas, y sigue creciendo constantemente. Recientemente, además, ha salido a la luz la versión española de FrameNet, Spanish FrameNet,⁴⁹ que contiene más de 1.000 unidades léxicas (verbos, sustantivos, adjetivos, adverbios, preposiciones) de un amplio rango de dominios semánticos extraídos principalmente de periódicos, reseñas de libros y ensayos.

Además de en este proyecto, la Semántica de Marcos se está aplicando en la actualidad en otros proyectos, tanto lexicográficos como terminológicos, debido a la importancia que ha adquirido el aspecto cognitivo en la Lexicografía y en la Terminografía y a las posibilidades que permite a la hora de realizar un trabajo sistemático y objetivo y, sobre todo, a la hora de tener en cuenta y de representar el aspecto pragmático de la comunicación, ya sea esta general o especializada. En el ámbito terminográfico, de hecho, se ha acuñado un término nuevo para hablar de este tipo de Terminografía denominado Terminografía basada en marcos (*frame-based terminology*) (Faber et ál., 2005; Faber et ál., 2006; Faber et ál., 2007; Faber et ál., 2009a), para hacer referencia al tipo de Terminografía que utiliza como base marcos semánticos para estructurar el conocimiento especializado y crear representaciones independientes de la lengua de trabajo. Faber et ál. (2009a: 5) indica los aspectos en los que se basa esta Terminografía basada en marcos: (1) la organización conceptual; (2) la naturaleza multidimensional de las unidades terminológicas, es decir, su dimensión comunicativa, lingüística y cognitiva (Cabré Castellví, 1993: 93), y (3) la extracción de información semántica y sintáctica mediante el uso de corpus especializados

⁴⁹ <<http://gemini.uab.es/SFN>>

multilingües. En nuestro trabajo de investigación, también nos basaremos en este tipo de Terminografía basada en marcos a la hora de realizar la parte práctica de la investigación, por lo que profundizaremos en este tema más adelante.⁵⁰

Para terminar con este apartado, consideramos necesario destacar la relación tripartita de las tres teorías lingüísticas expuestas anteriormente (Lingüística Cognitiva, Lingüística de Corpus y Lingüística Computacional) con la Terminografía. A pesar de que la Lingüística Cognitiva no se encuentra representada en todos los trabajos terminográficos de la actualidad y sí la Lingüística de Corpus y la Lingüística Computacional, consideramos imprescindible recalcar la relación de trabajo de las cuatro disciplinas, puesto que todas han aportado una parte correspondiente para que la Terminografía sea como es en la actualidad y para que los proyectos y estudios terminográficos se lleven a cabo de la forma en la que se realizan actualmente. Así pues, la relación se puede resumir de la siguiente manera: la Lingüística de Corpus permite tener en cuenta el funcionamiento de los términos dentro de su contexto comunicativo real, a fin de adquirir el conocimiento necesario acerca de su formación, de su utilización, de su frecuencia, de la predilección por un término u otro por parte de los especialistas, etc.; la Lingüística Cognitiva permite analizar y comprender la organización mental y de categorización que se realiza de los conceptos de un dominio de especialidad y su representación semántica, con objeto de entender el dominio y de estructurar adecuadamente el conocimiento nuevo adquirido así como de facilitar la elaboración de definiciones terminológicas; y, por último, la Lingüística Computacional sirve de enlace para estas dos disciplinas proporcionando las herramientas necesarias para llevar a cabo el estudio correspondiente de los términos en su uso real de la Lingüística de Corpus, como son los extractores de candidatos a términos, de concordancias, de equivalentes, etc., y del análisis de la categorización y organización del léxico mental de la Lingüística Cognitiva, como son los editores de ontologías, los extractores de relaciones semánticas, etc.

⁵⁰ Durante el tercer y cuarto capítulo presentaremos un trabajo terminográfico basado en corpus y en la representación de conocimiento, es decir, un trabajo ontoterminográfico.

1.4. LA DICOTOMÍA TERMINOGRAFÍA – LEXICOGRAFÍA

Una vez que hemos expuesto la parte teórica de la Terminología, así como las diferentes disciplinas lingüísticas que influyen en su concepción y sus diferencias con la Lexicología, pasaremos a tratar su rama aplicada: la Terminografía.

Como veíamos anteriormente (cf. 1.3.), la mayoría de los autores suelen establecer una correspondencia paralela entre, por un lado, la Lexicología, la rama teórica que se encarga del estudio y descripción del léxico de una lengua, y la Lexicografía, la vertiente aplicada de la Lexicología centrada en la elaboración de diccionarios; y por otro, la Terminología, encargada del estudio teórico y metodológico de las unidades especializadas, y la Terminografía, la rama aplicada de la Terminología, encargada de la elaboración de diccionarios especializados.

Estas disciplinas, aunque con frecuencia se reconoce que la Terminografía posee objetivos comunes a la Lexicografía, son disciplinas que se distancian si tenemos en consideración el objeto de estudio al que cada una de ellas se dedica y la aproximación que realizan a dicho objeto:

Terminology is the study and the field of activity concerned with the collection, description, processing and presentation of terms, i.e. lexical items belonging to specialised areas of usage of one or more languages. In its objectives it is akin to lexicography which combines the double aim of generally collecting data about the lexicon of a language with providing information, and sometimes even an advisory, service to language users. The justification of considering it a separate activity from lexicography lies in the different nature of the data traditionally assembled, the different background of the people involved in this work, and to some extent in the different methods used (Sager, 1990: 2).

Asimismo, el mismo autor (1990: 55) nos señala la principal diferencia entre la Lexicografía y la Terminografía, que radica en el material inicial que se recoge y se procesa:

One of the fundamental differences between lexicography and terminology lies in the attitude to the lexicon, the raw material to be collected and processed. The lexicographer in principle collects “all” the words of a language in order to sort them in various ways. His ideal dictionary covers all the words and all their meanings, even though in practice he will produce various types of subsets for diverse uses. The terminologist starts out from a much narrower position; he is only

interested in subsets of the lexicon, which constitute the vocabulary (or lexicon) of special languages. In order to arrive at these subsets he needs a structure of knowledge, which justifies the existence and the boundaries of special languages, so that he can attribute words to separate areas of this structure. Since a word can belong to more than one area of knowledge –the well-known phenomenon of homonymy- the terminologist has to distinguish meanings before he distinguishes words. (ibid.: 55).

A continuación, trataremos las cuatro distinciones que establece Rondeau (1983: 62-63) entre la Terminografía y Lexicografía. Estas diferencias parten de una aproximación wüsteriana de la Terminología, por lo que se engloban dentro de la Terminología tradicional y están basadas en los cuatro elementos principales de cualquier disciplina, a saber: el objeto de estudio; el punto de vista de los estudios, la metodología y, por último, el tratamiento lingüístico.

Para Rondeau, la palabra (unidad léxica) es el objeto de estudio de la Lexicografía, y el término (unidad terminológica) de la Terminología. Esta distinción está clara y se encuentra aceptada por todos los especialistas lingüísticos, aunque se han dedicado multitud de trabajos a la diferencia entre término-palabra y a las diferentes denominaciones (unidad terminológica-término). El problema radica principalmente en distinguir cuándo una unidad léxica es un término y cuándo no lo es, ya que en algunos contextos puede desempeñar esa función pero no en otros. Asimismo, la diferencia entre contextos especializados y generales también es ambigua a veces y esto provoca también dudas sobre la naturaleza de la unidad.

En segundo lugar, encontramos la diferencia de perspectivas. Para el autor, la Lexicografía es descriptiva, mientras que la Terminología es prescriptiva. Esta afirmación puede dar lugar a polémica, si tenemos en cuenta que ejemplos de Lexicografía prescriptiva (*Diccionario de la Real Academia de la Lengua*) y de Terminología descriptiva son fácilmente detectables. No obstante, hay que tener en cuenta que uno de los usos principales de la Terminología es la estandarización, proveniente normalmente de organizaciones internacionales (ISO)⁵¹ o nacionales (AENOR)⁵², y que persigue la homogeneización (internacional o nacional) de la

⁵¹ International Organisation for Standardisation (ISO). URL: <http://www.iso.org>

⁵² Asociación Española de Normalización y Certificación (AENOR). URL: <http://www.aenor.es/>

Terminología utilizada en las materias de especialidad. La estandarización no es necesario que se dé en organizaciones internacionales o nacionales, se puede emplear en organizaciones, empresas, instituciones, etc. con objeto de establecer los conceptos utilizados, como es el caso de la Terminología empleada en la Unión Europea o en grandes empresas como es el caso de Microsoft. Asimismo, otro fenómeno importante a la hora de la prescripción son los neologismos, e incluso los extranjerismos (anglicismos, sobre todo), ya que la creación o adaptación de términos en los campos especializados es una actividad muy frecuente y requiere de una normalización para su inclusión en la lengua en cuestión.

En tercer lugar, el autor nos especifica la distinción en la metodología. En Terminología, se utiliza un método onomasiológico (del concepto al término), mientras que para la Lexicografía se emplea un método semasiológico (de la palabra al concepto). Este punto presenta también cierta duda actualmente, ya que los primeros terminólogos, como Schlomann y Wüster, eran especialistas en su campo de especialidad y realizaban sus estructuras terminológicas desde su conocimiento y su experiencia, es decir, partían del concepto para alcanzar el término. Sin embargo, los terminólogos de hoy en día cuentan con formación y conocimientos lingüísticos, pero normalmente carecen de conocimientos específicos del campo en el que trabajan (medicina, química, automoción, etc.), por lo que trabajan a partir de candidatos a términos extraídos (ya sea de forma manual o automática) de textos especializados. De esta manera, la labor de los terminólogos modernos se asemeja más a la metodología semasiológica (es decir, a la lexicográfica) que a la inicial metodología onomasiológica defendida por la Terminología tradicional. Asimismo, en la actualidad se pueden encontrar trabajos lexicográficos con un enfoque onomasiológico y trabajos terminológicos con una perspectiva semasiológica.

No obstante, esta metodología se emplea también actualmente a la hora de diseñar la representación conceptual de un campo de especialidad. En este esquema, se representan los conceptos del campo en cuestión y se indican las diferentes relaciones que tienen entre ellos, lo que da lugar a una red semántica de un lenguaje de especialidad.

Por último, Rondeau trata la diferencia en el tratamiento lingüístico desde dos puntos: por un lado, la representación de la unidad principal; y, por otro, la forma de la definición. En el primer caso, la diferencia radica en el hecho de que, en Terminografía, la unidad terminológica se considera un todo a pesar de contar con más de un término, es decir, no se fragmenta en los diferentes componentes. Por ejemplo, en el caso de «*conceptual dependency theory*» se traduciría en español por «teoría de dependencias conceptuales» y en francés por «*théorie des dépendances conceptuelles*» (Otman, 1991). Sin embargo, en una fuente lexicográfica, esta misma unidad se presentaría de la siguiente manera: «*dépendances conceptuelles, théorie des*» o «*conceptuelles, théorie des dépendances*» (ibíd.).

Con referencia al tipo de definición, los terminólogos suelen utilizar un tipo de definición más sistemática que los lexicógrafos, ya que pretenden reflejar la cohesión de la estructura conceptual en la definición. En este sentido, se trata de una definición más estricta y controlada que la que se pueda encontrar en cualquier fuente lexicográfica, que tiende a ser más enciclopédica, en el sentido de incluir más información de la estrictamente necesaria para comprender el significado de una unidad léxica.

| | LEXICOGRAFÍA | TERMINOGRAFÍA CLÁSICA |
|---|---|--|
| Objeto de estudio | Unidad léxica | Unidad terminológica |
| Punto de vista | Descriptivo | Prescriptivo |
| Metodología | Semasiológico | Onomasiológico |
| Tratamiento: - Unidad principal - Definición | - Unidad indivisible - Sistemática, controlada y cohesiva. | - Unidad más flexible - Coherente, pero no tan sistemática ni controlada. |

Tabla 4. Diferencias entre la Lexicografía y la Terminografía según Rondeau (1983: 62-63).

Una vez determinadas las diferencias que hemos presentado anteriormente propuestas por Rondeau, donde se deja patente la distinción entre la Terminografía y la Lexicografía desde un punto de vista de la Terminología tradicional, debemos destacar que ambas disciplinas responden a una misma necesidad, que consiste en la sistematización. Cada disciplina afronta esta sistematización de forma diferente, desde

distintas perspectivas y según su objeto, pero en definitiva su objetivo principal es compartido, a saber: sistematizar una estructura, ya sea léxica o terminológica, la cual no es más que una estructura conceptual formalizada con fines simbólicos según las convenciones lingüísticas (Langacker, 1987: 98).

Además de esta finalidad común de sistematización, la Terminología moderna defiende el hecho de que estas dos disciplinas comparten un mayor número de aspectos y, por tanto, no considera adecuada la separación tan tajante que realiza la Terminología tradicional entre la Lexicografía y la Terminografía. Por este motivo, la Terminología moderna propone una aproximación a la metodología empleada en la Lexicografía, como veremos a continuación, alejándose, así, de la corriente terminológica tradicional.

1.4.1. APROXIMACIÓN DE LA TERMINOGRAFÍA A LA LEXICOGRAFÍA ESPECIALIZADA

Las diferencias entre el enfoque terminológico y el enfoque lexicográfico resumidas en el capítulo anterior se establecen en el seno de la Terminología tradicional, es decir, son Wüster y sus seguidores los que consideran que debe existir una división clara y determinante entre las dos disciplinas. La TGT establece una separación nítida entre el conocimiento especializado y la lengua general, de manera que considera que la Terminografía debe ocuparse de los términos y la Lexicografía, de las palabras, y, para ello, establece una separación de su metodología (onomasiológica-semasiológica), del punto de vista (prescriptivo-descriptivo) y, por último, del tratamiento que se debe dar a cada unidad (cf. Tabla 4).

Actualmente, las corrientes modernas de la Terminología rechazan esta distinción exhaustiva y consideran la Terminografía y la Lexicografía como disciplinas relacionadas y complementarias entre sí. Y es que, como señala Rey (1995: 118): «The relationships between terminology-terminography and lexicography are clearly complex and reciprocal». Esta reciprocidad implica, por un parte, que la Lexicografía puede, y debe, ser modelo y complemento de la Terminografía, ya que su larga existencia como práctica y representación sistemática de designaciones de una lengua contribuye con una amplia experiencia de técnicas de las que se puede beneficiar la Terminografía. Y por otra parte, esta permite el progreso de la Lexicografía aportando nuevos métodos

basados en las nuevas tecnologías informáticas para la documentación, la producción de textos impresos y el análisis de datos, así como para la creación y denominación de unidades y de recursos.

Según Bergenholtz y Tarp (1995: 198), la distinción entre Lexicología/-grafía y Terminología/-grafía de la TGT es demasiado rotunda, ya que se basa en la metodología de trabajo y en la identificación de objetos diferentes para cada disciplina. En la misma línea, Sager (1990: 56) señala lo siguiente:

Traditional terminological theory identifies its approach as "onomasiological", i.e. a "naming" approach, because, in principle, it starts from concepts and looks for the names of these concepts. By contrast the lexicographical approach is called "semasiological", i.e. a "meaning" approach, because it starts from words and looks for their meaning. In reality the onomasiological approach only characterises the scientist who has to find a name for a new concept (an invention, a new tool, measurement, etc.); the terminologist, like the lexicographer, usually starts from an existing body of terms to start with.⁵³

Es decir, Sager defiende el método semasiológico también para el inicio de cualquier trabajo terminográfico, ya que los terminólogos, a la hora de comenzar cualquier tarea necesitan partir de unidades terminológicas determinadas para poder acceder a los conceptos que designan. En otras palabras, al no ser expertos en la materia en la que están trabajando no tienen los conocimientos previos necesarios para designar los conceptos del área de especialidad en cuestión y, por tanto, su trabajo comienza a partir de unidades concretas. Solamente los expertos en las materias pueden empezar a trabajar a partir de conceptos y emplear, así, un método onomasiológico.

Los principales factores que han provocado esta rotura con la visión drástica y diferenciadora de la TGT han sido los cambios que se han producido en el mundo desde la mitad del siglo XX hasta hoy, como son por ejemplo:

- La multiplicación de los conocimientos especializados;
- La difusión imparable del conocimiento a través de los medios de comunicación;
- La descentralización del control de la información;

⁵³ Énfasis añadido.

- El surgimiento de las industrias de la lengua;
- La aparición de las nuevas tecnologías como instrumentos de trabajo.

Todo ello ha tenido gran repercusión en la relación de la Terminografía con la Lexicografía, al igual que en la evolución de la teoría de la Terminología como vimos anteriormente (cf. 1.3.2.2.) y ha dado lugar a un acercamiento de las dos disciplinas aplicadas. Sin embargo, han sido dos los factores principales que acercado el método de trabajo de la Lexicografía y la Terminografía (Cabré Castellví: 2007: 81):

- la entrada en el trabajo terminográfico de personas de formación lingüística o traductora, y
- la adopción de tecnología para el trabajo, que en algunos casos ha hecho cambiar la concepción misma de una materia de conocimiento.

Históricamente, los profesionales de la Terminología y los de la Lexicografía procedían de formaciones distintas: en Lexicografía, los diccionarios los elabora normalmente el profesional de formación lingüística (lexicógrafo); pero en Terminografía, los realizaban hasta hace unas décadas solamente los especialistas, como E. Wüster. En este contexto, es evidente que la procedencia formativa del profesional condicionaba el proceso de trabajo en cada caso y provocaba una distinción en la metodología. Sin embargo, desde los años cincuenta hasta ahora han entrado en el trabajo terminográfico profesionales de formación lingüística o traductora y, con ello, la diferenciación tan clara entre las aproximaciones (onomasiología y semasiología) que existía con anterioridad se ha desdibujado y ha dado lugar a una combinación de métodos de trabajo. Actualmente, la mayor parte de los profesionales de la Terminografía tienen la misma formación que los de la Lexicografía, lo que condiciona el proceso de trabajo, que pasa a ser semasiológico y deja de ser onomasiológico.

En la mayoría de los trabajos terminográficos, este se lleva a cabo en equipo, dividiendo así las competencias necesarias: 1. competencia en el tema de la especialidad; 2. competencia en las lenguas de trabajo, y 3. competencia en la metodología terminográfica (ibíd.: 82). En estos casos, es común que el especialista del dominio de especialidad solo actúe como asesor en la fase inicial y de elaboración o como revisor en la fase final, mientras que el lingüista o el traductor son los que llevan a

cabo realmente el trabajo terminográfico de base. Así pues, esta formación compartida por los profesionales de la Terminografía y por los de la Lexicografía hacen que se aproxime la metodología empleada en cada disciplina, es decir, el empleo del método semasiológico tanto para los trabajos terminográficos como los lexicográficos. No obstante, es importante recalcar que el trabajo de los expertos en cada ámbito de especialidad y su colaboración en el trabajo terminográfico, sobre todo en la fase inicial, sigue siendo onomasiológico.

En este contexto, a veces la separación nítida entre una disciplina y otra se presenta complicada y difusa, y la distinción entre unidades léxicas y unidades terminológicas no se puede determinar tajantemente. Por tanto, el trabajo terminográfico se encontrará a menudo entre las dos disciplinas, y lo mismo ocurrirá con el trabajo lexicográfico.

Asimismo, debido a esta complementariedad y a la división poco clara entre las disciplinas, no nos podrá sorprender encontrar en un diccionario general unidades léxicas que se podrían clasificar como términos (Gómez González-Jover, 2006: 68), o viceversa. Esta oscilación se debe principalmente a las características que presentan los lenguajes de especialidad, especialmente a la dificultad de diferenciarlos de la lengua común y a la existencia de la interjección entre estos lenguajes. En este sentido, y como ejemplo ilustrativo, podremos encontrar diferentes tipos de unidades terminológicas en cualquier lenguaje de especialidad, a saber:

- Las unidades terminológicas propias del lenguaje de especialidad, como es el caso de «viaje combinado», «animador turístico», «*Inn*», «*B&B*», «*Ökotourismus*», «*shuttle*».
- Las palabras del lenguaje cotidiano con acepciones propias del lenguaje de especialidad, como por ejemplo «categoría», «*Stern*», «*Team building*», «*Incentives*».
- Las unidades terminológicas que han visto ampliado su uso al lenguaje cotidiano, como es el caso de «parador», «*spa*», «*overbooking*».

El terminólogo deberá, por tanto, tener siempre muy en cuenta de qué unidad terminológica se trata para poder completar la información pertinente de dicha unidad

de la mejor forma posible y no dejar ningún dato importante fuera de la definición y de la información terminológica de la entrada del término.

Fruto de este acercamiento de las dos metodologías ha surgido una nueva denominación para esta disciplina aplicada, la denominada Lexicografía especializada (*LSP Lexicography*) (Bergenholtz y Tarp, 1995) o Lexicología terminológica (Lerat, 1997: 189). Estos autores consideran que no está justificada la distinción entre Terminología y Terminografía por un lado, y entre Lexicografía y Terminografía por otro, como materias con principios teóricos y prácticos totalmente diferenciados y, por tanto, con tratamientos individualizados. Por este motivo, apuestan por una denominación común para aglutinar en un único marco teórico diferencias que no son reales entre la Lexicografía y la Terminografía y establecen, así, un híbrido entre la metodología tradicional de la Terminografía y la de la Lexicografía. De esta manera, defienden la aplicación de métodos lexicográficos estándares para describir un dominio determinado (por ejemplo, el turismo de aventura) y se alejan del modelo prescriptivo clásico, el cual lo identifican con el término Terminografía.

Según Bergenholtz y Tarp (1995: 10-11), esta Lexicografía especializada tiene fundamentalmente las siguientes características: a) trabaja con términos; b) utiliza en sus productos tanto macroestructuras sistemáticas como alfabéticas; c) debe ser descriptiva y prescriptiva en mayor o menor grado, según la finalidad del trabajo terminográfico; d) el usuario de sus productos es tanto el lego como el experto; e) las obras que elabora pueden ser tanto codificadoras como descodificadoras.⁵⁴ De esta manera, se empiezan a establecer las diferencias con la Terminografía clásica y los límites de la nueva *Lexicografía especializada*.

En la siguiente tabla, se realiza una comparativa entre el enfoque tradicional de la Terminografía con los nuevos argumentos de los autores mencionados anteriormente:⁵⁵

⁵⁴ En este punto, los autores se refieren a las funciones que desempeña un recurso terminográfico a la hora de su utilización. En este sentido, un recurso tendrá una función codificadora, si su finalidad es asistir al usuario durante la escritura (o reescritura, en el caso de los traductores) de un texto. Por su parte, tendrá una función descodificadora, si su objetivo es ayudar a comprender un texto o cualquier tipo de información.

⁵⁵ Se ha procedido también a una adaptación propia de las características.

| CARACTERÍSTICAS | TERMINOGRAFÍA | LEXICOGRAFÍA ESPECIALIZADA |
|---------------------------------------|--|---|
| 1. Aproximación | Onomasiológica (se trabaja a partir de conceptos). | Principalmente semasiológica (se trabaja a partir de datos en corpus), aunque también se trabaja de forma onomasiológica cuando colaboran expertos de las especialidades estudiadas. |
| 2. Usuarios | Principalmente, profesionales de una disciplina concreta. | Cualquier usuario interesado (desde especialistas a legos de una materia). Los traductores, intérpretes, periodistas, documentalistas y otros profesionales que median en la comunicación también pueden ser usuarios. |
| 3. Situación comunicativa | Dentro del campo de especialidad correspondiente y entre especialistas. | Se tienen en cuenta los diferentes usuarios (diferentes niveles de conocimiento) y las posibles situaciones comunicativas, tanto dentro como fuera del campo de especialidad. |
| 4. Objeto de estudio | Conceptos relacionados con la materia de especialidad. | Términos, como unidades especializadas dentro de un campo de especialidad. |
| 5. Orden de la macroestructura | Sistemático, mediante sistemas conceptuales. | Alfabético, sistemático o ambos. ⁵⁶ |
| 6. Función | Prescriptiva y codificadora. | Prescriptiva y descriptiva, según la finalidad del trabajo terminográfico. Codificadora y descodificadora, también según la finalidad del trabajo terminográfico. |
| 7. Aspectos lingüísticos de productos | Ofrece solo información sobre diferencias gramaticales que estén relacionadas con la asignación del concepto a un término. Incluye principalmente nombres, aunque también algunos verbos y adjetivos. | Recoge aquella información lingüística (gramatical, sintáctica, morfológica, etc.) útil para el usuario potencial y según la función del recurso (codificadora o descodificadora). Incluye principalmente nombres, aunque estudia otras unidades (abreviaturas, sufijos, adjetivos, adverbios, etc.) |

Tabla 5. Comparativa de la Terminografía con la Lexicografía especializada, según Bergholtz y Tarp (1995) y adaptación propia.

Como hemos observado a través de esta tabla, se ha establecido una distinción clara entre la Terminografía prescriptiva, procedente de la Terminología clásica, y la Lexicografía especializada (Bergholtz y Tarp, 1995). Sin embargo, la denominación

⁵⁶ La elaboración de recursos terminológicos electrónicos en la actualidad se suelen presentar con la combinación de ambos tipos de organización, con objeto de facilitar el acceso a la información por parte del usuario.

de esta nueva rama aplicada ha recibido varias críticas y su uso no está muy extendido entre los terminógrafos.

El problema con su denominación proviene de la metalexigrafía, ya que en esta rama de la lingüística se ha venido utilizando este término para identificar distintas aplicaciones de la Lexicografía general (monolingüe y bilingüe) y, por tanto, puede resultar ambiguo. Algunas de estas aplicaciones son los diccionarios especializados por las unidades lingüísticas (diccionarios de verbos, de locuciones, etc.), por la representación de contenidos (diccionarios visuales, enciclopédicos, etc.), por la información lingüística (diccionario de sinónimos, de antónimos, etc.), por los usuarios (diccionarios de aprendizaje, infantiles, etc.) o por la temática (diccionarios técnicos, médicos, etc.). Dentro de la Lexicografía general, estos últimos tipos de diccionarios no se caracterizan por seguir los postulados ni la metodología de trabajo de la Terminografía estricta, ya que, por lo general, abarcan delimitaciones de la especialidad muy amplias (Lorente, 2001). Según esta autora, es en este punto donde se sitúan Bergholtz y Tarp (1995), «con el objeto de distanciarse de una orientación exclusivamente prescriptiva y para diversificar la oferta de diccionarios especializados por la temática».

Al igual que Lorente, otros muchos autores relevantes no han aceptado el término «Lexicografía especializada» (*LSP Lexicography*) y continúan utilizando el término «Terminografía» (haciendo referencia a la «Lexicografía especializada por la temática»), aunque con los principios teóricos y prácticos de la nueva metodología. En nuestro caso, también seguiremos esta postura y hablaremos de «Terminografía», en vez de emplear el término «Lexicografía especializada».

Para concluir con este apartado, es importante recalcar que con esta nueva metodología de trabajo, el trabajo terminográfico deja de ser estático y homogéneo para convertirse en un trabajo más dinámico y flexible que nunca, ya que el terminógrafo (o el equipo de trabajo, en su caso) deberá decidir antes de comenzar cada proyecto sobre serie de variables. Entre estas, destacamos las siguientes: la finalidad comunicativa, es decir, si el recurso terminográfico presenta una función de codificación (reescritura) o de descodificación (comprensión); la finalidad cognitiva, que hace referencia a si la finalidad del recurso consiste en proporcionar conocimiento o pretende resolver algún

tipo de problema; los usuarios potenciales, que pueden ser desde especialistas en la materia a legos, estudiantes, etc.; la infraestructura y el equipo con los que se cuenta para llevar a cabo un proyecto de Terminografía (programas, corpus, capacidad de almacenamiento, etc.); la futura edición del recurso (en papel, CD-ROM, etc.) (Brekke, 2001: 181), entre otras.

Estas variables son la base para la teoría de las funciones en Lexicografía (*function theory of lexicography*) (Bergenholtz y Tarp, 2002; Tarp, 2008a, 2008b), desarrollada por investigadores de la Universidad de Aarhus (Dinamarca). Esta teoría establece una serie de principios con respecto a la labor lexicográfica y hace especial hincapié en la observancia de las variables de los diferentes proyectos terminográficos, para poder alcanzar un resultado óptimo en el producto final. Una de las afirmaciones más relevantes de esta teoría establece que los usuarios en general no necesitan información general, sino información apropiada para sus necesidades y sus expectativas. Así pues, el tipo de información requerida es siempre concreta y depende tanto de los usuarios potenciales como de los tipos de situaciones concretas en las que utilicen los diccionarios. En otras palabras, las necesidades de información de los usuarios potenciales no pueden ser definidas sin especificar *quién* necesita *qué* y *en qué contexto* (Tarp, 2009: 46). Por este motivo, y como decíamos con anterioridad, el trabajo terminográfico se ha convertido en una labor dinámica y flexible que cambia de un proyecto a otro.

Esta adaptación y adecuación del trabajo terminográfico también están reflejadas en la Teoría Comunicativa de la Terminología (Cabré Castellví, 1993), la cual propone una metodología abierta y dinámica donde caben recursos terminográficos diferentes y metodologías diversas, según el usuario potencial y los fines que se persigan con el producto final. Según esta corriente, cada trabajo concreto adopta una estrategia que ha de adecuarse a las circunstancias específicas en las que se desarrolla, a excepción de los principios mínimos que sirven de marco.

La metodología pues, lejos de actuar como un corsé, se adapta a las circunstancias sin contravenir los principios; la adecuación metodológica está por encima de la unificación extrema. [...] En todos los casos, se respetarán los mínimos, pero cada uno adecuará la metodología a sus circunstancias (Cabré Castellví, 1999b: 137).

Así pues, la metodología es variable en función de la temática del trabajo, de sus objetivos, del contexto, de los elementos implicados y de la infraestructura con la que se cuenta, etc., lo que afectará al diseño, formato, soporte y difusión de los recursos terminográficos. De esta manera, la TCT propone un marco de trabajo flexible para abarcar la diversificación de recursos, de usuarios y de situaciones comunicativas diferentes. No obstante, la flexibilidad de este método de trabajo no contraviene los principios a los que se debe ajustar cualquier trabajo sistemático y riguroso de terminología y no se puede confundir, por tanto, con otros tipos de trabajo (como la práctica de la traducción) (Gómez González-Jover, 2006: 138).

A continuación, seguiremos profundizando sobre la práctica de la Terminografía y trataremos con mayor detalle la influencia que han supuesto la Lingüística de Corpus y la Lingüística Cognitiva en la Terminografía, así como las herramientas desarrolladas por la Lingüística Computacional para las diferentes tareas terminográficas.

1.4.2. LA TERMINOGRAFÍA BASADA EN CORPUS

Como presentamos anteriormente (cf. 1.3.2.3.1.), la Lingüística de Corpus ha sido una de las disciplinas lingüísticas que más ha influido en la Terminología y, por ende, en la Terminografía. Tanto es así que hoy en día la idea de la contextualización de los términos a través de los corpus textuales está totalmente aceptada en la comunidad de terminógrafos, y los corpus son considerados recursos indispensables para cualquier trabajo de esta naturaleza.

La Terminología moderna ha incorporado el empleo de los corpus textuales como fuente de información esencial en su metodología, y estos han permitido conseguir avances y mejoras en la calidad del producto final. Así pues, la compilación sistemática de terminología⁵⁷ está basada en corpus, con objeto de realizar sus estudios

⁵⁷ La terminología sistemática hace referencia al estudio que se realiza de un campo de especialidad determinado con el fin de construir un recurso terminológico completo, ya sea una base de datos terminológica, un diccionario técnico o un glosario. En oposición a esta, se distingue otro tipo de terminología: la terminología *ad hoc* o puntual, que está centrada en el texto especializado y concierne a la investigación sobre un término aislado o sobre un conjunto restringido de términos de una misma área temática (Sager, 1990: 188). Normalmente, este último caso hace referencia a las necesidades de traductores y redactores técnicos al enfrentarse puntualmente a términos que aparecen en textos de trabajo. Para profundizar en el tema, véase Wright y Wright (1997: 148 y ss.) y Cabré (1993: 337 y ss.).

sobre el comportamiento lingüístico real de un ámbito determinado y, así, poder descubrir nuevas palabras o palabras en desuso, observar los distintos significados de una misma forma léxica, detectar colocaciones o combinaciones terminológicas, obtener contextos reales, definiciones, etc. En otras palabras, el interés de la Terminografía por el empleo de los corpus textuales reside en las posibilidades que ofrece este para la investigación de material terminológico real.

En este sentido, en Terminografía, el uso de los corpus textuales se ve motivado por dos motivos fundamentales.

Por un lado, el trabajo terminográfico no consiste en la invención de denominaciones para unos conceptos previamente establecidos como propugnaban los estructuralistas (cf.1.3.2.1.), sino en «la identificación y recopilación de los términos que los especialistas utilizan en realidad» (Cabré Castellví, 1993: 113). Por este motivo, si un terminógrafo debe estudiar los términos que los especialistas utilizan en su trabajo diario, deberá consultar directamente con los especialistas del campo de especialidad en cuestión o realizar un estudio detallado de las producciones lingüísticas que estos especialistas crean para comunicarse entre ellos o con otros actores.

De estas dos opciones, la primera no es siempre posible, ya que puede resultar complicado disponer de los especialistas adecuados y, cuando se puede acceder a ellos, a menudo encuentran dificultades a la hora de explicar el significado y el uso del lenguaje que emplean, al fin y al cabo, de forma intuitiva. En palabras de Meyer y Mackintosh (1996: 264):

While experts obviously *know* their domains, they do not all *explain* their knowledge *clearly* (whether orally or in writing), *completely* (it is up to the knowledge acquirer to make sure that all important areas in the field are covered), or *consistently* (experts often disagree with each other or change their minds).

Por ello, la segunda opción, la consulta de documentación especializada en forma de corpus textual, es más accesible, rápida y directa y, por tanto, determinante en el trabajo terminológico.

Por otro lado, el segundo motivo que hace imprescindible el uso del corpus en Terminografía se refiere a la dimensión conceptual de los términos. Para poder

identificar y recopilar los términos que los especialistas emplean en la realidad, los terminógrafos necesitan estudiar las estructuras de conocimiento (los conceptos y sus relaciones) que los términos representan. Es decir, los terminógrafos deben familiarizarse con el tema específico de su trabajo y adquirir los conocimientos básicos que le permitan estructurar y delimitar adecuadamente el alcance de la obra que pretenden elaborar, lo que Cabré Castellví (1999a: 144) denomina «competencia cognitiva». De la misma forma que en el caso anterior, los terminógrafos pueden dirigirse a especialistas en el ámbito de especialidad en cuestión, con lo que se encontrarán con problemas similares a los citados anteriormente, o consultar documentación especializada. En la mayoría de las ocasiones será más fácil para el terminógrafo familiarizarse con el ámbito de especialidad, con sus conceptos y estructuras a través de la documentación especializada, así como consultar a los especialistas para realizarle preguntas acerca del uso, significado, definición, etc. específico de una unidad terminológica que pueda ver en un texto y, por tanto, en su contexto real.

La necesidad de la utilización de documentación especializada en el trabajo terminográfico sistemático queda patente con estas dos razones expuestas. Asimismo, el empleo de corpus textuales se vuelve imprescindible en esta labor si nos referimos al estudio de la variación denominativa y conceptual (Cabré Castellví, 1999b) y al estudio de la formación de nuevos términos.

La labor terminográfica basada en la documentación especializada o, incluso en corpus textuales, no es nueva, ya que autores como Rondeau (1983: 71), Dubuc (1980: 24) y Picht y Draskau (1985: 167) comentan la utilización desde siempre de fuentes de información lingüística y conceptual. De hecho, todos los manuales coinciden en resaltar la importancia del uso de la documentación especializada en el trabajo terminológico. Sin embargo, la gran mayoría de estas fuentes se refieren al análisis *manual* de los textos utilizados como fuentes de información, por lo que el trabajo terminográfico se vuelve tedioso y extenuante, además de complicado a la hora de realizar, por ejemplo, estudios sobre la variación denominativa o de localizar y ordenar la información lingüística, conceptual y pragmática de los textos. Por este motivo, la

mayor parte del trabajo terminográfico actual se realiza de forma semiautomática,⁵⁸ ya que se encuentra asistido por herramientas informáticas que permiten agilizar y facilitar la labor del terminógrafo.

Este cambio de metodología (de manual a semiautomática) se ha debido principalmente a la revolución tecnológica de las últimas décadas que ha dado lugar a grandes avances, sobre todo en el campo de la informática. Gracias a estos avances, se han incrementado sobremanera la velocidad y la capacidad de almacenamiento de los ordenadores, se han creado herramientas informáticas destinadas a procesar, analizar y recuperar un gran número de textos y, además, se ha facilitado el acceso y la recuperación de textos electrónicos de diferentes contenidos, formatos y ubicaciones a través de redes de información existentes, especialmente la red Internet. Todo lo cual ha permitido (y permite) la elaboración, el tratamiento y el procesamiento de corpus textuales de gran tamaño y riqueza almacenados en ordenadores, lo que se denominan corpus textuales informatizados o electrónicos. De esta manera, se ha hecho posible que el terminógrafo pueda tener a su disposición grandes cantidades de texto en formato electrónico, de forma que sus observaciones no estén basadas solo en sus intuiciones lingüísticas, sino en el estudio detallado del uso lingüístico, lo que Leech (1992: 106) reconoce como «a new way of thinking about language».

No obstante, y a pesar de todas las ventajas que presenta el uso de corpus textuales electrónicos en el estudio del lenguaje en uso, esta herramienta se ha incorporado muy recientemente en la labor terminográfica (Meyer y Mackintosh, 1996: 257) y, además, la difusión académica de este tipo de estudios terminológicos es todavía limitada, si lo comparamos con el campo lexicográfico. Asimismo, la Terminografía basada en corpus sigue utilizando las técnicas y herramientas de la Lexicografía basada en corpus, ya que el corpus está muy asentado en esta disciplina, lo que no siempre es recomendable como nos indican Meyer y Mackintosh (1996: 258):

⁵⁸ A pesar de que las herramientas informáticas permiten un manejo automático de la Terminología, como por ejemplo para la extracción de términos, de relaciones conceptuales, de cognados, etc., es necesaria la revisión y corrección posterior de los resultados extraídos automáticamente por parte de los terminólogos.

Hence, it is essential that terminographers begin to specify the types of corpora they need and the corpus-analysis tools that best suit their task. [...] in broad terms, the *specificity* of corpus terminography. We shall argue that terminography and lexicography are in many ways substantially different, and consequently, that the tools and techniques developed in corpus lexicography cannot be applied intact to terminography. On the contrary, some can actually be *dangerous*.

En nuestra opinión, y coincidiendo con Meyer y Mackintosh (ibíd.) y con las recomendaciones del informe POINTER (1996: 90),⁵⁹ la utilización en Lexicografía de los corpus textuales para estudiar el uso lingüístico ha de hacerse extensiva a la labor terminográfica y debe desarrollarse con sus propias estrategias en el campo de la Terminografía:

Indeed, it can be argued that the use of corpora, particularly electronic corpora, is even better motivated in terminology than in general-purpose lexicography as a source of "evidence" for the compilation of terminologies. Our reasons relate mainly to the notion that special-language texts deal with semantically-restricted domains and that these texts are produced largely by members of closely-defined discourse communities in order to disseminate their knowledge of these domains. Such restrictions can be said to reduce the degree of lexical and syntactic variation in text [...] and therefore lead to what might be described as a more objective kind of evidence for the terminologist than the much more variable texts of general language. It is not the case that there is no variation in special-language texts, simply that there is less of it than in general-language texts.

Actualmente, existen proyectos y centros de investigación⁶⁰ que están impulsando el uso de estos corpus para el estudio terminológico, pero es necesario que su uso se extienda a sectores más amplios, con objeto de determinar las necesidades específicas de los terminógrafos con respecto a este tipo de recursos, como por ejemplo los tipos de

⁵⁹ El informe POINTER hace referencia al informe final que se elaboró en el proyecto POINTER (1995-96). Este proyecto ha sido realizado por un consorcio en el que han participado centros de investigación europeos dedicados al desarrollo y difusión de recursos terminológicos (Infoterm, Termcat, CL Servicios Lingüísticos, Institut National des Telecommunications, Deutsches Institut für Terminologie, etc.), así como organizaciones comerciales de Austria, Bélgica, Francia, Alemania, Grecia, Italia, España, Suiza, los Países Bajos, Escandinavia y el Reino Unido. Su principal objetivo consistía en estudiar detalladamente los recursos terminológicos existentes en Europa y recomendar en consecuencia una serie de propuestas que aseguraran el acceso, distribución y reutilización de dichos recursos, así como la creación de recursos nuevos que se ajustaran a las demandas del creciente mercado europeo. Se puede acceder a sus informes a través de la siguiente URL: <http://www.computing.surrey.ac.uk/ai/pointer>.

⁶⁰ A final de este capítulo, se hablará de algunos de estos proyectos y centros de investigación y de sus trabajos actuales.

corpus que requieren, las características que deben presentar (tamaño, tipo de textos, nivel de especialización, etc.) y las herramientas más apropiadas para que puedan ser de utilidad en la extracción terminológica y conceptual, en la búsqueda de cognados, en el estudio de variación denominativa, etc. En definitiva, es esencial que, para desarrollar una Terminología basada en corpus que mejore la calidad de la investigación terminológica, se concierten un conjunto de técnicas y herramientas propias que permitan a esta disciplina realizar sus estudios y trabajos sobre la terminología de los lenguajes de especialidad y, así, abandonar las herramientas desarrolladas propiamente para la Lexicografía.

1.4.2.1. La Lexicografía basada en corpus

La Lexicografía basada en corpus se considera el antecedente de la Terminografía basada en corpus, ya que esta no comenzó a emplear tímidamente el corpus hasta finales de la década de los noventa cuando ya en la Lexicografía se había asentado como herramienta de trabajo y su uso estaba muy extendido. Por este motivo, la metodología utilizada en la Terminografía basada en corpus es muy similar a la empleada en la Lexicografía, que, como hemos mencionado, contaba ya con un aparato crítico propio y una metodología específica cuando se incorporó el corpus a la Terminografía.

Al igual que en Terminografía, los lexicógrafos siempre habían utilizado diferentes tipos de fuentes de información para obtener los datos necesarios para la descripción lingüística. Tradicionalmente, estas fuentes han sido (a) la intuición, (b) otros diccionarios y (c) fuentes tradicionales de recopilación manual de información sobre el uso de las palabras (citas de autores reconocidos, periódicos, libros, etc.) (Sinclair, 1993: 5). Sin embargo, estas tres fuentes de información carecían a menudo de la fiabilidad y determinación requerida para llevar a cabo la labor lexicográfica y, por tanto, podían dar lugar a errores o sesgos, a descripciones y unidades obsoletas, a falta de atención en algunas estructuras sintácticas o colocacionales. En definitiva, estas fuentes, aunque parcialmente válidas, no aseguraban un trabajo totalmente objetivo y fiable.

Con la introducción del uso de los corpus textuales informatizados (corpus electrónicos o *machine-readable*), las posibilidades de análisis lingüístico que se pueden llevar a cabo en el trabajo lexicográfico se han multiplicado exponencialmente, y la fiabilidad de los resultados también ha aumentado de manera considerable. Con los corpus electrónicos, se pueden llevar a cabo estudios detallados de la lengua en un contexto real, que pueden ayudar a descubrir regularidades (e irregularidades) en el uso de la lengua, a observar estructuras sintácticas, etc., con una perspectiva más amplia y con índices de frecuencia relativa más fiables (Pérez Hernández, 2002). En palabras de uno de los lexicógrafos más reconocidos de la actualidad, Patrick Hanks, los corpus electrónicos se han convertido en herramientas indispensables para el trabajo lexicográfico: «From a lexicographical point of view, a large corpus is an indispensable tool of the trade for serious compilation of paper dictionaries and computational lexicons alike» (Hanks, 2003: 54).

La primera vez que se empleó el corpus en la elaboración de diccionarios fue en el novedoso proyecto Collins COBUILD (Collins Birmingham University International Language Database) en la década de los ochenta, llevado a cabo de forma conjunta por la editorial Harper Collins y la Universidad de Birmingham bajo la dirección de J. M. Sinclair. Fruto de este proyecto fue el diccionario *Collins Cobuild Dictionary of English Language*, que supuso sin duda alguna una revolución en la historia de la Lexicografía y marcó un antes y un después en la elaboración de diccionarios: «In 1987, with the publication of the COBUILD dictionary [...], a new development in lexicography emerged: the corpus-based lexicography» (Hanks, 2003: 53). Desde entonces, se ha considerado el corpus como la herramienta base para cualquier trabajo lexicográfico y, hoy por hoy, la creación de diccionarios no se contempla sin la utilización de un corpus textual electrónico. Tanto es así que en la década de los noventa, las editoriales de diccionarios británicas, en especial aquellas encargadas de la publicación de diccionarios pedagógicos, invirtieron grandes cantidades de dinero para revisar sus diccionarios y modificar sus macroestructuras, reestructurar las definiciones y actualizar el uso de palabras, etc. de acuerdo con las nuevas investigaciones con corpus (ibíd.: 59), lo que dio lugar a un cambio radical en la mayoría de los diccionarios que se habían elaborado hasta ese momento.

Lo más destacable de ese proyecto fue que, por primera vez, se llevó a cabo la compilación de un diccionario mediante un análisis detallado de un corpus representativo de textos ingleses, orales y escritos, con un total de 200 millones de palabras.⁶¹ La selección de los textos se realizó de forma cuidadosa y equilibrada para proporcionar gran precisión y objetividad al estudio sistemático del inglés de uso diario.

Los lexicógrafos del equipo COBUILD fueron los pioneros en el uso de los corpus textuales electrónicos, aunque hoy en día casi todas las editoriales importantes también utilizan este recurso en el proceso de elaboración de sus diccionarios en mayor o menor medida. Tanto Oxford University Press como Addison-Wesley Longman y Larousse Kingsfisher Chambers han colaborado activamente en la creación del British National Corpus (BNC) y lo han utilizado en la elaboración de sus diccionarios, como por ejemplo en la quinta edición del *Oxford Advanced Learner's Dictionary of Current English* (Crowther, 1995), la tercera edición del *Longman Dictionary of Contemporary English* (Summers, 1995) y el *Times-Chambers Essential English Dictionary* (Higgleton y Seaton, 1995), respectivamente. Por su parte, Cambridge University Press ha basado también sus diccionarios en corpus, como es el caso del *Cambridge International Dictionary of English* (Procter, 1995), que está basado en un corpus de 100 millones de palabras (Baugh et ál., 1996). En España, varias editoriales también cuentan con corpus de diferentes tamaños y características: Vox Bibliograf posee un corpus de 10 millones de palabras; la editorial SM, uno de 60.000 palabras, y la editorial SGEL posee el corpus CUMBRE, de 8 millones de palabras (Sánchez et ál., 1995), entre otras.

La producción de diccionarios para la enseñanza de segunda lenguas, especialmente en la lengua inglesa, ha sido sin ninguna duda una de las aplicaciones más destacadas del corpus electrónico y, tal y como nos apunta Hunston (2002: 96), la Lexicografía actual no se entendería sin la introducción de este recurso:

⁶¹ El *Bank of English* ha ido creciendo desde su compilación y actualmente cuenta con aproximadamente 650 millones de palabras de inglés oral y escrito. URL: <http://mycobuild.com/about-collins-corpus.aspx>.

Corpora have so revolutionised the writing of dictionaries and grammar books for language learners (or rather, for learners of English) that is by now virtually unheard-of for a large publishing company to produce a learner's dictionary or grammar reference book that does not claim to be based on a corpus. As a result, this is probably the application of corpora that is most far-reaching and influential, in that even people who have never heard of a corpus are using the product of corpus investigation.

Gracias a las posibilidades que ofrecen los corpus electrónicos de «estudiar los patrones de uso auténtico de la lengua en una selección representativa de textos» (Lavid, 2005: 139), el estudio de la Lexicografía con corpus ha dado lugar a multitud de estudios y trabajos de variada naturaleza, como por ejemplo estudios sobre variación lingüística (Biber et ál., 1998), sobre colocaciones y fraseología (Moon, 1998; Sinclair, 1991), sobre la compilación de diccionarios generales (Baugh, Harley y Jellis, 1996), sobre la influencia ideológica y cultural en el lenguaje (Teubert, 2000; Flowerdew, 1997), sobre la práctica de la traducción (Baker, 1993, 1995; Corpas Pastor, 2001, 2004), sobre la frecuencia relativa de uso de palabras o de significados de una palabra en diferentes registros (Conrad y Biber, 2000; Mindt, 2000; Halliday, 1992) y, en general, en prácticamente todas las vertientes de la lingüística aplicada (cf. Hunston, 2002).

Asimismo, los lexicógrafos han visto aumentadas las posibilidades de recopilación de información sobre gran cantidad de aspectos lingüísticos, como por ejemplo las diferentes acepciones de una palabra, las colocaciones y fraseología, las relaciones semánticas que se establecen entre unidades, los usos innovadores u obsoletos de palabras, los aspectos pragmáticos, las palabras de nueva creación. Todo ello, ha influido en la macroestructura de los diccionarios, que también ha cambiado en las últimas décadas para adaptarse a las nuevas líneas de investigación y, así, incidir cada vez más en la información relacionada con la lengua y su uso (Pérez Hernández, 2002). En definitiva, la Lexicografía basada en corpus ha supuesto la revisión de numerosos aspectos de la Lexicografía tradicional, que afectaban a la selección, organización y definición de las unidades que se incluían en los diccionarios.

Como consecuencia de esta incursión y propagación del corpus en la Lexicografía, la Terminografía también ha recibido la influencia de esta herramienta de trabajo, lo que ha dado lugar a la Terminografía basada en corpus. Consecuentemente, se ha producido una ruptura con la tradicional metodología terminográfica, lo que ha

dado lugar al abandono del trabajo artesanal y a una serie de avances y mejoras en la compilación de recursos terminográficos, en las técnicas de trabajo y en los resultados obtenidos, como veremos en los siguientes apartados. En este sentido, se ha puesto en valor el contexto en el que aparece un término y se ha dado un giro a una visión más empírica de los conceptos.

1.4.2.2. Algunas consideraciones en torno al corpus: concepto y definición

Today, the corpus is considered the default resource for almost anyone working in linguistics. No introspection can claim credence without verification through real language data. Corpus research has become a key element of almost all language study. This is an indication that the paradigm of linguistics is finally becoming again more pluralistic (Teubert, 2005: 1).

En principio, un *corpus* se puede considerar como cualquier colección que contenga más de un texto (corpus como cuerpo textual). Así pues, en un sentido amplio del término, un *corpus* se entiende como el inventario o el conjunto de datos que se utilizan en un trabajo de investigación, que están compuestos por formas lingüísticas, como oraciones o palabras, y también por categorías como sílabas, vocales, etc. (Alcaráz Varó y Martínez Linares, 1997: 151). Sin embargo, cuando empleamos este término en la lingüística actual, el corpus adquiere una serie de implicaciones que van más allá del conjunto de datos que sirven para el análisis y la investigación lingüística, como veremos a continuación. En este contexto, para que un corpus sea considerado como tal es necesario que atienda a una serie de criterios de diseño y que presente unas características básicas, como son el formato legible por ordenador y su diseño de acuerdo con unos criterios específicos en función de la aplicación que vaya a tener (Leech, 1991: 10; Atkins et ál., 1992: 5; Ahmad y Rogers, 2001: 726).

Hasta la fecha, el concepto de *corpus* ha recibido gran variedad de definiciones, aunque la más aceptada sigue siendo la propuesta por EAGLES (Expert Advisory Group on Language Engineering Standards):

A corpus is a collection of pieces of language that are selected and ordered according to explicit linguistic criteria in order to be used as a sample of the language (EAGLES, 1996: 4).

En esta definición, podemos encontrar tres aspectos fundamentales que deben tenerse en cuenta en la definición de los corpus para la investigación lingüística aplicada. En primer lugar, habla de «a collection of pieces of language» y «a sample of the language», es decir, el corpus no pretende incluir toda la lengua o lenguaje de especialidad que se esté analizando, sino que se limita a una muestra de ella que permita realizar estudios y extrapolarlos a toda la lengua o lenguaje de especialidad. En este sentido, hace referencia a uno de los criterios más relevantes de la Lingüística de Corpus, la representatividad,⁶² es decir, el corpus debe ser representativo y, con la representación de la realidad que contiene, permitir realizar estudios generalizados de toda la lengua o lenguaje de especialidad. En segundo lugar, encontramos «selected and ordered according to explicit and linguistic criteria», es decir, la inclusión de textos en un corpus no puede ser nunca arbitraria, sino que la compilación necesita cumplir una serie de requisitos y criterios establecidos previamente para que se pueda crear un corpus ordenado y adecuado. Y, por último, esta definición se refiere en todo momento a usos reales de la lengua, es decir, a la metodología empirista de la Lingüística de Corpus. Con ello, nos indica el propósito real del corpus: mostrar la lengua tal y como se utiliza en un contexto real.

Todos los estudiosos dedicados al corpus están de acuerdo en que estos son aspectos fundamentales en la creación y definición de los corpus, aunque no por ello dejan de ser aspectos controvertidos y que en ocasiones han dado lugar a posturas diferentes (Pérez Hernández, 2002).

Un aspecto importante que no menciona esta definición es el formato del corpus. Actualmente, no se concibe el corpus sin un soporte informático, por lo que Sinclair (2004) actualiza su definición para incluir el formato y, así, incluir la referencia a los corpus electrónicos (*machine-readable corpus*):

⁶² La representatividad puede definirse como «the extent to which a sample includes the full range of variability in a population» (Biber, 1993: 243). Para profundizar en este tema, recomendamos la consulta de Seghiri Domínguez (2006: 371 y ss.), Biber (1993, 1994) y apartado 1.4.3.2.1. de esta tesis doctoral.

A corpus is a collection of pieces of language text in electronic form, selected according to external criteria to represent, as far as possible, a language or language variety as a source of data for linguistic research.

En esta segunda definición, el autor nos muestra de nuevo las principales características de los corpus, a saber: a) la representatividad, b) el protocolo de diseño según unos criterios marcados, c) el uso real de la lengua (o variedades de esta), y añade: d) el formato electrónico.

Por su parte, Bowker y Pearson (2002: 9) consideran que un corpus debe contar con cuatro importantes características, las cuales lo diferencian de cualquier otro tipo de compilación: «[...] authentic, electronic, large and specific criteria. These characteristics are what make corpora different from other types of text collection».

Con respecto a la primera característica, estas autoras coinciden con Sinclair (2004) y con la mayoría de los lingüistas en que un corpus debe ser *original (authentic)*, es decir, debe estar formado por un conjunto de textos que sean «naturally occurring and has not been created for the express purpose of being included in order to demonstrate a particular point of grammar, etc.» (Sinclair, 1991: 171). En otras palabras, un corpus debe mostrar las unidades dentro de sus contextos lingüísticos reales (*in vivo*) para poder estudiar y analizar la información de forma objetiva y evitar, así, cualquier tipo de introspección por parte del lingüista, en nuestro caso del terminógrafo.

En segundo lugar, el corpus debe ser *electrónico (electronic)*, es decir, los textos que componen el corpus deben encontrarse en formato digital, a fin de que puedan ser procesados por un ordenador y la información pueda ser consultada también a través de este de múltiples maneras. De nuevo, las autoras coinciden con Sinclair, y con otros muchos autores de renombre (Leech, 1992; Bowker, 1996; McEnery y Wilson, 1996; Hunston, 2002; Corpas Pastor, 2001, entre otros) en esta característica, al igual que en las demás.

En tercer lugar, la característica de *extenso (large)* hace referencia al tamaño que debe tener un corpus. Esta característica se enfatiza en la definición de corpus aportada por Biber et ál. (1998: 246): «a corpus is not simply a collection of texts. Rather, a corpus seeks to represent a language or some part of a language». A este respecto, existen opiniones encontradas sobre la extensión que debe presentar un corpus textual:

por un lado, encontramos, por ejemplo, las de EAGLES (1996), que afirma que «the corpus should be as large as could possibly be envisaged with the technology of the time»; mientras que, por otro, Bowker y Pearson (2002: 10), por ejemplo, consideran que «*large* means a greater number of texts than you would be able to easily collect and read in printed form». Consecuentemente, no hay ninguna regla determinada sobre las dimensiones a las que debe limitarse un corpus aunque, como hemos comentado con anterioridad, siempre debería ser representativo de la lengua o del lenguaje de especialidad que se esté estudiando, ya que, como afirma Sinclair (1991: 171), «a corpus is a collection of naturally-occurring language text, chosen to characterize a state or variety of a language».

Dentro de esta característica relativa a la cantidad, encontramos también el aspecto relacionado con su tamaño limitado, sobre la que encontramos también diferentes opiniones. Los autores McEnery y Wilson (1996: 24), por ejemplo, consideran que un corpus es «a finite-sized body of machine-readable texts sampled in order to be maximally representative of the language variety under consideration», es decir, un conjunto de textos cerrado y limitado que deja de alimentarse en un momento determinado cuando se ha conseguido el tamaño establecido al comienzo de la compilación. Estos corpus “cerrados” o finitos representan la realidad de la lengua en un momento determinado y, como tal, no pretenden mostrar la variación que sufre a lo largo del tiempo ni la actualización de los campos de especialidad. Sin embargo, esta afirmación de McEnery y Wilson (ibíd.) no es universal, puesto que también existen corpus “abiertos” (no estáticos), como el corpus COBUILD, en inglés, o el CUMBRE, en español), es decir, corpus que se alimentan constantemente de nuevos textos y crecen indefinidamente y que, además, permiten ser actualizados mediante la inclusión o eliminación de algunos textos según los cambios de la lengua.⁶³

Por último, los criterios específicos (*specific criteria*) hacen referencia al hecho de que, como veíamos anteriormente, un corpus no puede estar formado por textos compilados de forma arbitraria, sino que es necesario que su compilación siga una serie de criterios establecidos previamente para otorgarle homogeneidad y representatividad.

⁶³ Volveremos sobre la característica de la representatividad del corpus en el apartado 1.4.3.2.1., ya que se trata de uno de los criterios principales que hay que establecer al comenzar la compilación de un corpus.

Como nos sugieren Ahmad y Rogers (2001: 726), si el corpus va a compilarse de forma sistemática, los textos que formen parte de este «need to be collected together, evaluated, selected and organized in a systematic way according to explicit criteria». En este sentido, según la finalidad que persigamos, se seleccionarán unos textos y no otros, de acuerdo con unos criterios explícitos que determinarán la pertinencia o no de los textos en el corpus. Estos criterios de compilación tienen gran relevancia y se les debe prestar especial atención, ya que, de lo contrario, no se conseguirían los objetivos planteados de forma adecuada. Por este motivo, los trataremos con mayor detenimiento en el siguiente apartado.

1.4.2.2.1. Criterios generales para la compilación de corpus

It is a truism that a corpus is neither good or bad in itself, but suited or not suited to a particular purpose (Hunston, 2002: 26).

Como indicábamos con anterioridad, uno de los aspectos más importantes de la compilación de un corpus se basa en el establecimiento de unos criterios previos que nos permitan alcanzar los objetivos propuestos mediante el corpus compilado. Algunos autores (Atkins et ál., 1992; Alvar Ezquerro y Corpas Pastor, 1994; EAGLES, 1996; Biber et ál., 1998) han realizado diferentes propuestas en torno al diseño de los corpus textuales, con opiniones a veces dispares con relación a la representatividad y tipologías textuales. Por su parte, en el informe de trabajo del grupo EAGLES (1996: 4) se establecen unos criterios mínimos que deben cumplir los corpus textuales para que puedan ser considerados como tales, a saber: cantidad, calidad, simplicidad y documentación.⁶⁴

De estos cuatro criterios, los más polémicos han sido los referentes a la cantidad y a la calidad, los cuales mantienen una relación muy estrecha entre sí. Por un lado,

⁶⁴ La cantidad hace referencia a la extensión del corpus en número de palabras. La calidad implica que el material debe ser auténtico, en el sentido de que debe proceder de situaciones comunicativas originales, a la par que representativo y equilibrado. La simplicidad hace referencia a la codificación, es decir, a la cantidad y al tipo de información que se le añade al texto (información relacionada con el origen, el autor, la fecha o el medio de publicación). Y, por último, la documentación hace necesaria el registro de las referencias exactas de las fuentes documentales utilizadas.

encontramos aquellos que defienden la postura de que el corpus sea lo más representativo y equilibrado posible (calidad), frente a otros que destacan la importancia de que el corpus sea lo más grande posible (cantidad). Sinclair (1991: 18), por ejemplo, considera que el corpus debe ser «as large as possible, and should keep on growing». Tanto es así que el mismo autor (EAGLES, 1996: 4) establece que un corpus tiene por defecto asignado el valor de “grande” y, en caso de no ser así, se clasificaría como un *corpus especial*.

La concepción que presenta este autor en este sentido parece más aceptada si el objetivo consiste en llevar a cabo un estudio lingüístico sobre la totalidad de la lengua. Sin embargo, si la finalidad del corpus consiste en realizar un estudio de una parte de la lengua, como por ejemplo el estudio de un lenguaje de especialidad, el tamaño pierde importancia con relación a la calidad. Existe un consenso generalizado respecto al hecho de que los corpus para el estudio de la lengua general han de ser mucho más amplios (de millones de palabras) que los corpus destinados al estudio lingüístico de variedades de la lengua o lenguajes especializados (cf. Rogers y Ahmad, 1994; Bergenholtz y Tarp, 1995). Por su parte, Zanettin (1998: 2) también comparte esta opinión y afirma que «en el caso de textos con un elevado grado de tecnicidad, escritos por especialistas, es muy probable que la información más relevante se pueda obtener incluso con un pequeño número de textos». Este hecho se debe principalmente a lo que se conoce como «densidad terminológica» (Ahmad y Rogers, 2001: 726), es decir, en la comunicación especializada, ya sea oral o escrita, el número de unidades que transmiten el conocimiento (los términos) es muy elevado. En este sentido, si un corpus de textos especializados no cuenta con un número tan elevado de palabras como los que presentan los corpus de lengua general pero se consigue una representatividad adecuada del campo de estudio y un alto grado de tecnicidad (elevada densidad terminológica), no habrá ningún obstáculo para la consecución de los objetivos planteados en el estudio lingüístico.

Como hemos visto, el criterio de cantidad, y por ende el de representatividad, sigue siendo hoy día un concepto polémico, ya que no parece existir un consenso acerca del número de palabras necesario para que exista tal representatividad, como nos indica en la siguiente cita López Arroyo (2000: 326):

La cuestión del tamaño es controvertida dentro de los estudios basados en corpus lingüísticos, ya que si bien se postula que un corpus debe contener un gran número de palabras para justificar la representatividad de los resultados que se extraigan del mismo, en ninguno de estos estudios se ha concretado qué se entiende por representatividad, número elevado de palabras, etc.

No obstante, parece que la mayoría de los autores sí que comparten la idea de que un corpus debe ser representativo, es decir, que la información que se extraiga de él debe poder generalizarse a un corpus hipotéticamente mayor (Leech, 1991:27). En la actualidad, la concepción de que «cuanto más grande sea un corpus, más representativo será» parece estar siendo sustituida por el interés hacia la densidad léxica (o terminológica) que presentan los textos, así como hacia el cumplimiento de ciertos criterios externos e internos establecidos previamente a la compilación de corpus textuales, como son la autoría, autenticidad, actualidad, restricción geográfica y temporal, etc. Asimismo, se están realizando estudios sobre la posibilidad de medir cuantitativamente la representatividad de los corpus *a posteriori* mediante aplicaciones informáticas y poder tener, por tanto, un fundamento objetivo de la representatividad del corpus compilado.⁶⁵

Además de la representatividad que debe proporcionar un corpus, el *equilibrio* también adquiere gran relevancia en la definición y caracterización de esta herramienta. De acuerdo con Atkins et ál. (1992: 6), un *corpus equilibrado (balanced corpus)* se define de la siguiente manera:

By “balanced corpus” is meant a corpus so finely that it offers a manageably small scale of model of the linguistic material which the corpus builders wish to study.

En otras palabras, un corpus se considerará equilibrado cuando está formado por textos o números de palabras distribuidos de forma equitativa (con relación al género textual, a la lengua, a los autores, a los subcampos de la especialidad, etc.), con objeto de mostrar la realidad de manera homogénea. Bowker (1996: 42), además, introduce el término «equilibrio conceptual», que hace referencia a la necesidad de recoger en el corpus todos los subcampos en los que se divide una parcela del conocimiento o campo de

⁶⁵ A este respecto, véanse las publicaciones de Corpas Pastor y Seghiri Domínguez (2006, 2007a, 2007b, 2009, 2011/en prensa) y Seguiri Domínguez (2006) acerca de la aplicación informática *ReCor*.

actividad, así como todos los ámbitos que guarden relación con esta cuando tiene un carácter multidisciplinar, es decir, la estructuración conceptual del ámbito objeto de estudio se convierte en una herramienta de trabajo esencial para el terminógrafo. Para que el corpus sea representativo desde el punto de vista del equilibrio, es necesario, a su vez, que tenga en cuenta la finalidad del estudio para el que va a ser compilado, lo que supondrá considerar unos criterios u otros para la selección de los textos que conformen el corpus.

The balance of a special-language corpus is perhaps more easily judged than that of a general-language corpus, especially where the corpus is being compiled with a particular purpose in mind, such as a bilingual terminology of earth-moving machinery for the translation of marketing materials, or a monolingual terminology of powered industrial trucks for the writers of safety manuals (Ahmad y Rogers, 2001: 734).

Como nos indican Atkins et ál. (1992: 5), el equilibrio de un corpus destinado al análisis lingüístico se consigue gracias al cumplimiento de unos criterios internos (o parámetros lingüísticos) y externos (o extralingüísticos) durante el diseño del corpus. Esta misma idea la defienden también otros autores como Biber (1993: 243), EAGLES (1996), Bowker (1996), entre otros. Para estos autores, el corpus debe seguir ambos criterios; de lo contrario, estaría dejando de lado un aspecto importante de su diseño:

A corpus selected entirely on internal criteria would yield no information about the relation between language and its context of situation. A corpus selected entirely on external criteria would be liable to miss significant variation among texts since its categories are not motivated by textual (but by contextual factors) (Atkins et ál. (1992: 5).

Los criterios externos son esencialmente criterios no lingüísticos que sirven para determinar todos los posibles aspectos que influyen en los textos que se van a incorporar al corpus, como por ejemplo los participantes, el contexto situacional, los límites geográficos y temporales, el género textual, los niveles de lengua, el formato o la función del texto. Por lo general, estos criterios se establecen durante la fase de diseño del corpus y, por tanto, previamente a la compilación de este.

Posteriormente, se evalúan los criterios internos o lingüísticos, los cuales permiten identificar los textos en función de una serie de categorías o atributos lingüísticos. A diferencia de los criterios externos, que varían en función del tipo de

corpus que se desee compilar y son más subjetivos, los criterios internos son más objetivos, puesto que se basan en pruebas empíricas realizadas a los propios textos que conforman el corpus. Un ejemplo de estos criterios internos sería la clasificación de un texto en registro formal o informal de acuerdo con sus rasgos lingüísticos como son el léxico o la sintaxis.

Como nos indica Gómez González-Jover (2006: 263), no todos los autores reconocen los mismos atributos dentro de estos dos tipos de criterios, ni los definen de la misma manera ni les otorgan la misma relevancia. No obstante, sí que parece existir un acuerdo generalizado respecto a la clasificación y selección de textos mediante un enfoque que combine criterios externos e internos. Como veremos en el capítulo 3, la compilación de nuestro corpus para el trabajo práctico de esta investigación también se basará en el establecimiento de unos criterios internos y externos, con objeto de conseguir el equilibrio y la representatividad en nuestro campo de especialidad para alcanzar nuestro objetivo propuesto: la elaboración de una base de datos ontoterminográfica para el turismo de aventura.

1.4.2.2.2. Tipos de corpus

La variedad de corpus existentes se debe a los diferentes parámetros que se establecen para compilarlos, lo que ha dado lugar a un numeroso abanico de corpus. A continuación, trataremos la propuesta de clasificación más completa realizada hasta el momento desde nuestro punto de vista, que se encuentra en varias obras de Corpas Pastor (2001: 157-159; 2004b: 225-228; 2008: 117-118). Esta autora presenta su tipología de corpus utilizando como base los informes EAGLES (1994 y 1996) y la tipología propuesta por Torruella y Llisterri (1999), así como aportaciones de Baker (1995) y Johansson (1998) con relación a los corpus multilingües. Asimismo, la completaremos con referencias de otros autores (Hunston, 2002; Atkins et ál., 1992; Lemnitzer y Zinsmeister, 2006).

En primer lugar, podemos clasificar los corpus según la distribución y el porcentaje de los diferentes tipos de textos que lo componen. Así pues, este parámetro permite distinguir entre:

- *Corpus grande* (o *extenso*): aquél que se caracteriza por no estar voluntariamente delimitado y contar con un número de palabras muy elevado sin tener en cuenta cuestiones como el equilibrio o la representatividad.
- *Corpus equilibrado*: aquél que está formado por textos que representan diversas variedades de la lengua en porcentajes similares.
- *Corpus piramidal*: aquél cuyos textos se distribuyen en niveles según su riqueza temática. Cuanto mayor sea la variedad temática, más alto será su nivel y menor será el número de textos de cada variedad.
- *Corpus monitor*: aquél que se encuentra en constante actualización de textos, incluyendo y eliminando textos cada cierto tiempo, lo que da lugar a un corpus vivo y dinámico como la propia lengua (Torruella y Llisterri, 1999: 10). También se habla de *corpus abierto* (Bowker y Pearson, 2002: 12-13).
- *Corpus paralelo*: aquél formado por un conjunto de textos en lengua original junto con sus correspondientes traducciones (a una o varias lenguas). Cuando hay solo dos lengua implicadas (lengua original y sus traducciones a una sola lengua meta), se denomina *corpus paralelo bilingüe*; mientras que cuando el corpus cuenta con la lengua original y más de una lengua meta se habla de *corpus paralelo multilingüe*.
- *Corpus comparable*: aquél que está constituido por un conjunto de textos originales de similares características, a saber: pertenecen al mismo campo de especialidad o temática, comparten la misma función comunicativa y son de contenido similar. Al igual que en el tipo anterior, podemos distinguir entre el *corpus comparable bilingüe*, que se limita a la compilación de textos originales en dos lenguas y el *corpus comparable multilingüe*, que incluye tres o más lenguas. Además, en este caso, también se encuentra el *corpus comparable monolingüe*, que incluye textos originales y traducidos en un mismo idioma.

Un segundo criterio de clasificación está relacionado con la especificidad o la funcionalidad de los textos que forman los corpus, que se pueden dividir en:

- *Corpus general*: aquél que es representativo de la lengua común y, por tanto, incluye textos pertenecientes a situaciones comunicativas cotidianas o normales. Corpas Pastor (2001: 156) añade que, por oposición al corpus general, aparecen los siguientes corpus restringidos: corpus especializado, genérico, canónico, periódico o cronológico y diacrónico.
- *Corpus especializado*: aquél que, por oposición al corpus general, está formado por textos que hacen referencia a un tipo particular de lengua y que se han compilado siguiendo unos criterios establecidos previamente para alcanzar un propósito específico.

Dentro de este último grupo, algunos autores distinguen entre *corpus especializado*, *corpus especial* y *corpus para fines específicos* (Pérez Hernández, 2002), pero la mayoría de los autores están de acuerdo en que son términos que se utilizan a menudo indistintamente (Pearson, 1998: 45). Según EAGLES (1996: 7):

Those which do not contribute to a description of the ordinary language, either because they contain a high proportion of unusual features, or their origins are not reliable as records of people behaving normally.

- *Corpus genérico*: aquél que está formado por textos pertenecientes a un determinado género.
- *Corpus canónico*: aquél que contiene las obras completas de un autor determinado, independientemente de los géneros.
- *Corpus periódico o cronológico*: aquél que está formado por textos que se han producido durante un periodo de tiempo concreto. Este tipo de corpus también recibe el nombre de *corpus sincrónico* (Bowker y Pearson, 2002: 12).
- *Corpus diacrónico*: aquél que incluye textos de diferentes periodos de tiempo con objeto de realizar estudios sobre la evolución de la lengua. También se denomina *corpus histórico* (Abaitua, 2002).

En tercer lugar, encontramos el criterio de clasificación relacionado con la cantidad de palabras de cada texto. Así pues, se distingue entre:

- *Corpus textual*: aquél que incluye textos completos a fin de ofrecer una muestra representativa de la lengua común o de sus variedades más relevantes. Se trata del corpus más utilizado para la creación de diccionarios, gramáticas y otros recursos similares.
- *Corpus de referencia*: aquél que incluye fragmentos de textos, normalmente todos de un tamaño similar, cuyo objetivo recae en el aspecto de la lengua representado. El corpus de referencia cuyo objetivo es el aspecto léxico de los textos que lo forman se denomina *corpus léxico*.

El siguiente criterio de clasificación, la codificación y anotación, hace referencia a una de las características señaladas por el informe EAGLES (1996: 4), la simplicidad. Como ya apuntábamos anteriormente, la simplicidad hace referencia al formato en el que se encuentra el texto, según el tipo de información lingüística o metatextual que se le añade al texto. Este criterio permite distinguir entre dos tipos de corpus:

- *Corpus simple o no anotado*: aquél que se encuentra en formato de solo texto o ASCII y no presenta ningún tipo de etiquetado.
- *Corpus anotado*: aquél a cuyos textos se les han añadido etiquetas para anotar diferentes aspectos lingüísticos (morfosintaxis, sintaxis, semántica, pragmática) o metatextuales (indicación del título, del autor, de la estructura del texto, etc.), ya sea manual o automáticamente.

En el quinto lugar, encontramos el criterio de la documentación, que permite clasificar los corpus en:

- *Corpus no documentado*: aquél que carece de archivos asociados o cabeceras que indiquen algún tipo de referencia del texto.
- *Corpus documentado*: aquél que tiene incluida una cabecera (*header*) o un archivo DTD (*Document Type Definition*) con objeto de identificar su procedencia.

Por último, debemos distinguir entre el modo en que se encuentran los datos primarios del corpus. Así, se pueden dividir en los siguientes:

- *Corpus oral*: aquél que en el que se recogen transcripciones de conversaciones informales y espontáneas, que han tenido lugar sin la intervención de ningún medio de comunicación.⁶⁶
- *Corpus escrito*: aquél que está constituido por textos escritos, que se han elaborado para ser leídos y no para ser transmitidos oralmente.
- *Corpus multimodal*: aquél que está formado por diferentes materiales como pueden ser vídeo, audio, transcripción y textos escritos.

Y, por último, otro criterio que establece la autora (2008: 117-118) atiende al carácter profesional del corpus. Así pues, establece la siguiente división:

- *Corpus profesional*: aquél que estaría formado por todas aquellas traducciones realizadas por profesionales de la traducción.
- *Corpus semiprofesional*: aquél formado por todas aquellas traducciones realizadas por estudiantes de los últimos cursos de la licenciatura y grado de Traducción e Interpretación.⁶⁷

Todos estos criterios de clasificación que hemos venido observando, los cuales tienen en cuenta los documentos que forman el corpus, ya sean estos textos completos o fragmentos, sean especializados o generales, sean monolingües o bilingües, etc., se pueden englobar a su vez bajo una tipología de corpus en función del diseño puntual o estable y el método de compilación empleados (Corpas Pastor, 2004b: 226-227). Así pues, se puede realizar una distinción funcional entre los *corpus virtuales*, también denominados corpus *ad hoc* (Corpas Pastor, 2001), *electronic* (Corpas Pastor, 2001), *disposable corpora* (Varantola, 2002), *do-it-yourself/DIY corpora* (Zanettin, 2002), *web corpora* (Fletcher, 2004), en contraposición a los denominados *corpus estables* (*robust*,

⁶⁶ El informe EAGLES pone de manifiesto que no existe consenso sobre lo que debe considerarse un corpus oral. Para algunos académicos, se trata de la definición ofrecida arriba; sin embargo, en otros casos, el uso del término se amplía para referirse a cualquier tipo de lengua en la que los hablantes se comportan de forma oral, como por ejemplo en los textos escritos para ser hablados.

⁶⁷ Siguiendo esta terminología, se podría establecer también el tipo de *corpus no profesional*, que estaría formado por traducciones realizadas por estudiantes de Traducción e Interpretación de los primeros años, así como por otras personas sin formación específica en traducción.

proper o *ready-made corpus*). Los primeros están formados por textos que se encuentran originalmente en formato electrónico, puesto que, por lo general, proceden de la red Internet, aunque también pueden ser textos extraídos de CD-ROM o incluso de memorias de traducción (en formato .TMX). Además, no suelen tener ningún tipo de anotación o etiquetado y raras veces contienen referencias de las fuentes electrónicas de las cuales proceden. Y, por último, se caracterizan por la finalidad concreta para la que han sido elaborados (docencia, traducción, investigación, etc.) y por su facilidad para ser compilados.⁶⁸

Por su parte, los segundos son los corpus de referencia más tradicionales y se diferencian de los anteriores principalmente en el hecho de que su tamaño suele ser mucho mayor y fijo y los usuarios los encuentran ya listos para su explotación mediante programas de recuperación de información o gestión de corpus. De esta manera, mientras que en los primeros el compilador y el usuario pueden converger en la misma persona, es decir, el usuario puede crearlo a partir de fuentes electrónicas y posteriormente utilizarlo; en los segundos, el usuario rara vez es el propio compilador y utiliza el corpus ya totalmente compilado. Como afirma Corpas Pastor (2004b: 227), no es fácil establecer límites precisos entre ambos tipos de corpus, y estos representan más bien los dos extremos ideales de una escala a lo largo de la cual se sitúan todos los corpus disponibles en la actualidad.

Para terminar, es necesario destacar que están apareciendo concepciones nuevas sobre los corpus, como el uso de la red Internet a modo de corpus gigantesco (*web as corpus*) (Kilgariff, 2001; Kilgariff y Grefenstette, 2003),⁶⁹ que pueden dar lugar a un cambio de perspectiva, tanto en el diseño como en la clasificación de los corpus. No obstante, como veremos en el capítulo siguiente, para utilizar los datos de la red Internet como corpus para fines terminográficos, la información debe ser correctamente evaluada, seleccionada y organizada de un modo sistemático, de acuerdo con unos

⁶⁸ Cada vez más se emplea el término «corpus electrónico», sinónimo de «corpus virtual» o «Electronic corpora» o «Web corpora», para denominar a los corpus elaborados con material electrónico, ya sea extraído de Internet, de CD-ROM, de material escaneado o de otro formato similar, y diferenciarlos así de los corpus con una función más inmediata y *ad hoc*. De esta manera, todos los corpus *ad hoc* serían considerados corpus electrónicos, pero no a la inversa.

⁶⁹ Se están desarrollando herramientas gratuitas y comerciales que permiten la compilación de corpus directamente de Internet, como por ejemplo WebCorp (<http://www.webcorp.org.uk/>), WebBootCaT (<http://www.sketchengine.co.uk/>), Jaguar (<http://rc16.upf.es/cgi-bin/jaguar/jaguar.pl>), entre otras.

criterios internos y externos explícitos, por lo que la aplicación de la nueva concepción de *web as corpus* no está todavía muy extendida dentro del ámbito terminográfico.

Hasta ahora, hemos venido tratando cuestiones generales sobre los corpus, a fin de proporcionar una visión global de la definición y diseño de los corpus. Una vez realizada dicha introducción, pasaremos a detallar la utilización de los corpus en el trabajo terminográfico y, en consecuencia, a centrarnos en los corpus especializados.

1.4.2.3. Criterios específicos para la elaboración y diseño de los corpus especializados para la terminografía

En la tipología de corpus que ofrecimos en el apartado anterior, definíamos un *corpus especializado* como un tipo de *corpus especial* que ha sido diseñado con un propósito específico y que tiene la finalidad de ser representativo de un tipo particular de lengua, como por ejemplo un campo de especialidad o un grupo particular de hablantes. Será, pues, este tipo de corpus el que se tendrá en cuenta en la labor terminográfica y al que dedicaremos los siguientes apartados de este trabajo de investigación.

La especificidad de la Terminografía basada en corpus (Meyer y Mackintosh, 1996: 258), en contraposición a la Lexicografía basada en corpus u otras aplicaciones de los corpus (traducción, enseñanza de segundas lenguas, etc.), obliga al establecimiento de una serie de requisitos o criterios específicos para el trabajo terminográfico. Algunos de ellos serán comunes a los criterios generales de la compilación y diseño de los corpus y otros, como veremos, presentarán ciertas diferencias.

Antes de comenzar con los criterios específicos de la compilación de corpus en terminografía, consideramos necesario exponer brevemente las fases en las que se divide el trabajo de un terminógrafo,⁷⁰ con objeto de indicar explícitamente las necesidades del empleo de los corpus en cada fase y, así, poner de relieve la importancia de la compilación de los corpus: en primer lugar, los terminógrafos deben familiarizarse con el dominio en el que están trabajando, a fin de establecer sus límites, relaciones con

⁷⁰ Estas fases se tratan en profundidad en el capítulo siguiente.

otros dominios y la organización interna de este, es decir, los subdominios en los que puede dividirse; en segundo lugar, deben identificar las fuentes de conocimiento que les proporcionarán tanto la información lingüística como conceptual y comunicativa de los términos; en tercer lugar, los terminógrafos pasan a considerar, con la ayuda de las fuentes de conocimiento, un conjunto de candidatos a término para empezar a trabajar y crear la conceptualización del dominio; en cuarto lugar, analizan la nomenclatura identificada en la fase previa, así como la información terminológica de los textos compilados (colocaciones, relaciones semánticas, etc.) y elaboran la base de datos terminográfica extrayendo del corpus información para realizar definiciones y seleccionar contextos adecuados; y, por último, resuelven los posibles problemas presentados, realizan las validaciones y editan el recurso terminológico en cuestión. En estas fases, se observa la importancia que tiene la documentación en la labor terminográfica y, por ende, la relevancia que presentan los corpus electrónicos para el terminógrafo, tanto a la hora de obtener un trabajo sistemático y objetivo como para facilitar y acelerar las tareas de cada fase. Así pues, la documentación, y en consecuencia, los corpus y las herramientas de análisis de corpus permiten llevar a cabo:

- La adquisición de conocimiento conceptual y la familiarización del terminógrafo no experto del dominio mediante la consulta de documentos, a fin de identificar la estructura interna del dominio, las relaciones con otros campos de especialidad y las fuentes de conocimientos adecuadas.
- La identificación de unidades terminológicas dentro de un discurso de especialidad, es decir, la nomenclatura.
- El análisis y la preparación de entradas mediante la adquisición de información lingüística, pragmática y semántica extraídas del corpus, como por ejemplo definiciones, contextos o colocaciones.
- La detección y descripción de nuevos conceptos, así como la identificación de etiquetas léxicas que se están atribuyendo a dichos conceptos dentro del dominio. En algunos casos, también la propuesta de un neologismo adecuado (cuando todavía no se ha acuñado un término en una lengua).

- La estandarización de términos sinónimos o cuasisinónimos que son utilizados por diferentes expertos.
- La clarificación de dudas y preguntas sobre inconsistencias y otros asuntos que, de lo contrario, solo se podría llevar a cabo mediante consultas a expertos del dominio.

Una vez expuestas las fases del trabajo terminográfico así como las funciones de la documentación en Terminografía, podemos resumir los roles de la documentación en dos: por un lado, para el análisis lingüístico (extracción de términos, elaboración de definiciones, etc.), y por otro, para la adquisición de conocimiento (conceptualización del dominio, relaciones semánticas, etc.). A fin de poner en práctica estos roles de la documentación desde el punto de vista lingüístico, pragmático y conceptual, será necesario que los corpus estén compilados de forma apropiada y según unos criterios generales y específicos establecidos dentro de un contexto terminográfico.

Dentro de estos criterios requeridos, podemos reconocer una serie de criterios generales de compilación que son comunes a cualquier rama lingüística que utilice corpus textuales, a saber: cantidad, calidad, documentación y simplicidad, aunque en terminografía presentan unas características concretas que deben ser tenidas en cuenta durante la compilación. Además de estos criterios generales, encontramos otros específicos para el trabajo terminográfico: delimitación de fronteras, apertura y pragmática. A continuación, comentamos los diferentes criterios generales y específicos en el contexto terminográfico:

El criterio de la cantidad que nos indicaba Sinclair (EAGLES, 1996:4) (cf. 1.4.2.2.1.) es un aspecto polémico para la compilación de corpus especializados. Para algunos (Meyer y Mackintosh, 1996: 268), un corpus especializado puede ser mucho más pequeño que uno de propósito general, pero para otros no sería necesario poner un límite a la cantidad de texto que se recopile (Pearson, 1998: 57), siempre y cuando siguiera unos criterios establecidos previamente. En definitiva, podemos confirmar que actualmente no existe un acuerdo acerca del volumen que debe tener un corpus especializado para que sea representativo. Desde la lexicografía basada en corpus se recomienda que el volumen del corpus sea lo más elevado posible; por el contrario, en terminografía se hace mayor hincapié en la densidad terminológica que en el volumen

propio del corpus utilizado y, así pues, se considera que un número de palabras elevado no determina la representatividad del corpus. En este sentido, no es posible establecer *a priori* el número de textos o de palabras que necesita un corpus para ser representativo, puesto que dependerá de diferentes variables (género textual, lengua, lenguaje de especialidad, nivel de especialización, etc.), aunque sí es posible medir actualmente su representatividad *a posteriori* mediante el uso de aplicaciones informáticas como *ReCor*, una herramienta informática que permite estimar de forma objetiva y cuantitativa la representatividad de los corpus electrónicos en función de su densidad terminológica una vez que se han compilado (cf. Seghiri Domínguez, 2006; Corpas Pastor y Seghiri Domínguez, 2006, 2007a, 2007b, 2011/en prensa).⁷¹

Con respecto al criterio de calidad, los textos deberán cumplir una serie de requisitos: en primer lugar, los textos deben ser recientes y actuales, es decir, se debe compilar un corpus sincrónico formado por textos que denoten el estado actual del dominio, a fin de proporcionar los términos utilizados en el campo de especialidad en ese momento y la información conceptual actualizada; en segundo lugar, los textos deberán estar escritos por diferentes autores expertos (ya sean especialistas, instituciones, organizaciones, etc.) en la medida de lo posible para ofrecer fidelidad en el contenido y neutralizar cualquier tipo de idiosincrasia del autor; en tercer lugar, los textos deben ser preferentemente originales y no traducciones, ya que solo de esta manera se puede asegurar que la terminología empleada es la adecuada y la original en cada ámbito especializado,⁷² así como es recomendable también que los textos estén escritos por hablantes nativos, para evitar cualquier tipo de influencia de la lengua materna del autor, errores (léxico, gramaticales, etc.) y cualquier giro o uso no correcto; en cuarto lugar, los textos incluidos en los corpus deben ser textos completos, es decir, se debe evitar el uso de fragmentos con objeto de evitar la posible descontextualización o error en la información contenida; y, por último, según Pearson (1998: 60), los textos deben haber sido publicados previamente a su inclusión, con la idea de que adquieran

⁷¹ En nuestro corpus, hemos empleado esta herramienta para estimar su representatividad, cuyos resultados y conclusiones se pueden consultar en el capítulo 3 (cf. 3.4.).

⁷² A pesar de que el corpus compilado para un trabajo terminográfico debe ser comparable, es decir, formado por textos originales, para poder extraer la terminología adecuada del ámbito de especialidad en cada lengua de trabajo, también suele haber un subcorpus paralelo, es decir, constituido por textos originales y sus traducciones, que permite detectar y extraer los posibles equivalentes en otras lenguas, siempre y cuando sea un trabajo bilingüe o multilingüe.

un valor más formal, serio y respetable dentro del ámbito especializado. Asimismo, si el trabajo terminográfico está enfocado para unos países y hablantes determinados, se deberá establecer unos límites geográficos para la búsqueda de los textos y, por tanto, el terminógrafo deberá comprobar que todos los textos pertenecen o han sido publicados dentro de esos límites geográficos. Con todos estos criterios de calidad se pretende conseguir la mayor fiabilidad posible de los textos incluidos en un corpus con un objetivo terminográfico.

El criterio de documentación ocupa un lugar muy relevante en la compilación, ya que permite llevar a cabo un trabajo sistemático y homogéneo. Consiste en registrar las referencias de los textos contenidos en el corpus (autor, lugar de publicación, fecha o actualización, etc.) con objeto de realizar un seguimiento de los textos seleccionados y de las fuentes de información utilizadas. También es recomendable utilizar un código unívoco entre el registro del texto y el propio texto para establecer una vinculación entre los datos registrados y el texto que nos permita acceder a su información de referencias.

Por último, el criterio de simplicidad hace referencia al tipo de información añadida al texto original. Esta información es principalmente morfológica, sintáctica, léxica y semántica y permite al terminógrafo utilizar el corpus de forma más eficiente y más precisa a la hora de realizar búsquedas, estudios concretos y clasificar la información contenida en los textos. En terminografía, las anotaciones más utilizadas son las semánticas y léxicas, que se utilizan para el estudio del discurso especializado, extracciones terminológicas o de patrones semánticos, aunque también se puede encontrar otro tipo de estudios lingüísticos. Además de estas anotaciones en el texto, se pueden llevar a cabo anotaciones similares a las indicadas en el criterio de documentación aunque en el mismo texto (autor, institución, fecha de producción, género textual, nivel de especialización, etc.).

Además de estos criterios generales comentados hasta el momento, consideramos que son útiles otros, de carácter más específico en el contexto terminográfico, aunque muy relacionados con los anteriores:

El criterio de delimitación de fronteras, que exige que, antes de comenzar con cualquier proyecto, los terminógrafos delimiten el campo de especialidad que tiene que representar. Como indica Meyer y Mackintosh (1996: 268), esta delimitación deberá ser

de dos tipos: «side boundaries», es decir, con respecto a los campos de especialidad más cercanos al objeto de estudio y «upper boudaries», con relación a los niveles de especialización que va a incluir el corpus (de más especializado a menos). En relación con este criterio, los terminógrafos deberán compilar un corpus que represente lo mejor posible el campo de especialidad delimitado y, para ello, deberán seleccionar textos que permitan un equilibrio en todos los aspectos del dominio, cubriendo todos los subdominios y dominios relacionados de la forma más equitativa posible.

Otro criterio específico es la apertura del corpus. En Terminografía, debido a la rapidez con la que se producen los cambios, el corpus especializado debe ser abierto, a fin de que se pueda ir actualizando, ya sea eliminando o incluyendo textos, con el paso del tiempo.

Por último, encontramos el criterio de pragmática, que se encuentra en la misma línea que el criterio de delimitación de fronteras, aunque en este caso hace referencia a la situación comunicativa para la que va destinado el producto final. Este criterio exige tener en cuenta la situación comunicativa, a saber: los receptores, el contexto, el tipo textual y el nivel de especialización que habrá en el momento en el que se emplee el recurso final. Así pues, dependiendo de quiénes sean los usuarios, para qué utilicen y cuándo utilicen el recurso final, el corpus estará formado por un tipo de textos u otros.

En nuestra opinión, estos criterios generales y específicos expuestos anteriormente serían los criterios que deberían tenerse en cuenta para compilar cualquier corpus especializado en un contexto terminográfico y, en definitiva, para garantizar la calidad del producto final. De lo contrario, la calidad del producto no estaría garantizada.

A continuación, presentaremos las principales herramientas y técnicas existentes para llevar a cabo el análisis de un corpus especializado, con las que se pueden satisfacer las necesidades específicas de los terminógrafos.

1.4.2.4. Herramientas y técnicas de análisis y procesamiento de corpus: necesidades específicas de los terminógrafos

El trabajo terminográfico sistemático orientado a la elaboración de recursos especializados, ya sean monolingües, bilingües o multilingües, recibe la asistencia de aplicaciones informáticas en prácticamente todas sus fases, especialmente en la fase de análisis y preparación del trabajo; la fase de recopilación de información (corpus y recursos terminológicos); la fase de gestión terminológica y conceptualización del dominio; la fase de la elaboración de la base de datos ontoterminográfica y, por último, la fase de resolución de problemas, validación y edición de la terminología.⁷³ Con la ayuda de las herramientas informáticas se consigue especialmente mejorar los resultados del trabajo terminográfico y agilizar las tareas del terminógrafo, que ve reducido el trabajo largo y tedioso que realizaba anteriormente de forma manual.

Cada vez encontramos más, mejores y más variadas herramientas informáticas disponibles dirigidas al terminólogo, tanto de libre acceso como comerciales, y continúan las investigaciones de aplicaciones que beneficien el trabajo terminográfico en todas sus fases; no obstante, sigue habiendo gran cantidad de herramientas que no distinguen su aplicación en la Lexicografía y en la Terminografía y, por tanto, hacen caso omiso a las necesidades dispares entre los lexicógrafos y los terminógrafos, como se ha venido defendiendo hasta el momento en esta investigación.

Actualmente, encontramos investigaciones y proyectos dirigidos a la creación o mejora de diferentes tipos de herramientas: por un lado, herramientas que se dedican a una única tarea dentro de todo el trabajo terminográfico, como por ejemplo extractores de terminología y de concordancias, anotadores sintácticos, etc.; y, por otro, paquetes de herramientas, es decir, paquetes donde se integran diferentes tipos de herramientas que se emplean en un trabajo terminográfico, por ejemplo un paquete donde se incluyera un extractor de terminología y de concordancias, un analizador de frecuencias, etc. La tendencia actual se dirige, no obstante, hacia la creación de aplicaciones donde se integren todas o casi todas las herramientas que se necesitan para llevar a cabo un trabajo terminográfico cualquiera desde el principio hasta su edición. Como ejemplo de

⁷³ Remitimos al capítulo 2 para consultar las fases de un trabajo ontoterminográfico.

ello, tenemos actualmente TERMINUS (del grupo IULATERM),⁷⁴ que integra herramientas para llevar a cabo la mayoría de las tareas de un trabajo sistemático, por ejemplo incluye una aplicación que permite la elaboración de mapas conceptuales, la gestión y compilación de corpus (búsqueda y recuperación de textos de Internet,⁷⁵ extracción terminológica, análisis de frecuencia y concordancias), la gestión de términos (registro y manipulación de términos y su información asociada a la base de datos que incluye) y la edición final del recurso. A pesar de ser una herramienta muy completa e integradora, carece de ciertas aplicaciones que serían necesarias para un trabajo terminográfico. De acuerdo con Vargas Sierra (2009: 43), estas tareas pendientes serían las siguientes:

- (a) la digitalización de textos en papel (uso de un escáner, selección de los fragmentos para ser procesados por un programa de reconocimiento óptico de caracteres, y revisión ortográfica a fin de detectar los errores de reconocimiento);
- (b) el registro de los atributos textuales (datos bibliográficos, función principal del texto, tenor, lengua, nombre del fichero electrónico, temática, etc.);
- (c) el etiquetado del corpus;
- (d) la búsqueda y visualización de concordancias en forma bilingüe;
- (e) el acceso a otros recursos terminológicos de referencia en línea para su consulta;
- y (f) la edición personalizada del repertorio para su publicación.

De esta manera, podemos observar que todavía quedan metas por alcanzar en el campo de la terminótica, a pesar de que se están realizando grandes avances en este terreno.

Como hemos venido diciendo, la aplicación de herramientas informáticas en las diferentes fases del trabajo terminográfico está muy extendido, pero en este apartado trataremos únicamente las herramientas relacionadas con la gestión y procesamiento del corpus, que son realmente las que ocupan un lugar preponderante en este tipo de trabajo. En primer lugar, hablaremos de las herramientas de forma individual y, posteriormente, presentaremos algunos de los paquetes informáticos más utilizados en la actualidad.

En nuestra opinión, las herramientas necesarias para trabajar con un corpus especializado con fines terminográficos una vez que ha sido compilado se pueden clasificar en dos grupos diferentes: por un lado, las herramientas destinadas al

⁷⁴ <http://melot.upf.edu/Terminus2009/index_es.html>

⁷⁵ Recordamos aquí la concepción de la compilación de corpus de *web as corpus*, es decir, la búsqueda y recuperación de textos a través de buscadores en Internet.

preprocesamiento de los textos incluidos en el corpus; y por otro, las herramientas para la gestión y análisis textual.

El primer grupo, las herramientas de preprocesamiento del corpus, tienen como objetivo preparar los textos que forman parte del corpus para su posterior análisis. En este sentido, encontramos diferentes tipos de aplicaciones útiles para la labor terminográfica:

- Aplicaciones que realizan análisis gramaticales del corpus, como son los *Part-of-Speech taggers* (*POS taggers* o etiquetador gramatical),⁷⁶ cuya finalidad consiste en asignar a cada palabra de un texto su categoría gramatical, es decir, indicar si se trata de un sustantivo, un adjetivo, un verbo auxiliar, etc. Es el tipo de anotación más básica y, a menudo, es una fase previa a la extracción terminológica, como veremos más abajo, y para distinguir homógrafos. Podemos observar las etiquetas POS en la columna situada en el centro de la Tabla 6.

| | | |
|-------------|--------------------|-------------|
| POR | PREP ⁷⁷ | por |
| SIERRA | NC | sierra |
| DE | PREP | de |
| LAS | ART | las |
| NIEVES | NC | nieve |
| Senderismo | NC | senderismo |
| Descripción | NC | descripción |
| El | ART | el |
| Pinsapo | NP | <unknown> |
| es | VSfin | ser |
| un | ART | un |
| abeto | NP | abeto |
| endémico | ADJ | endémico |
| del | PDEL | del |
| sur | PE | sur |
| de | PREP | de |
| España | NP | España |
| Este | DM | esto |
| tipo | NC | tipo |
| de | PREP | de |
| árbol | NC | árbol |
| solo | ADV | solo |

Tabla 6. Ejemplo de etiquetado gramatical (*POS-tagging*) y lematización con *TreeTager*.⁷⁸

⁷⁶ Para profundizar en el tema, se recomienda consultar Voutilainen (2003).

⁷⁷ Los códigos de las etiquetas se pueden consultar en el siguiente enlace: <<http://www.ims.uni-stuttgart.de/ftp/pub/corpora/spanish-tagset.txt>>

⁷⁸ Las herramientas que hemos utilizado para elaborar la base de datos terminológica, así como para realizar búsquedas de términos (que se presentan en el capítulo cuarto de esta tesis doctoral) utiliza el

- Aplicaciones que realizan análisis morfológicos del corpus, como son los lematizadores (*lemmatizers*). Estas herramientas llevan a cabo la relación automática de palabras que morfológicamente pertenecen a una misma familia, las cuales se reducen a una forma común o lema. Por ejemplo, las formas *comí*, *comían*, *han comido*, que comparten la misma raíz, corresponderían a un mismo lema: *comer*. El objetivo de esta herramienta consiste principalmente en la eliminación de recuentos múltiples de palabras, ya que las variantes morfológicas anteriores no se contarían como palabras distintas, sino como ocurrencias de de una misma palabra (el lema correspondiente) (véase la columna de la derecha de la Tabla 6).

- Aplicaciones que realizan análisis sintácticos del corpus (*parsing*),⁷⁹ como son los analizadores sintácticos o *parsers*. Con estas herramientas, el objetivo consiste en producir una representación de la estructura de la frase que refleje sus relaciones sintácticas mediante reglas gramaticales. Los corpus que han sido anotados sintácticamente se denominan *treebanks* y en ellos se observa lo siguiente:

```
the dog.n was.v playing.v with a ball.n on the yard.n
      Constituent tree:
      (S (NP the dog)
        (VP was
          (VP playing
            (PP with
              (NP a ball))
            (PP on
              (NP the yard))))))
```

Figura 5. Ejemplo de etiquetado sintáctico de la frase «The dog was playing with a ball on the yard» con el analizador sintáctico *Link Parser*.⁸⁰

programa *TreeTagger* (<http://www.ims.uni-stuttgart.de/projekte/corplex/TreeTagger/>), un etiquetador POS de libre acceso en Internet que ha sido desarrollado por el Instituto de Lingüística Computacional de la Universidad de Stuttgart, pero existen otros muchos programas disponibles, como *CLAWS* (<http://ucrel.lancs.ac.uk/claws/>) o *EngCG-2* (<http://www.conexor.fi/analysers.html>).

⁷⁹ Para profundizar en el tema, se recomienda consultar Carroll (2003).

⁸⁰ <<http://www.link.cs.cmu.edu/cgi-bin/link/construct-page-4.cgi#submit>>. Otros ejemplos de anotadores sintácticos son: *VISL* (Visual Interactive Syntax Learning) (<http://visl.hum.sdu.dk/itwebsite/visl/>), *The*

- Aplicaciones que realizan anotación semántica del corpus (*semántica tagging*). En este grupo, se pueden identificar dos tipos de anotación semántica: a) anotación de relaciones semánticas entre unidades del texto, por ejemplo entre pacientes y agentes de determinadas acciones; y b) anotación de categorías conceptuales en el texto. En general, este tipo de anotación sirve para incluir indicaciones sobre las relaciones semánticas de una unidad con otras unidades dentro de un dominio de especialidad, para obtener desambiguación léxica y para establecer las categorías conceptuales de las unidades lingüísticas.

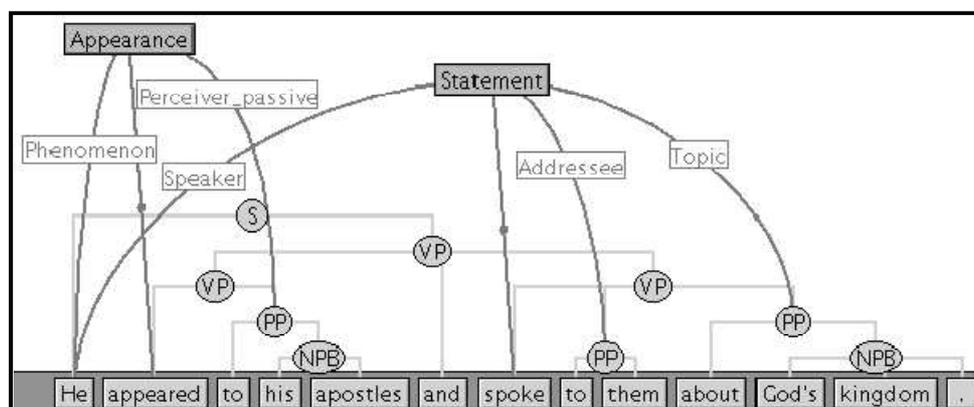


Figura 6. Ejemplo de anotación semántica del treebank *TIGER*.⁸¹

- Aplicaciones que realizan el alineado de corpus bilingües paralelos (*corpus alignment* o *bitext*),⁸² que consisten en identificar las correspondencias de traducción a nivel de palabra, de unidad fraseológica, de oración, de párrafo o de estructura entre el subcorpus de origen y su subcorpus meta. El objetivo de estas herramientas consiste en facilitar la identificación de unidades terminológicas y sus equivalentes al permitir la visualización de las unidades de forma próxima. Por ejemplo, en el caso de

Proximity Parser (<http://proximity.franklin.com/Parse.htm>) Conexor's English Analyzers (http://www.connexor.com/demos/syntax_en.html), entre otros.

⁸¹ Ejemplo extraído del FrameNet alemán, *SALSA* (<<http://www.coli.uni-saarland.de/projects/salsa/>>). Otros anotadores semánticos son, por ejemplo: KIM Platform (<http://www.ontotext.com/kim/index.html>), Ontomat (<http://annotation.semanticweb.org/ontomat/index.html>), SemConst (<http://sourcesup.cru.fr/xmg/SemConst/>), etc.

⁸² Para profundizar en el tema, se recomienda consultar Véronis (2000).

una alineación a nivel de párrafo, se podría visualizar el párrafo del texto original en la parte de la izquierda y la del texto meta en la de la derecha (Figura 7) o el párrafo original en la parte superior y su traducción debajo:

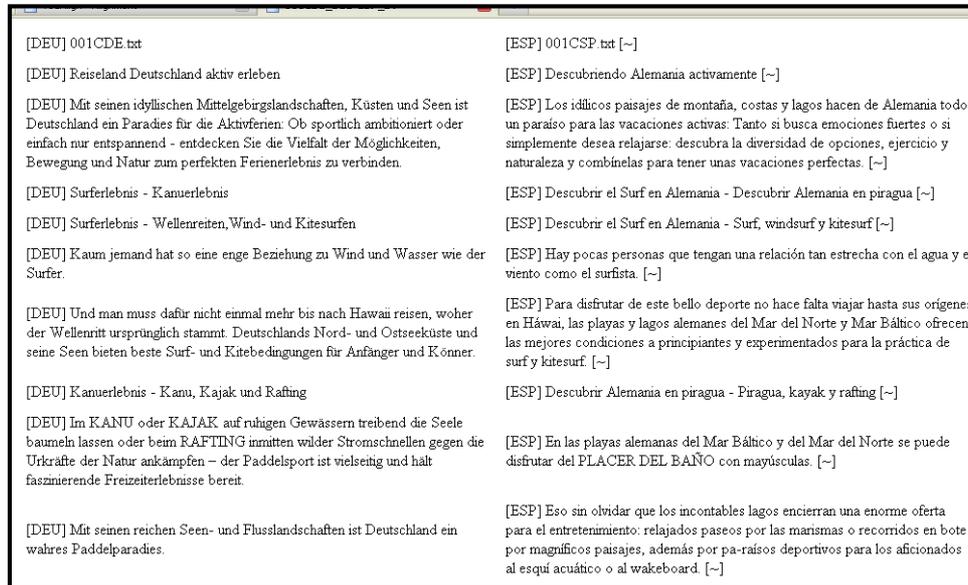


Figura 7. Ejemplo de alineación bilingüe a nivel de párrafo con *YouAlign*.⁸³

Estas herramientas que hemos visto hasta ahora nos permiten realizar un preprocesamiento de los textos que forman el corpus especializado, las cuales se pueden emplear en cualquier ámbito lingüístico, es decir, no estarían limitadas a su uso en la Terminografía. Debido a la importancia de la adecuada anotación de los corpus para la consecución de los objetivos planteados al comienzo de un proyecto, EAGLES (1996b) elaboró una serie de recomendaciones para la anotación morfosintáctica de los corpus, *Recommendations for the Morphosyntactic Annotation of Corpora*, a fin de establecer unos estándares y unas etiquetas comunes que facilitarían la reutilización de los corpus así como la comprensión de las etiquetas entre expertos. Estas recomendaciones son, a día de hoy, utilizadas a la hora de anotar los corpus y de desarrollar nuevos anotadores lingüísticos.

⁸³ <<http://www.youalign.com/>>. Otros ejemplos de alineadores pueden ser: WinAlign (dentro del paquete de programa Trados), Tagaligner (<http://sourceforge.net/projects/tag-aligner/>), NATools (<http://linguateca.di.uminho.pt/natools/>), Giza++ (<http://www.fjoch.com/GIZA++.html>), etc.

A pesar de que estas aplicaciones de preprocesamiento se encuentran muy extendidas en el ámbito de la Lingüística Computacional, así como en otras disciplinas lingüísticas, la tendencia en las tareas terminográficas se inclina más al análisis y gestión del corpus “en crudo” (*raw corpus*), es decir, un corpus que carece de etiquetados (Orăsan et ál., 2007: 27), que a un corpus anotado, ya sea sintácticamente, semánticamente, etc. No obstante, el terminógrafo también utiliza estos etiquetados en algunas ocasiones, por ejemplo para la extracción terminológica basada en reglas morfosintácticas (como veremos a continuación), para la extracción semiautomática de relaciones semánticas, etc. El principal problema de esta anotación se basa en la dificultad de producir corpus anotados y en la imposibilidad de reutilizarlos en otros dominios o, incluso, en otros recursos para el mismo dominio. Por esto, en Terminografía, no se suelen emplear a menudo estas herramientas de preprocesamiento, aunque sí que se emplean en algunos proyectos y sí que pueden ser útiles en momentos puntuales de la tarea terminográfica.

Al contrario que las herramientas de preprocesamiento, el segundo grupo de herramientas, las de análisis y gestión de corpus especializados, son muy populares en Terminografía y se emplean en la mayor parte de los proyectos terminográficos. A continuación, comentamos las más relevantes:

En primer lugar, los extractores de terminología (*term-extraction tools*).⁸⁴ Dentro de este grupo, se pueden distinguir dos grupos: por un lado, las herramientas de extracción de terminología asistida por ordenador; y por otro, los extractores automáticos. Los primeros son muy básicos y consisten en la extracción de los términos una vez que han sido etiquetados por parte del terminógrafo o del experto del campo de especialidad; es decir, en un primer momento se etiquetan unidades lingüísticas del

⁸⁴ Estas técnicas se suelen aplicar a corpus monolingües, aunque también se pueden utilizar con corpus bilingües, concretamente a corpus paralelos siempre y cuando se haya procedido a una alineación previa de los textos, ya sea a nivel de palabra, de oración o de párrafo (Maia, 2003; Lemay et ál., 2005; Guinovart y Simões, 2009). Con los corpus comparables bilingües también se están llevando a cabo extracciones terminológicas, aunque los resultados de la extracción con este tipo de corpus son todavía limitados. Se están utilizando diferentes técnicas como, por ejemplo, la extracción por cognados (Ruslan et ál., 2007), el uso de contextos semilla generados a partir de diccionarios bilingües externos (Gamallo Otero y Pichel Campos, 2008) o pequeños corpus paralelos (Gamallo Otero, 2007). Para profundizar en el tema, véase Jacquemin y Bourigault (2003) y la bibliografía indicada anteriormente. Asimismo, Estopà Bagot (1999) y Cabré Castellví et ál. (2001) realizan un estudio comparativo muy completo de los principales sistemas de extracción de términos disponibles en ese momento.

texto como términos y, posteriormente, el programa extrae dichos términos etiquetados y crea fichas terminológicas para cada uno, en las que también puede incluir el contexto, la fuente original y datos similares. Debido a la necesidad de una identificación manual de los términos, este método no es recomendable para textos (y corpus) de grandes dimensiones, aunque sí que puede ser útil para el análisis de textos de corta longitud. Un ejemplo de este tipo de programas es YVANHOÉ[®], una aplicación desarrollada en la Translation Bureau de Canadá para uso interno.

El segundo tipo, los extractores automáticos de terminología, está mucho más extendido que el primer grupo y sus resultados son relativamente buenos. El objetivo de estas herramientas consiste en la extracción automática de una lista de candidatos a términos que, *a posteriori*, necesita ser examinada, analizada y validada por el terminógrafo antes de su inclusión en la base de datos. En este sentido, a pesar de denominarse extractores automáticos, esta técnica se considera semiautomática, puesto que se necesita la intervención humana para determinar si un candidato a término es considerado, finalmente, un término y para realizar las modificaciones finales pertinentes. Los mecanismos de extracción terminológica se suelen clasificar en tres grupos: lingüísticos, estadísticos e híbridos (Ha et ál., 2008a y 2008b).

El método lingüístico intenta identificar candidatos a términos a través su estructura lingüística (morfológica y sintáctica). Para esta técnica, es imprescindible que los textos hayan sido previamente anotados por anotadores morfosintácticos (*POS taggers*), lematizadores y *parsers*. Dentro de este enfoque, Ha et ál. (2008b) distinguen los métodos intrínsecos, que consisten en la filtración de los candidatos a término según su estructura interna, es decir, según su estructura morfológica (cf. Ananiadou, 1994); y los métodos extrínsecos, que identifican los candidatos a término mediante el análisis de la estructura morfosintáctica de una palabra o secuencia terminológica. Para este último enfoque, se encuentran a su vez diferentes opciones: los marcadores de frontera, de Bourigault (1992, 1995); las secuencias de categoría gramatical (Heid, 1999); los contextos definitorios (Pearson, 1998), entre otros. Como ejemplo de extractor que emplea este tipo de técnica podemos citar el programa LEXTER (Bourigault, 1995).

El método siguiente, el estadístico, se basa en unas reglas estadísticas para obtener los candidatos a términos. La idea que subyace en este tipo de técnicas consiste

en que los documentos especializados se caracterizan por utilizar una serie de términos y combinaciones de palabras que se repiten a menudo. Para ello, estos sistemas intentan detectar candidatos a términos basados en la alta frecuencia de aparición de estas unidades. Esta técnica puede aplicarse de dos formas diferentes: por un lado, la frecuencia de lo que se denominan *n-gramas* (combinaciones de *n* palabras consecutivas); y por otro, con la frecuencia de las unidades que aparecen en un corpus especializado con respecto a un corpus general.⁸⁵

Con la primera técnica, se realizan diferentes análisis del corpus según los *n-gramas* que se seleccionen, a saber: *1-grama*, sería la combinación formada por una única unidad; *2-grama*, sería la que consta de 2 unidades, y así sucesivamente. Esta técnica se completa con la utilización de listas de palabras vacías (*stopwords*), es decir, palabras que carecen de contenido semántico (artículos, preposiciones, números, demostrativos, etc.), con la idea de filtrarlas y eliminarlas de los resultados finales y, así, obtener una lista más reducida de unidades extraídas. Además de la extracción de los *n-gramas*, se trabaja también con la frecuencia. De esta manera, los resultados del análisis con esta técnica se muestran en forma de una lista que contiene los *n-gramas* indicados previamente y ordenados según la frecuencia de aparición en el corpus especializado. Posteriormente, esta lista deberá ser revisada manualmente por el terminógrafo, al igual que la anterior. Un programa que sirve de ejemplo para este tipo de técnicas es *kfNgram*, desarrollado por William H. Fletcher.⁸⁶

Si comparamos esta técnica con la lingüística, podemos prever que con esta el ruido es mucho mayor, ya que lo único en lo que se basa es en la aparición frecuente de combinaciones de palabras y no en el contenido semántico o en la categoría gramatical de estas, que podrían ayudar a la mejor detección de los términos. Sin embargo, esta técnica presenta una ventaja con respecto a la anterior y es su independencia lingüística, es decir, al basarse únicamente en un análisis cuantitativo puede trabajar perfectamente de forma independiente a la lengua de trabajo. De esta manera, se podrá emplear para

⁸⁵ Para profundizar en las diferentes técnicas cuantitativas utilizadas por los extractores automáticos de términos, véase Porta Zamorano (2006).

⁸⁶ <<http://www.kwicfinder.com/kfNgram/kfNgramHelp.html>>

cualquier lengua con la que esté trabajando el terminógrafo, siempre y cuando pueda reconocer los caracteres.

La segunda técnica basada en la estadística es un poco más compleja y sus resultados obtienen una menor cantidad de ruido que la primera, aunque todavía su presencia sigue siendo elevada. Su inconveniente consiste en el hecho de que hay que contar, además de con el corpus especializado de trabajo, con un corpus general que sirva de referencia. Con este método, se compara la frecuencia de unidades y oraciones extraídas de un corpus especializado con su frecuencia en los textos de un corpus general con la idea de que los términos tienden a aparecer con una mayor frecuencia en los textos especializados que en los textos de la lengua general. Un ejemplo de este tipo de extractor puede ser *TermoStat* (Drouin, 2004).⁸⁷

Ambas técnicas estadísticas presentan otro problema además del ruido de los resultados, que es el silencio. Este “silencio” hace referencia a la omisión de candidatos a términos por su baja frecuencia de aparición. Como sabemos, un corpus tiene que ser lo más representativo posible de un campo de especialidad, pero algunos términos, aun formando parte de la terminología de un dominio, presentan una aparición más baja y no por ello deben ser omitidos de la base de datos terminológica. Por este motivo, ha surgido también una técnica híbrida, que consiste en la combinación de métodos estadísticos y lingüísticos, a fin de obtener los mejores resultados posibles (con menos ruido y menos silencio).

Generalmente, la parte principal del algoritmo de este método híbrido es estadístico, aunque también se incorpora información lingüística mediante filtros sintácticos que permiten que únicamente algunas estructuras sintácticas puedan ser consideradas candidatos a término (Ha et ál., 2008). Las estructuras sintácticas que se utilizan normalmente son las siguientes: N (nombre)+ Adj (adjetivo), N + N, Adj + N, N + prep (preposición) + N, etc. Para esta técnica, se necesita, de nuevo, que el corpus de trabajo haya sido previamente anotado, al menos, con un etiquetador morfosintáctico (*POS-tagger*). Los resultados de esta técnica son mejores que los obtenidos por las anteriormente descritas; sin embargo, estos sistemas requieren de un desarrollo más complicado y de una constante actualización, por lo que la creación se realiza

⁸⁷ <http://olst.ling.umontreal.ca/~drouinp/termostat_web>

normalmente en el seno de algún proyecto de investigación y su acceso suele estar restringido. Como ejemplos de este tipo de programas, podemos citar *TermExtractor* (Sclano y Velardi, 2007)⁸⁸ y *YATE* (Vivaldi, 2003).⁸⁹

Las segundas herramientas de análisis y gestión de corpus que vamos a comentar aquí son los analizadores de concordancias (*concordancers*).⁹⁰ Estas aplicaciones se utilizan para la búsqueda de palabras que preceden y siguen al término que se ha introducido en el sistema, es decir, su contexto más próximo; y el resultado consiste en una lista de ejemplos del término en cuestión, que aparece destacado en el centro, junto con las unidades que lo anteceden y lo siguen en la oración en la que aparece. Los datos obtenidos con la búsqueda son establecidos por el usuario y, así, este puede determinar si desea observar una palabra antes y después del término, dos, tres, cinco o más. Con estos resultados, el terminógrafo puede analizar los patrones lingüísticos que aparecen con una determinada frecuencia en el corpus y detectar las relaciones semánticas propias del dominio junto con las combinaciones de palabras más usuales y relevantes del campo de especialidad para, posteriormente, extraerlas e incluirlas en los registros terminológicos correspondientes.

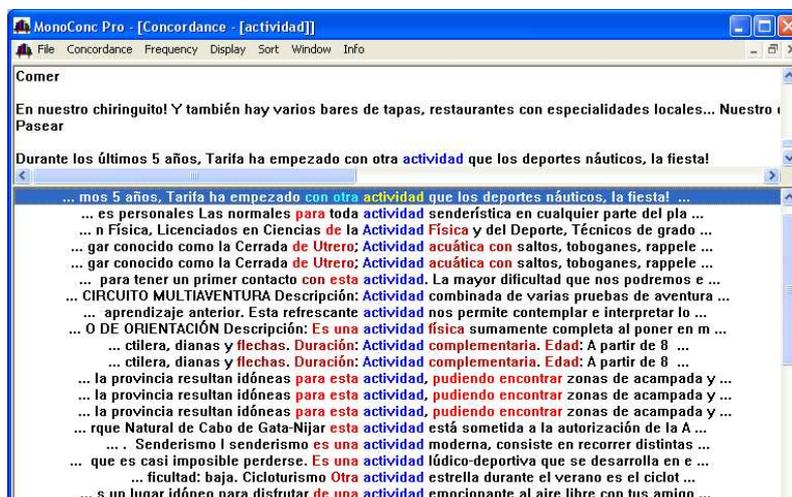


Figura 8. Ejemplo de líneas de concordancia monolingüe con el programa *MonoConc Pro*.⁹¹

⁸⁸ <<http://lcl2.di.uniroma1.it/termextractor/>>

⁸⁹ <<http://igraine.upf.edu/cgi-bin/Yate-on-the-Web/yotwMain.pl>>

⁹⁰ En <<http://www.euralex.org/tools/Concordancer.html>> se puede consultar un informe elaborado por B. M. Schulze y U. Heid sobre los diferentes programas de concordancias disponibles en el mercado.

⁹¹ <<http://www.athel.com/mono.html>>

Además de los analizadores de concordancias monolingües, también existen actualmente otros programas de concordancias bilingües, es decir, que permiten la extracción de líneas de concordancias en dos idiomas, a saber: los textos originales y sus traducciones en un corpus paralelo. Estos programas funcionan de manera similar a los anteriores, es decir, destacan la unidad de búsqueda así como sus unidades precedentes y siguientes a fin de proporcionar su contexto. En Terminografía bilingüe, estas aplicaciones son muy útiles, ya que ofrecen la posibilidad de observar a la vez los resultados de concordancias en las dos lenguas de trabajo, sin perder las funciones propias de los analizadores monolingües y, de esta manera, detectar posibles equivalentes. Ejemplos de programas de concordancias bilingües son *Multiconcord*⁹² y *Paraconc*.⁹³

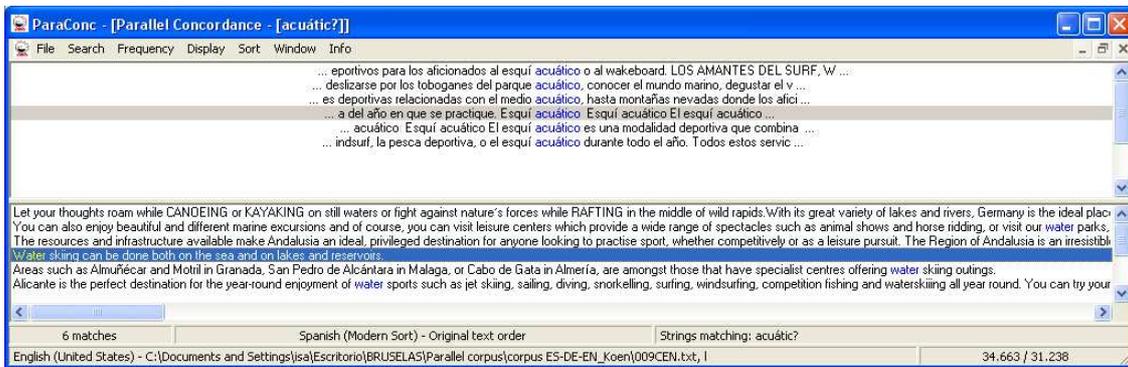


Figura 9. Ejemplo de líneas de concordancia bilingüe con el programa *ParaConc*.

A parte de los analizadores de concordancias, que únicamente proporcionan el contexto de la unidad que se ha buscado, existen los analizadores de colocaciones (*collocation extractors*). Estas aplicaciones funcionan de forma similar a las anteriores, pero se basan en otras reglas de cálculo y su objetivo es proporcionar colocaciones, es decir, unidades fraseológicas que presentan una relación de asociación elevada y, por tanto, tienden a aparecer juntas (cf. Corpas Pastor, 1997). Las reglas de cálculo más utilizadas para la obtención de colocaciones se basan en el cálculo de asociación de las unidades y son principalmente dos: por un lado, el índice de información mutua (*Mutual Information Score, MI-Score*), que calcula la fuerza de asociación de una unidad con

⁹² <http://artsweb.bham.ac.uk/pKing/multiconc/l_text.htm>

⁹³ <<http://www.athel.com/para.html>>

otra comparando la probabilidad de que esas unidades aparezcan juntas bien porque forman una unidad fraseológica o bien porque su aparición sea solo casual (McEnery y Wilson, 2004: 86). Cuanto mayor sea el resultado de ese cálculo, más fuerte será la asociación entre esas dos unidades; y al contrario, si las unidades aparecen juntas de forma casual, el resultado será muy bajo. Y, por otro lado, el índice estadístico *t-score*, que mide el grado de certeza con que se puede decir que existe una asociación de unidades y no, como el anterior, la fuerza de la asociación de dichas unidades. Las palabras que poseen un índice de frecuencia más alto en el corpus (preposiciones, pronombres o artículos) ofrecerán también un índice de colocación *t-score* mayor, de forma que los índices significativos de esta medida suelen señalar colocaciones muy fuertes o asociaciones entre palabras léxicas y gramaticales (por ejemplo, preposiciones con verbos o con adjetivos), mientras que el índice de información mutua suele indicar asociaciones que son estadísticamente significativas (aunque la frecuencia de aparición de los elementos de la colocación en el corpus sea muy baja), por lo que suele señalar asociaciones semánticas entre palabras o elementos de una unidad fraseológica (Pérez Hernández, 2002). Un ejemplo de estos analizadores de colocaciones es el programa *Collocate*.⁹⁴

Las siguientes herramientas de gestión y análisis de corpus de las que vamos a hablar a continuación son los extractores de relaciones semánticas, que se emplean a menudo para extraer así como para elaborar definiciones terminológicas.

La extracción de relaciones semánticas entre términos se basa en la premisa de que dichas relaciones se expresan en el texto y que, por tanto, pueden ser extraídas mediante técnicas de Procesamiento del Lenguaje Natural (PLN). El primer paso para extraer las relaciones de un dominio determinado requiere de la identificación de los conceptos propios de la materia objeto de estudio mediante la consulta a expertos, glosarios especializados y la extracción de términos (Orăsan et ál., 2007: 29). Una vez que se han identificado dichos conceptos, es posible llevar a cabo la extracción de relaciones en los corpus. De los métodos empleados en este proceso, el más relevante se basa en la extracción de relaciones a partir de patrones de conocimiento (*knowledge*

⁹⁴ <<http://www.athel.com/colloc.html>>

patterns),⁹⁵ que son estructuras similares a “X está formado por Y”, “X produce Y”, “X es un Y”, “X es similar a Y”, etc. El segundo paso consiste en emplear los patrones de conocimiento seleccionados en el primer paso para identificar las relaciones entre los términos del corpus. Así pues, cada vez que se encuentre un patrón en el corpus, se considerará que existe una relación entre sus componentes, siempre y cuando estos sean términos. Es importante tener en cuenta que las técnicas de PLN actuales no son capaces aún de llevar a cabo identificaciones muy precisas de las relaciones entre términos y, por tanto, el terminógrafo deberá realizar una revisión de los resultados obtenidos tras la extracción.

La extracción de relaciones semánticas en Terminografía tiene gran relevancia, puesto que permite al terminógrafo analizar el funcionamiento de las unidades, categorizarlas y establecer la terminología del dominio de especialidad objeto de estudio. Un ejemplo de este tipo de herramientas es SeRT (*Semantic Relations in Text*) (Barrière, 2002). Asimismo, la extracción de relaciones se utiliza a menudo para establecer definiciones terminológicas. Estas definiciones derivadas de la extracción de relaciones están basadas sobre todo en la relación de hiperonimia, es decir, IS-A, aunque también pueden elaborarse mediante relaciones de metonimia (parte-todo), relaciones causales, instrumentales, etc. (García de Quesada, 2001; Sierra et ál., 2008).⁹⁶

En la actualidad, existen diferentes paquetes informáticos disponibles en el mercado, tanto de libre acceso (por ejemplo *TextWorld*, desarrollado por Alan Reed,⁹⁷

⁹⁵ Los patrones de conocimiento se definen como «a linguistic pattern, which is repetitive and expresses domain knowledge about the terms» (Meyer, 2001: 290). Marshman et ál. (2002) distinguen tres tipos de patrones lingüísticos, a saber: los patrones léxicos, los patrones gramaticales y los patrones paralingüísticos. Los primeros son los más visibles y consisten en palabras o grupos de palabras que indican las relaciones conceptuales; los segundos implican combinaciones de estructuras gramaticales que ofrecen relaciones semánticas entre conceptos aunque de un modo más limitado que los patrones léxicos; mientras que los terceros hacen referencia a los elementos textuales como, por ejemplo, comas, paréntesis, puntos, etc. No obstante, en la extracción de relaciones semánticas, los patrones más usados son los léxicos.

⁹⁶ En este tema profundizaremos en el apartado 2.6.1.3.3., cuando presentemos la etapa de extracción de relaciones semánticas dentro de las fases de nuestra propuesta metodológica para el trabajo onoterminográfico. Asimismo, trataremos explícitamente estas relaciones en el apartado 4.2.1.3., cuando describamos las relaciones encontradas en nuestro corpus formado por textos promocionales del turismo de aventura.

⁹⁷ <<http://www.textworld.com/scp/>>

System Quirk, de la Universidad de Surrey⁹⁸ o *TACT*, del departamento de Computing in the Humanities and Social Sciences de la Universidad de Toronto⁹⁹) como comerciales (*WordSmith Tools*, desarrollado por Michael Scott para Oxford University Press),¹⁰⁰ que integran la mayoría de las herramientas que hemos comentado anteriormente, con la intención de proporcionar una mayor facilidad de uso a los usuarios en general, y concretamente a los terminógrafos. Sin embargo, ningún programa o paquete informático actual está formado por todas las herramientas necesarias para llevar a cabo la labor terminográfica. Así pues, la sistematización terminográfica se lleva a cabo mediante la utilización de varias aplicaciones simultáneamente y buscando métodos de cooperación entre ellas.

El paquete informático más utilizado en Terminografía, así como en otros ámbitos lingüísticos (traducción, lexicografía, enseñanza de lenguas, etc.) es el programa de gestión de corpus *WordSmith Tools* (Vargas Sierra, 2006b; López Rodríguez et ál., 2006; Marshman, 2008), que contiene un conjunto de tres herramientas fundamentales: *Wordlist*, que permite crear una lista con todas las palabras o grupos de palabras de un texto (o un corpus textual), ordenadas por orden alfabético o por frecuencias; *Concord*, que posibilita observar las líneas de concordancias de cualquier unidad que se especifique; y *KeyWords*, con la que se pueden identificar las palabras clave de un corpus especializado comparándolo con un corpus general. Como se observa, este programa integra dos de las herramientas presentadas anteriormente, a saber: un extractor de terminología basado en reglas estadísticas y un analizador de concordancias.

Hoy día, además de los paquetes informáticos como *WordSmith Tools*, también se cuenta con algunos sistemas integrados de gestión terminológica, como por ejemplo el gestor terminológico TERMINUS, desarrollado por el Institut Universitari de Lingüística Aplicada de la Universitat Pompeu Fabra.¹⁰¹ Este programa es un gestor integral para el trabajo terminográfico que se hizo público a mediados del año 2009 y

⁹⁸ <<http://www.computing.surrey.ac.uk/SystemQ/>>

⁹⁹ <<http://www.chass.utoronto.ca/cch/tact.html>>

¹⁰⁰ <<http://www.lexically.net/wordsmith/>>

¹⁰¹ <<http://igraine.upf.edu/Terminus2/index.html>>

que incluye aplicaciones informáticas para apoyar el trabajo terminográfico estándar,¹⁰² a saber: a) la compilación de un corpus, b) la creación de mapas conceptuales, c) el análisis y gestión de dicho corpus (con analizador de concordancias, de extracción terminológica y de colocaciones), d) la compilación de un glosario y e) su edición final.¹⁰³ Con este programa, los terminógrafos trabajan en un entorno donde, como decíamos, tienen a su disposición las herramientas necesarias para llevar a cabo su labor terminográfica. No obstante, tiene algunos inconvenientes, tales como la elevada cantidad de ruido obtenido en la extracción de términos o la carencia de extractores de relaciones semánticas, que se pueden solucionar mediante el uso de otras herramientas externas más potentes. No obstante, en nuestra opinión, se trata de un programa de gran utilidad para el trabajo terminográfico estándar, puesto que cubre las necesidades básicas de los terminógrafos y ofrece flexibilidad en su uso.

Como hemos visto, la variedad actual de herramientas y programas informáticos es muy variada y cada uno presenta una serie de características y funcionalidades que pueden ser más o menos parecidas entre sí. Por ese motivo, el terminógrafo, o el equipo de trabajo, deberá realizar la elección de una u otra herramienta, o de un u otro paquete informático según las necesidades de cada proyecto, los objetivos planteados y, por supuesto, la disponibilidad económica y de tiempo de cada uno.

1.4.2.5. Ventajas e inconvenientes del empleo de corpus en Terminografía

The value of a corpus or “the corpus” should, however, not be over- or underestimated –it cannot replace the lexicographer, nor should it be regarded as inferior to the knowledge of the lexicographer at any respect (Prinsloo, 2009: 181).

Para finalizar con el apartado de los corpus especializados, trataremos las ventajas y los inconvenientes (o limitaciones) que pueden presentar a la hora de llevar a

¹⁰² Cuando hablamos de «trabajo estándar», nos referimos al trabajo que se lleva a cabo en Terminografía de forma usual, sin incorporar la gestión y creación de ontologías como herramientas de apoyo y fuentes de información, ya que en su caso hablaríamos de Ontoterminografía.

¹⁰³ Las diferentes fases del trabajo terminográfico se tratarán con mayor detenimiento en el capítulo 2.

cabo un proyecto terminográfico, a fin de ofrecer un compendio de las ventajas que hemos ido presentando durante toda esta parte del capítulo así como indicar algunos de los inconvenientes que se le podrían encontrar a su uso.

Como hemos visto hasta ahora, la gran mayoría de los autores y terminógrafos en la actualidad defienden el uso de los corpus por todas las ventajas que presentan:

En primer lugar, el corpus puede servir al mismo tiempo de fuente, referencia y documentación (Bourigault, 2000).¹⁰⁴ Como *fuentes*, el corpus actúa como depositario de los conocimientos de un ámbito de especialidad y, a través de su análisis y estudio, permite la localización de los términos y equivalentes, las relaciones entre los conceptos, u otras informaciones terminológicas necesarias (contexto, usos típicos, combinatoria, derivados, etc.). Como *referencia*, el corpus supone trabajar sobre datos auténticos y realmente utilizados por los especialistas, lo que se denomina datos *in vivo*; y, por ello, se trata de un material al que acudir para validar o descartar hipótesis, comprobar información extraída, verificar datos anteriores, etc. Por último, el corpus constituye la *documentación* del recurso terminológico, ya que autentifica los términos, las definiciones u otro tipo de datos que incluye, así como da cuenta del origen de la información, de la fecha de creación, etc.

En segundo lugar, y muy relacionado con lo anteriormente expuesto, el corpus es capaz de mostrar el dinamismo característico de la lengua, ya sea esta general o especializada. Mediante la actualización constante de sus textos, se alcanza a observar el sistema dinámico y cambiante de relaciones entre conceptos, y de estos con sus contextos y con el lenguaje, lo que permite observar la riqueza de los lenguajes de especialidad, así como su actualización y creaciones constantes.

En tercer lugar, el corpus permite la gestión y análisis de datos de forma rápida, precisa y consistente a un bajo coste, gracias a la cantidad de aplicaciones informáticas que existen en la actualidad, las cuales se encuentran en continuo desarrollo. Asimismo, estas herramientas facilitan el acceso y la manipulación de grandes cantidades de datos, así como el procesamiento (semi)automático de los textos. Sin este tipo de aplicaciones y sin el uso del corpus electrónico, los terminógrafos se verían obligados a realizar todas

¹⁰⁴ Citado en Gómez González-Jover (2006: 252).

las tareas de forma manual, lo cual daría lugar a una gran pérdida de tiempo y a unos resultados con una menor fiabilidad.

En cuarto lugar, otro aspecto positivo de los corpus es su reutilización en múltiples usos y aplicaciones, siempre y cuando no esté etiquetado, las etiquetas se puedan eliminar del corpus o las etiquetas existentes sirvan para el siguiente proyecto.

En quinto lugar, son un recurso muy valioso para realizar estudios contrastivos entre diferentes lenguas y lenguajes de especialidad, así como para llevar a cabo estudios diacrónicos y de variación al facilitar la comparación, la consulta y la extracción de información.

Todas estas ventajas están directamente relacionadas con la fiabilidad y credibilidad que aporta el uso del corpus especializado para fines terminográficos, en contraposición al método de la introspección, ya que sus resultados se basan en métodos empíricos y estadísticos. El corpus muestra exactamente cómo es el lenguaje de especialidad objeto de estudio de forma objetiva y directa, al contrario que la intuición, que se trata de un método subjetivo dependiente del conocimiento del terminógrafo o del especialista consultado y de los sesgos que estos pudieran hacer. Por ello, el empleo del corpus está tan extendido en la comunidad terminográfica.

No obstante, no podemos olvidarnos de los inconvenientes y limitaciones que pueden presentar los corpus en algunas ocasiones.

La principal limitación de un corpus es su representatividad. Todavía no se han podido acordar la cantidad de datos necesarios para que un corpus sea representativo y, aunque lo fuera, todos los intentos por establecer generalizaciones extraídas de los corpus serían extrapolaciones. Como indica Hunston (2002: 23), una afirmación realizada de datos contenidos en un corpus es una afirmación sobre ese corpus, y no sobre el lenguaje de especialidad del cual el corpus es una muestra. Por este motivo, la información y los resultados obtenidos de un corpus deben tratarse con cautela y no como datos definitivos y generales.

Otra limitación relevante es la disponibilidad de corpus para las tareas terminográficas. La gran mayoría de terminógrafos que llevan a cabo un proyecto deben compilar su propio corpus especializado para poder analizar los datos y extraer la

información necesaria, puesto que existen muy pocos corpus especializados que se puedan utilizar. Para esto, es necesario invertir mucho tiempo de evaluación, selección y compilación de los textos, sumado al tiempo que se requiere para la gestión y análisis posterior. Este motivo hizo que el corpus se introdujera tan tarde en la Terminografía. No obstante, este inconveniente se está intentando solucionar de diferentes maneras, por ejemplo con la búsqueda automática en Internet para la compilación de un corpus (*web as corpus*), como veníamos más arriba, o con sistemas informáticos de compilación de corpus semiautomáticos basados en la densidad de aparición de patrones semánticos (cf. Barrière, 2006).

Por último, citaremos una limitación vinculada a la buena elección de los textos para compilar un corpus. La necesidad de seleccionar textos adecuados, que cumplan los criterios establecidos previamente, que sean representativos del dominio de especialidad y que mantengan un equilibrio interno es una cuestión a simple vista sencilla; sin embargo, a menudo estos criterios se vuelven muy exigentes y no siempre es posible satisfacerlos o, incluso, se incumplen. En estos casos, la calidad del corpus no será adecuada y los resultados, por tanto, no serán relevantes ni fiables. Para evitar este inconveniente, es imprescindible que el terminógrafo establezca todos los criterios previamente al comienzo del proyecto y que los siga durante todo el proceso de compilación y que, además, intente alcanzar un corpus equilibrado (ya sea de géneros textuales, subdominios, lenguas, etc., dependiendo del objetivo del corpus). De lo contrario, se podrían producir sesgos, interpretaciones erróneas, etc. que provocarían unos análisis equivocados.

Para finalizar, nos parece oportuno destacar que, desde nuestro punto de vista, el corpus se ha convertido en una herramienta imprescindible para el trabajo terminográfico y que, a pesar de las limitaciones que presenta, es un recurso de gran valor para la labor del terminógrafo, que, gracias a él, ha visto reducido su trabajo a la par que facilitado por las herramientas de gestión y análisis (semi)automáticas.

1.4.3. LA TERMINOGRAFÍA BASADA EN LA REPRESENTACIÓN DE CONOCIMIENTO

La Terminografía basada en la representación de conocimiento no se entiende como una rama independiente de la Terminografía basada en corpus, es decir, no son dos líneas de investigación distintas entre sí, sino que la primera emplea la metodología de la Lingüística de Corpus¹⁰⁵ y, como en el resto de líneas de investigación de la Terminografía actual, el corpus viene a considerarse como una herramienta indispensable para llevar a cabo sus análisis y sus proyectos terminográficos. Así pues, como indica Faber (2004: 386), la combinación de la terminología descriptiva basada en el conocimiento (Meyer y Mackintosh, 1996) y la Lingüística de Corpus ofrece una base idónea para extraer, de forma semiautomática, el conocimiento especializado así como para gestionar dicho conocimiento y aplicarlo al desarrollo de recursos terminográficos.

Como veíamos con anterioridad, los estudios basados en corpus hacen explícito el contexto del término y ponen de manifiesto la información sobre su significado y sobre su uso (Pearson, 1998: 191), todo lo cual los terminógrafos extraen mediante el empleo de herramientas como extractores de términos, analizadores de concordancias y colocaciones, etc. En la Terminografía basada en el conocimiento, el corpus también permite extraer información relevante para la creación de estructuras conceptuales y ontologías, como son los términos asociados, elementos definatorios, relaciones semánticas, a la par que el contexto hace posible la verificación de la información conceptual extraída de obras de referencia y la elaboración de definiciones de forma sistemática y homogénea.

Esta posibilidad que ofrece el corpus de encontrar la estructura conceptual de los dominios de especialidad reflejados en la lengua en uso ha permitido que la investigación en representación semántica adquiera un protagonismo del que hasta ahora no había gozado, en concreto en el ámbito de la terminografía, pero también en traducción y en otras ramas de la lingüística como son la semántica y la sintaxis. Este contexto ha motivado el interés creciente por la elaboración de estructuras conceptuales

¹⁰⁵ Ya vimos en el capítulo anterior cómo la Lingüística de Corpus no puede considerarse una disciplina tal cual, sino más bien una metodología de trabajo.

y de ontologías en la última década, puesto que se han apreciado las posibilidades que ofrecen al servir como puente entre la representación mental del conocimiento y el lenguaje (Faber, 2009: 290) y como estructura subyacente en cualquier base de datos. De acuerdo con Gillam et ál. (2007: 51):

[...] The key role of an ontology is to identify areas of knowledge, associate these areas of knowledge with other areas, and demarcate these areas with key terms. The practical import of an ontology is to create systems for storing and retrieving fragments of knowledge.

No obstante, es necesario tener en cuenta que la estructuración conceptual no es algo nuevo en Terminografía, puesto que uno de los principales objetivos del terminólogo consiste en establecer sistemas conceptuales entre los conceptos de un campo de especialidad, además de elaborar y definir las relaciones conceptuales que existen entre esos conceptos, crear taxonomías y especificar cuáles son las unidades léxicas que se emplean en las diversas lenguas objeto de su estudio para hacer referencia a dichos conceptos. De hecho, encontramos referencias a la necesidad de desarrollar esta actividad en cualquier manual de Terminología (Sager, 1990; Cabré Castellví, 1993; Wright y Budin, 1997). Como nos recuerda Sager en la siguiente cita:

A theory of terminology is usually considered as having three basic tasks: it has to account for sets of concepts as discrete entities of the knowledge structure; it has to account for sets of interrelated linguistic entities which are somehow associated with concepts grouped and structured according to cognitive principles; it has, lastly, to establish the link between concepts and terms, which is traditionally done by definitions (Sager, 1990: 21).

De esto, se ha producido una variación en el trabajo terminográfico con respecto a la creación de estructuras conceptuales basada en la concepción que se tiene actualmente del concepto y de las relaciones entre los conceptos en la Terminología moderna; la utilización del corpus textual como herramienta básica de trabajo; las aplicaciones informáticas que permiten un trabajo sistemático, homogéneo y más automatizado, entre otras cosas, todo lo cual ha permitido una evolución de la Terminografía con relación a la representación del conocimiento especializado. Asimismo, la colaboración de la Terminografía con otras disciplinas con las que comparte ciertas tareas, como son Inteligencia Artificial y la Ingeniería del Conocimiento, ha permitido disponer de

sistemas de representación conceptual formalizados, explícitos, estandarizados y consensuados, lo que ha facilitado parte de la labor del terminólogo (Moreno Ortiz, 2008) y ha dado lugar a la creación de ontologías en Terminografía.

Con respecto a las similitudes que se observan entre la Terminografía y la Ingeniería del Conocimiento y la Inteligencia Artificial, se pueden distinguir principalmente las dos siguientes (Vargas Sierra, 2007: 52):

En primer lugar, los dominios objetos de estudio se organizan conceptualmente y se lleva a cabo la extracción de los términos y de sus relaciones con otros términos a partir de un corpus textual.

En segundo lugar, tanto en la elaboración de terminologías (en Terminografía) como de ontologías (en Ingeniería del Conocimiento o Inteligencia Artificial)¹⁰⁶ se combinan dos aproximaciones para capturar y representar el conocimiento de un ámbito especializado: el deductivo (*top-down*) y el inductivo (*bottom-up*). En el primer caso, se establece la organización conceptual del ámbito en cuestión con la colaboración de expertos del dominio mediante la identificación inicial de los conceptos básicos, que constituyen el punto de partida en la identificación de otros elementos de la ontología. En el segundo caso, se parte de los textos para adquirir conocimiento sobre el dominio y detectar las unidades terminológicas en contexto, las cuales se ubicarán dentro de la estructura conceptual del dominio analizado.

Estas similitudes han dado lugar a esta nueva línea de investigación en Terminografía, la Terminografía basada en la representación conceptual, que actualmente, debido a su reciente aparición, no cuenta aún con una denominación única y compartida por todos los expertos, sino que recibe diferentes denominaciones, a saber: *termontografía* (*termontography*, Kerremans et ál., 2003), *ontoterminología* (*ontoterminologie*, Roche, 2007) o simplemente terminografía basada en ontologías o en la representación de conocimiento, entre otros. Todos los términos hacen referencia a la misma idea: la incorporación de las ontologías a la Terminografía, y las sitúan en el punto central de cualquier trabajo terminográfico, pero, sin embargo, la terminología

¹⁰⁶ El término *ontología* se emplea en el ámbito de la Ingeniería del Conocimiento para hacer referencia a un conjunto de conceptos organizados jerárquicamente, representados en algún sistema informático cuya utilidad es la de servir de soporte en diversas aplicaciones que requieren de conocimiento específico sobre la materia que la ontología representa (Moreno Ortiz, 2004: 31).

adoptada no nos parece adecuada: la primera denominación, el término *termontography* se encuentra inmerso en la Teoría Sociocognitiva de la Terminología (Temmerman, 2000),¹⁰⁷ por lo que los términos que se emplean y la perspectiva de trabajo está restringida por completo a dicha teoría; el segundo término, la *ontoterminología* habla de *terminología*, es decir, de la rama teórica de la disciplina, lo que deja fuera la terminografía, sin realizar ninguna distinción entre una y otra y, además, su metodología se aproxima a un trabajo terminográfico más tradicional, a saber: onomasiológico, prescriptivo, sin uso de corpus como base, entre otros puntos, lo que se aleja de la Terminología moderna que defendemos; y, por último, el resto de términos son simplemente descriptivos y podrían formar múltiples variables. En nuestro caso, defendemos que es necesario proporcionarle una denominación concreta debido al creciente interés que está despertando en los últimos años y proponemos la acuñación del término *ontoterminografía*, por no estar de acuerdo con las denominaciones anteriores, que consideramos excluyentes. En este sentido, la ontoterminografía se definirá como la terminografía que cuenta con la incorporación de las ontologías en el trabajo terminográfico en un marco de la Terminología moderna, incluyendo los principios que comparten la Teoría Comunicativa de la Terminología y la Teoría Sociocognitiva de la Terminología, como son la necesidad de prestar atención a la dinámica de los términos y a su polisemia, a la situación comunicativa donde aparecen las unidades, al rechazo de la univocidad absoluta, al empleo de los corpus textuales al trabajo terminográfico, la categorización de los dominios, etc.

En definitiva, la Ontoterminografía establece un nuevo contexto de trabajo en la elaboración de recursos terminológicos que sitúa la creación de ontologías como paso previo a la construcción de cualquier base de datos terminológica. Actualmente, existen todavía escasos estudios en los que se ha llevado a cabo este tipo de trabajo terminográfico, especialmente por la falta de herramientas disponibles para abarcar todas las fases de elaboración de un recurso terminográfico dentro del marco ontoterminográfico. No obstante, existen buenos ejemplos de aplicación de esta rama

¹⁰⁷ De hecho, sus creadores la definen de la siguiente manera: «Termontography is a multidisciplinary approach in which theories and methods for multilingual terminological analysis of the sociocognitive approach (Temmerman 2000) are combined with methods and guidelines for ontological analysis» (Temmerman y Kerremans, 2003).

terminográfica en la actualidad, de los cuales comentaremos algunos al final de este capítulo.

1.4.3.1. La representación de conocimiento especializado en la TGT y otros modelos de representación

En Terminografía, uno de los primeros pasos imprescindibles antes de abordar cualquier trabajo consiste en la categorización de los conceptos propios de un campo de especialidad, entendiendo esta como el proceso mental a través del cual una mente humana agrupa los conceptos en distintas categorías según las características y propiedades de dichos conceptos. De esta manera, antes de realizar la representación conceptual de un dominio, es necesario haber llevado a cabo la categorización previa, a fin de haber creado los diferentes apartados donde se podrán ir añadiendo los conceptos propios del campo de especialidad en cuestión y donde se podrán establecer los prototipos de acuerdo con la Teoría de Prototipos, de Rosh.

Una vez establecida esta categorización del campo de especialidad en cuestión, se pasará a la representación, es decir, a la explicitación de la información que contiene dicho campo de especialidad. Como se recoge en la siguiente cita, esta representación puede ser tan detallada o tan básica como exija el trabajo en curso o la finalidad última que se persiga con el trabajo terminográfico:

Representation is about making things explicit, about resolving ambiguities and above all, particularly in the context of artificial intelligence, about creating a surrogate of a class of things that exist in the real world on a computer system. (Proyecto POINTER, 1998: 97)

A pesar de que el objetivo de la representación conceptual y de la categorización es compartido por todos los expertos, a saber: ilustrar la posición de los conceptos dentro de un sistema conceptual completo así como la relación de dicho concepto con respecto a los demás conceptos que forman la representación, no existe hoy en día consenso en cuanto a la metodología utilizada para esta representación del conocimiento, que además ha ido cambiando con el paso del tiempo. Esta evolución en la representación conceptual se ha debido principalmente a dos razones: por un lado, a la evolución que

ha sufrido la Terminología desde que se considera disciplina, es decir, el cambio de la Terminología tradicional a la Terminología moderna; y por otro, a la disponibilidad de herramientas para el desarrollo de representaciones conceptuales más complejas.

Con respecto a la primera razón, ya veíamos anteriormente que los avances que se han producido en la Terminología desde la TGT, tanto con relación a los términos como a la concepción de los lenguajes de especialidad, las relaciones de sinonimia y polisemia, el carácter prescriptivo, etc. Recordemos que, según la Terminología tradicional, todo trabajo terminográfico presenta un carácter normalizador y debe otorgar prioridad al concepto sobre el término con un enfoque onomasiológico, es decir, comienza desde el concepto y finaliza en el término, con el que establece una relación preferentemente unívoca. Con respecto a la estructura conceptual, esta teoría defiende la posición estática, uniforme y universal de los conceptos en dicha estructura, independientemente de la lengua en la que se trabaje, y el establecimiento de relaciones lógicas u ontológicas entre todos los conceptos de un mismo dominio de especialidad, los cuales se diferencian entre sí mediante características intrínsecas y extrínsecas propias. Por el contrario, en la Terminología moderna se postula que los conceptos no siempre están bien delimitados ni la relación 1:1 entre concepto y término es siempre posible (Temmerman, 2000: 223). Asimismo, se defiende la existencia de otros tipos de relaciones (no jerárquicas) además de las relaciones jerárquicas (lógicas y ontológicas) que han de ser codificadas si se quiere dar cuenta de la arquitectura conceptual global en cada dominio de especialidad (García de Quesada, 2001: 249). Por último, se sostiene que todo dominio de especialidad está en continuo cambio y, por tanto, los conceptos, términos y relaciones, ya sean estas jerárquicas o no jerárquicas, que se establecen entre ellos cambian de la misma manera. Todo esto ha dado lugar a un cambio de visión acerca del concepto y de la estructura conceptual de un dominio de especialidad que ha derivado en las representaciones ontológicas que se realizan en la actualidad, donde se atiende especialmente a las diferentes relaciones (jerárquicas y no jerárquicas) que mantiene un concepto con otros conceptos del mismo sistema, al dinamismo y a la multidimensionalidad de dichos conceptos dentro del dominio de especialidad a la par que a la parte pragmática de la comunicación.

La segunda razón que ha dado lugar a las nuevas representaciones conceptuales tiene que ver con las herramientas que se han ido desarrollando en la Lingüística Computacional y en la Ingeniería del Conocimiento y que han facilitado la representación de sistemas conceptuales más complejos y completos y ha permitido una mayor profundización en la representación de relaciones conceptuales.¹⁰⁸

Antes de pasar a la descripción de las diferentes representaciones conceptuales en Terminología, es importante tener en cuenta los sistemas de clasificación empleados previamente y que también tenían como objetivo la organización del conocimiento. Estos sistemas principalmente pertenecen al ámbito de la Documentación y la Biblioteconomía y se pueden considerar como los antecedentes de las representaciones conceptuales realizadas dentro de la Terminografía. Los principales sistemas de clasificación que se pueden considerar en este contexto son las taxonomías y los tesauros.

Una taxonomía es una colección de términos de un vocabulario controlado¹⁰⁹ organizada jerárquicamente, es decir, cada término en una taxonomía tiene una o más relaciones hiperonímicas con otros términos incluidos en ella, donde los términos que ocupan los niveles más altos son los términos genéricos. En este sentido, una taxonomía podría confundirse con una ontología; sin embargo, al contrario que las ontologías, las taxonomías están formadas por términos y no por conceptos, y únicamente incluyen relaciones genéricas, es decir, de hiperónimos e hipónimos.

Por su parte, un tesoro puede definirse de dos formas diferentes, dependiendo de si se atiende a su estructura o su función (Cordón García et ál., 2001: 97). Por su estructura, un tesoro es un vocabulario controlado y dinámico de términos que

¹⁰⁸ En el apartado 1.4.3.4., presentaremos los principales tipos de herramientas que existen en la actualidad relacionadas con la elaboración de sistemas conceptuales basados en ontologías.

¹⁰⁹ Un vocabulario controlado consiste en una lista de términos que ha sido elaborada explícitamente por una autoridad, cuyo objetivo consiste en proporcionar definiciones no redundantes ni ambiguas de las diferentes unidades que lo forman. Básicamente, este tipo de recurso tiene tres funciones fundamentales, a saber: a) tiende a reducir las ambigüedades semánticas en un dominio de especialidad, mediante la diferenciación de los distintos significados de los homógrafos; b) intenta mejorar la consistencia en la representación de la materia mediante el control de los sinónimos y cuasisinónimos; y c) pretende facilitar la realización de búsquedas amplias de información estableciendo una estructura que una los términos relacionados semánticamente.

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. En definitiva, tanto los tesauros como las ontologías son herramientas que sirven para estructurar conceptualmente determinados ámbitos del conocimiento por medio de vocabularios controlados; sin embargo, la diferencia entre los tesauros y las ontologías radica en la complejidad, puesto que estas últimas introducen un mayor nivel de profundización semántica y proporcionan una descripción lógica y formal que puede ser interpretada tanto por personas como por máquinas, mientras que los tesauros solo pueden ser interpretados por humanos (Lamarca Lapuente, 2009).

Una vez que hemos visto brevemente los sistemas de organización del conocimiento más extendidos y antiguos fuera de la Terminología, como son las taxonomías y los tesauros, vamos a pasar a tratar la evolución que ha sufrido la representación de conocimiento desde la TGT hasta llegar a la representación basada en ontologías, pasando por algunos sistemas de representación utilizados en la Terminografía que no pertenecen a la perspectiva tradicional pero que tampoco utilizan ontologías como base.

Partiendo de los supuestos de la TGT, es evidente que las estructuras conceptuales que se realicen en el marco de esta teoría discrepen con los que se puedan realizar siguiendo otras premisas teóricas, ya sea la TCT, la TST u otra, ya que, como decíamos antes, la TGT propone que los sistemas conceptuales sean invariables y universales, dejando de lado la parte comunicativa y pragmática de la lengua. No obstante, todas las teorías terminológicas comparten un mismo objetivo: la generación de sistemas que permitan ordenar los conceptos en función de las relaciones que se establecen entre ellos dentro de un dominio específico (Wright y Budin, 1997: 337).

La TGT propugna la necesidad de estudiar y crear sistemas de conceptos, los cuales vienen definidos como un conjunto ordenado de conceptos que tienen relaciones conceptuales jerárquicas entre sí, ya sean de tipo lógico u ontológico. En el caso de que las relaciones no pertenezcan a ninguno de esos dos tipos, es decir, sean relaciones no jerárquicas, Wüster (1998: 57) habla de clasificaciones temáticas y no de sistemas de

conceptos, ya que, según el autor, las relaciones no jerárquicas son relaciones que solo aparecen ocasionalmente en un documento y que no afectan a la esencia de un concepto y que, por tanto, no se les debe prestar atención en Terminografía. Siguiendo este marco teórico, la norma ISO 704 (2000: 12) establece las funciones que deben cumplir los sistemas de conceptos:

- Modelar las estructuras conceptuales a partir de los conocimientos propios de un dominio de especialidad;
- Clarificar las relaciones entre conceptos;
- Servir de base para una Terminología uniforme y estandarizada;
- Facilitar el análisis comparativo de los conceptos y los términos en diferentes lenguas;
- Facilitar la elaboración de definiciones.

Esta misma norma identifica diferentes sistemas de conceptos según el tipo de relaciones conceptuales que se elijan para su estructuración. Así pues, establece tres tipos de sistemas de conceptos: sistemas genéricos, sistemas partitivos y sistemas mixtos, para los cuales ofrece una serie de recomendaciones para representar gráficamente las distintas relaciones conceptuales.

En primer lugar, los sistemas genéricos de conceptos corresponden con unos sistemas en los que todos los conceptos se relacionan con otros conceptos de forma vertical como conceptos específicos y genéricos, es decir, únicamente mediante relaciones lógicas («Y es un X»). La recomendación de esta norma a la hora de representar gráficamente este tipo de sistema de conceptos es una representación en forma de árbol, que refleje las series genéricas verticales (Figura 10), o bien una lista de conceptos con una notación alfanumérica separada por puntos para indicar los diferentes niveles (Figura 11).

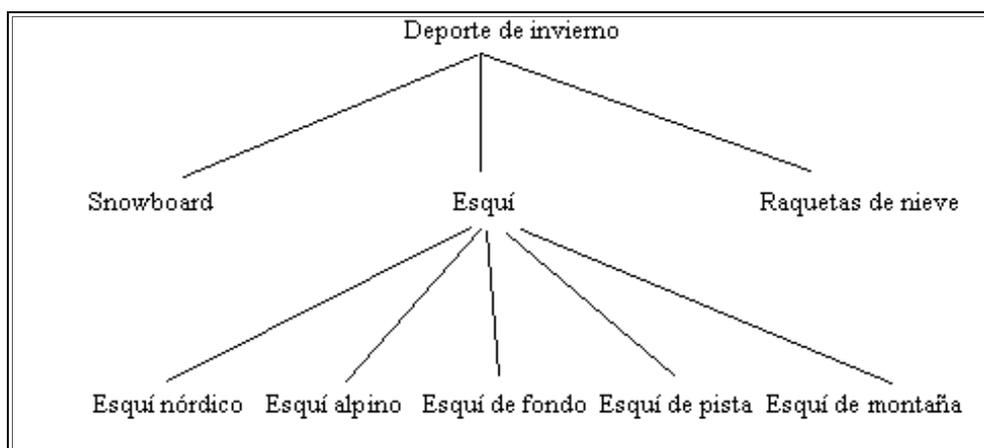


Figura 10. Representación de sistema genérico de conceptos en forma de árbol con algunos deportes de invierno.

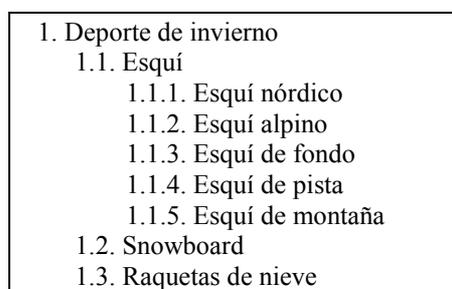


Figura 11. Representación alfanumérica de sistema genérico de conceptos con algunos deportes de invierno.

En segundo lugar, el sistema partitivo de conceptos se trata de un sistema en el que todos los conceptos se relacionan entre sí de forma vertical mediante la relación parte-todo. Estos sistemas se suelen representar gráficamente mediante una estructura de corchetes (Figura 12), y su notación numérica es semejante a las de las relaciones genéricas.

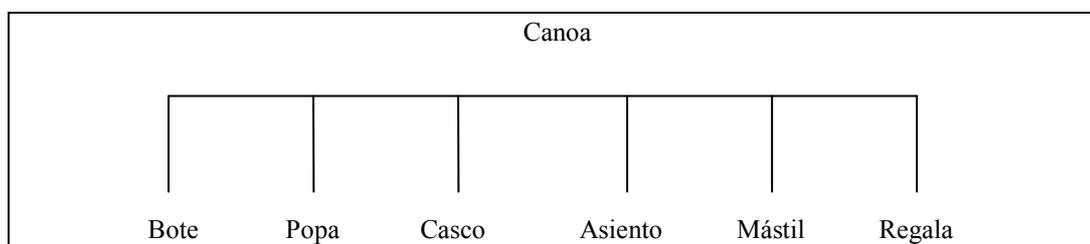


Figura 12. Representación de sistema partitivo de conceptos con corchetes con las partes de una canoa.

En tercer lugar, tenemos los sistemas de conceptos mixtos, los cuales combinan los dos sistemas de conceptos anteriores, es decir, las relaciones genéricas y partitivas y sus correspondientes representaciones gráficas (Figura 13).

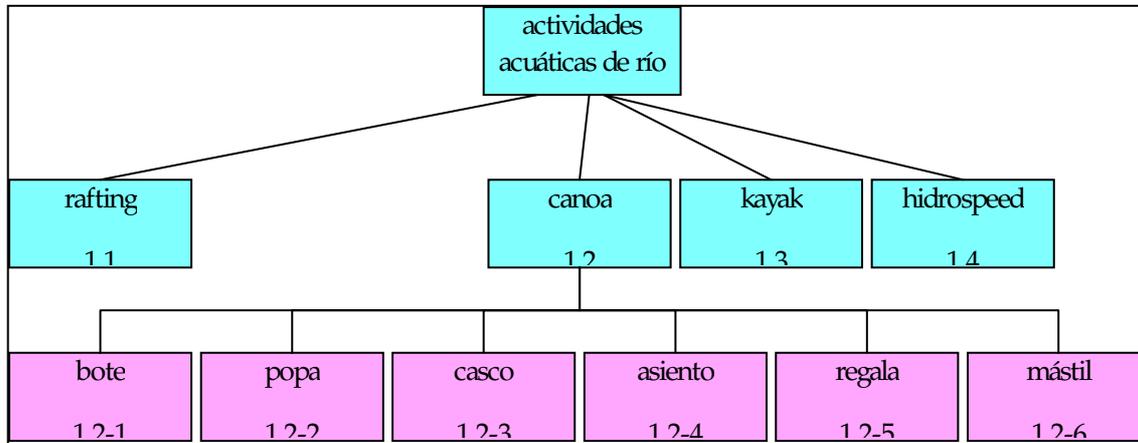


Figura 13. Representación de sistema de conceptos mixto con algunas de las actividades acuáticas de río (sistema de conceptos genérico) y las partes de una canoa (sistema de conceptos partitivo) realizada con el programa *GeneSis*.¹¹⁰

Como se puede observar, el sistema de conceptos está dividido en dos partes: por un lado, se distingue un sistema genérico de conceptos con un color turquesa y con una representación gráfica en forma de árbol; y por otro, un sistema partitivo de conceptos con un color morado y con una representación con corchetes. En un primer nivel, encontramos el concepto superordinado <actividades acuáticas de río>, que se divide en conceptos subordinados en un segundo nivel. En el tercer nivel, distinguimos el sistema de conceptos de parte-todo, que incluye las diferentes partes en las que se divide la canoa.

Además de estos tres tipos de sistemas de conceptos, la norma ISO 704 (2000: 13) también contempla los sistemas de conceptos asociativos, los cuales son definidos como sistemas en los que todos los conceptos se relacionan entre sí mediante relaciones de asociación que pueden variar dentro de un sistema. Estas relaciones de asociación son las denominadas según la TGT como relaciones no jerárquicas y, como decíamos

¹¹⁰ El programa *GeneSis* es una aplicación informática diseñada para la generación de sistemas de conceptos, tanto de forma alfanumérica como gráficamente (Monterde Rey, 2003). URL: <http://www.webs.ulpgc.es/terminol/page3.html>

antes, no son las relaciones de mayor interés para la Terminología tradicional. Entre estas relaciones, se distinguen diferentes tipos (Monterde Rey, 2006: 303):

- Causales: son las relaciones que se establecen entre un hecho u objeto y sus causas.
- Pragmáticas: relaciones temáticas entre conceptos.

Otros tipos:

- Genéticas: relación *productor-producto*.
- De producción: relación *material-producto*.
- De transmisión: relación *emisor-receptor*.
- Instrumentales: relación *instrumento-aplicación*.
- Funcionales: relación *argumento-función*.

Para estos sistemas de conceptos, esta norma ISO no especifica ninguna recomendación para llevar a cabo la representación gráfica. No obstante, los ejemplos que incluye ofrecen una representación mediante flechas con dos punteros. Como ejemplo de esta representación, podría servir la siguiente:

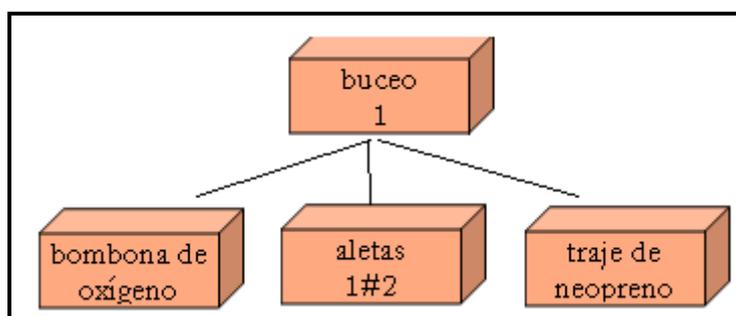


Figura 14. Representación de sistema de conceptos asociativo (relación instrumental) con la actividad de buceo y algunos de sus instrumentos básicos realizada con el programa *GeneSis*.

Esta formalización expresada por la TGT y recogida en las diferentes normas internacionales (ISO 704: 2000, ISO 1087-1: 2000, DIN 2331: 1993, etc.) permite la representación de sistemas de conceptos basados principalmente en relaciones jerárquicas (genéricas y partitivas), puesto que el resto de relaciones no jerárquicas son difíciles de representar de forma clara y concisa. En este sentido, la propuesta de

representación de la Terminología tradicional mediante los sistemas de conceptos se considera muy limitada a la hora de representar todos los posibles tipos de relaciones que se pueden dar en un dominio de especialidad, así como de mostrar la multidimensionalidad de los conceptos.

Por este motivo, y a pesar de que la representación de sistemas simples se considera útil para la terminografía básica (Wright, 1997: 90), los terminólogos han reconocido los problemas que presentan estos sistemas de conceptos simples, estáticos y restringidos por los que aboga la TGT. Uno de los principales problemas se refiere a la simplicidad que denotan estos sistemas, la cual rara vez concuerda con la realidad: «Real concept systems, of course, are normally more complex than the “neat-and-tidy” tree structures found in terminology textbooks» (Meyer et ál., 1997: 106), ya que los conceptos se interrelacionan formando redes complejas, pueden contar con más de un concepto genérico, no son estáticos ni se encuentran delimitados por fronteras concretas, se pueden clasificar en función de unas características u otras, etc. Este último fenómeno es lo que se denomina la multidimensionalidad, que se define como «a phenomenon of conceptual classification that arises when concepts are classified in more than one way within a concept system» (Bowker, 1997: 133). Un ejemplo de multidimensionalidad puede ser el caso de la representación del concepto <vehículo>, que se podría clasificar según diferentes características: una, como *medio de transporte* puede ser <vehículo aéreo>, <vehículo terrestre> o <vehículo acuático>; otra, se puede clasificar según el método de propulsión como <vehículo motorizado> o <vehículo no motorizado>, y por último, según el tipo de carga, podría ser <vehículo de pasajeros>, <vehículo de mercancías> o <vehículo de pasajeros y mercancías> (ibíd.: 134).

Gracias al desarrollo de herramientas informáticas que permiten la representación de relaciones entre conceptos de forma más compleja, se ha evolucionado enormemente en la representación de este tipo de sistemas conceptuales. De hecho, la posibilidad de modelar sistemas de representación del conocimiento más complejos ha dado lugar a un crecimiento en el número de relaciones que se pueden dar entre conceptos, lo que ha provocado un cambio en la aproximación teórica de la Terminología. Según Sager (1990: 55), opinión que secundamos en nuestra investigación:

El punto de vista simplista del pasado de representar las relaciones entre los conceptos mediante tres tipos (genérico, partitivo, otros) ya no tiene validez ni apoyo. Hoy en día se admite que para las aplicaciones prácticas se puede establecer prácticamente cualquier número y tipo de de relación conceptual y declararla como requisito para una necesidad concreta.

Por este motivo, se han ido sucediendo diferentes propuestas de representación conceptual en Terminografía, cuyo objetivo ha sido el superar las limitaciones de los sistemas de conceptos propuestos por la aproximación tradicional y ofrecer la posibilidad de organizar el campo de especialidad de la forma más completa posible.

Como antecedentes de las representaciones conceptuales basadas en formalismos complejos (especialmente ontologías), encontramos a Quillian (1968) y Shapiro y Woddmansee (1971)¹¹¹, y podemos distinguir principalmente dos tipos tradicionales de redes semánticas:¹¹²

- **Grafos conceptuales**, en los que existen dos tipos de nodos: conceptos y relaciones.
- **Redes de marcos** (*frames*), en los que los puntos de unión de los enlaces son parte de la etiqueta del nodo.

De estos dos tipos de redes semánticas, el esquema basado en marcos es el que ha alcanzado una mayor aceptación por parte de los investigadores de ciencia cognitiva y lingüística, puesto que es el que permite mayor flexibilidad y complejidad de los tres. Por su parte, el esquema de grafos también ha despertado mucho interés en la actualidad (Antia, 2000), pero cuenta a su vez con muchos detractores.

Además de estos dos esquemas conceptuales, podríamos incluir aquí las redes semánticas basadas en la relación IS-A, las denominadas taxonomías (cf. Pérez Hernández, 2002), pero consideramos que estas deberían quedar fuera de una tipología

¹¹¹ Recogidos en Moreno Ortiz (2000).

¹¹² En una red semántica, los conceptos (estructuras, clases, marcos, dependiendo del esquema concreto) están organizados en una red en la que existe un nodo superior (*top*) al que se le asigna uno o varios nodos hijos, que a su vez tienen otros conceptos hijos y así sucesivamente hasta que se alcanza el final (*bottom*). Para profundizar en estos tipos de sistemas de representación de conocimiento, véase Moreno Ortiz (2000), que realiza un resumen de los más destacados.

de esquemas de representación conceptual complejos por dos razones básicamente: en primer lugar, por su simplicidad a la hora de representar un conocimiento conceptual, ya que únicamente emplean las relaciones genéricas; y, en segundo lugar, porque se suelen incorporar a los otros dos tipos de redes semánticas.

El esquema basado en marcos (*frames*), que, como decíamos, es el tipo de representación al que se le he prestado mayor atención por su flexibilidad y modularidad, se basa en la Semántica de Marcos (cf. 1.3.2.3.3.1.), por la cual un marco semántico se definía de la siguiente manera:

A semantic frame is a schematic representation of a situation type (e.g. speaking, eating, judging, moving, comparing, etc. -activities and situation which make up our everyday life) together with a list of the typical participations, props, and concepts that are to be found in such a situation; these are the semantic roles, or "frame elements" (FEs) (Atkins y Rundell, 2008: 145).

Y Minsky (1975: 211) definía la teoría de marcos como sigue:

Here is the essence of frame theory: When one encounters a new situation (or makes a substantial change in one's view of a problem), one selects from memory a structure called a *frame*. This is a remembered framework to be adapted to fit reality by changing details as necessary.

La representación del conocimiento mediante marcos se basa, pues, en la siguiente premisa: cuando nos enfrentamos con una situación determinada, intentamos ajustarla a otra parecida de la que ya tenemos experiencia previa y esperamos que aparezcan un número de elementos comunes y sucedan algunas situaciones conocidas. Por tanto, el objetivo de este tipo de representación es la elaboración de marcos semánticos que formalicen el contexto en el que aparecen los conceptos a través de relaciones semánticas, todo lo cual permitirá la activación de procesos y la recuperación de la información almacenada según el contexto. A su vez, estos marcos se pueden relacionar entre sí creando una base de conocimiento basada en marcos, donde estos se organizan de forma jerárquica, según un número de criterios estrictos y otros principios más o menos imprecisos tales como el de similitud entre marcos (Pérez Hernández, 2002).

Las actuales representaciones conceptuales mediante ontologías se basan en mayor o menor medida en un modelo de marcos, aunque, en general, no emplean

sistemas de activación de procesos por tratarse de representaciones con un marcado carácter estático, enfocadas a surtir de información a otras aplicaciones informáticas que contienen los procedimientos.

1.4.3.2. Algunas consideraciones en torno a las ontologías: definición y componentes

Durante la última década, se está produciendo un creciente interés por la representación del conocimiento especializado mediante ontologías en un buen número de disciplinas, especialmente en informática e ingeniería del conocimiento debido a lo que se conoce como la Web Semántica. Según Berners-Lee et ál. (2001):

The Semantic Web is an extension of the current web in which information is given well-defined meaning, better enabling computers and people to work in cooperation.

Esta *nueva* Web es una extensión de la anterior que tiene como objetivo mejorar la búsqueda y la recuperación de información en Internet mediante la anotación de los documentos electrónicos con información sobre su contenido. El lenguaje que se utiliza actualmente, el HTML, solo aporta información acerca de los aspectos visuales o formales de los documentos, como el tipo de letra o las imágenes que incluye, pero no ofrece ningún dato sobre el contenido informativo del texto. Al contrario, la Web Semántica contiene metadatos de carácter semántico, que permiten una recuperación de información de forma más precisa; por este motivo, esta nueva Web es la tendencia actual a la elaboración de páginas web.

Esta Web está formada por varios elementos diferentes y uno de los más importantes son las ontologías, que organizan el conocimiento de un dominio de forma comprensible para un ordenador y son un recurso ideal para que la recuperación de información se haga de un modo más efectivo. En este proyecto de Web, en concreto en el tema de la ontologías, están participando algunos centros de investigación dedicados a la Lingüística Computacional y a la Terminología, pues es evidente que su experiencia con el lenguaje, los conceptos, los términos y las aplicaciones lingüísticas o terminológicas que generan son aspectos que pueden resultar en un desarrollo

beneficioso para la Web Semántica (Vargas Sierra, 2007). Con esta colaboración entre ingenieros y lingüistas se obtiene un beneficio mutuo: por una parte, los lingüistas y terminólogos pueden aportar su conocimiento sobre el funcionamiento y las peculiaridades del lenguaje natural y de la terminología; y por otra parte, la investigación y las herramientas relacionadas con las ontologías pueden resultar de sumo interés para terminólogos y lingüistas, puesto que podrán aplicarlas en las fases del trabajo.

Concretamente en Terminografía, se está asistiendo a un gran aumento de la utilización de ontologías para llevar a cabo los trabajos terminográficos. Ejemplo de ello son las numerosas publicaciones y congresos que tienen lugar en torno a esta línea de investigación, por citar algunas: el número especial dedicado en la revista *Terminologies nouvelles* al tema «Terminología e Inteligencia Artificial» en 1999, los congresos anuales sobre Terminología e Inteligencia Artificial (TIA, *Terminologie et Intelligence Artificielle*), *Terminology and Knowledge Engineering* (TKE), *OntoLex*, entre otros, y revistas como *Ontological Applied*, que dedica una sección a la terminología, confirman el acercamiento y la cooperación entre la Terminología y la Ingeniería del Conocimiento. Esta relación ha dado lugar al fructífero desarrollo de ontologías y de aplicaciones informáticas para la gestión de conceptos en el ámbito de la Terminología, lo que ha permitido un gran avance en la línea de investigación de la ontoterminografía.

No obstante, antes de profundizar en la representación conceptual mediante ontologías y los proyectos que se han llevado a cabo, nos parece conveniente comenzar con la definición de *ontología* y los elementos básicos que la constituyen, así como sus diferentes aplicaciones.

El término *ontología* se ha empleado desde hace muchos siglos en el campo de la filosofía y del conocimiento y, desde la década de los noventa, ha cobrado especial relevancia en el campo de la Inteligencia Artificial, la ingeniería, la representación del conocimiento y la Lingüística Computacional como modelos de representación del conocimiento. Debido a esto, en la actualidad, coexisten dos usos diferenciados del término «ontología», que corresponden a dos ramas del saber diferentes y, por tanto, le atribuyen características y propiedades distintas.

En el campo de la filosofía, la ontología se ha considerado una rama de la metafísica que se ocupa de la naturaleza y organización de la realidad. En el DRAE, se recoge únicamente una acepción, que hace referencia a esta definición en el campo de la filosofía: «[Una ontología es la] parte de la metafísica que trata del ser en general y de sus propiedades trascendentales». Y, por su parte, en el ámbito de la Ingeniería del Conocimiento, se emplea con otro significado más pragmático y aplicado derivado de su significado original. En este ámbito, el término «ontología» se usa para referirse a un conjunto de conceptos organizados jerárquicamente, representados en algún sistema informático cuya utilidad consiste en servir de soporte a diversas aplicaciones que requieren de conocimiento específico sobre la materia que la ontología representa (Moreno Ortiz, 2004: 31) y, a menudo, se emplea como sinónimo de conceptualización. En nuestra investigación, trataremos el término *ontología* con este segundo significado, entendido como un cuerpo estructurado que ofrece una representación compleja del conocimiento de un ámbito de especialidad.

Dentro del ámbito de la Inteligencia Artificial y de la Ingeniería del Conocimiento existen numerosas definiciones de ontologías, pero la más conocida y citada por los autores es la propuesta por Gruber (1993: 199): «*An ontology is a formal, explicit specification of a shared conceptualization*». Esta definición, convertida ya en estándar, incluye los aspectos más relevantes de una ontología y sus principios básicos, a saber:

- El término «conceptualización» (*conceptualization*) se refiere a un modelo abstracto del dominio o de algún fenómeno del mundo que representa del que se pretenden identificar los conceptos más relevantes.
- Con «explícita» (*explicit*) se pretende expresar que los conceptos que forman una ontología deben ser definidos de forma expresa.
- El adjetivo «formal» (*formal*) se refiere al hecho de que la ontología debe materializarse en algún lenguaje informático mediante un formalismo siempre idéntico, de manera que pueda ser reutilizada y entendida por cualquier máquina independientemente del lugar o de la plataforma o idioma del sistema que lo emplee.

- «Compartida» (*shared*) refleja la noción de que una ontología representa un conocimiento compartido y aceptado por el grupo o la comunidad epistemológica a la que se refiere, y no solo la de un individuo.¹¹³

Otras definiciones también muy utilizadas en este ámbito son las siguientes:

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 (Neches *et al*, 1991: 40).

An ontology is a database describing the concepts in the world or some domain, some of their properties and how the concepts relate to each other (Weigand, 1997).

Por último, mostramos aquí la definición que ofrece Roche (2003: 2), que pretende resumir las definiciones más utilizadas en la Ingeniería del Conocimiento:

An ontology is a shared description of Concepts and relationships of a domain expressed in a computer readable language.

Estas definiciones son más completas que la ofrecida por Gruber (1993), puesto que hacen explícitos algunos de los componentes de las ontologías, como son los conceptos, las relaciones y las propiedades, además de incluir la especificación de que deben ser entendidas por el ordenador. Siguiendo con este tema, veremos a continuación cuáles son los componentes de las ontologías, los cuales servirán para representar el conocimiento de algún dominio. Según Gruber (1993), los componentes de las ontologías serían los siguientes:

- **Conceptos** (también llamados **clases**): representan ideas sobre los objetos físicos o abstractos que configuran un dominio y constituyen la pieza básica de estructuración del conocimiento. Los conceptos pueden ser clases de objetos, métodos, planes, estrategias, procesos de razonamiento, etc., y normalmente se organizan en forma de taxonomías a las que se pueden

¹¹³ En la práctica, es prácticamente imposible conseguir el consenso de todos los involucrados en un dominio específico (piénsese, por ejemplo, en una ontología sobre sanidad: ¿quién pone de acuerdo a médicos, pacientes, profesores, alumnos, gestores, etc., en la Terminología, valores y relaciones en el dominio de la medicina?). Por ello, se evitan las ontologías de carácter genérico y se tienden a desarrollar ontologías en ámbitos mucho más restringidos y limitados, puesto que el consenso es más factible. Los creadores de ontologías intentan mantener un equilibrio entre la generalidad y la especificidad de las necesidades de un trabajo concreto (Noy, 2003: 183).

aplicar mecanismos de herencia. De este modo, una clase puede dividirse en subclases que representen conceptos más específicos.

- **Relaciones:** representan un tipo de asociación entre los conceptos de un dominio. Suelen formar la taxonomía del dominio. La mayoría de relaciones enlazan un concepto con otro, por lo que son de tipo binario. Algunas relaciones semánticas básicas son: subclase-de, parte-de, tipo-de, etc.

- **Funciones:** son un tipo concreto de relación donde se identifica un elemento mediante el cálculo de una función que considera varios elementos de la ontología. Por ejemplo, pueden aparecer funciones como: asignar-fecha, categorizar-clase, etc.

- **Instancias:** se utilizan para representar individuos o elementos específicos de una ontología que ejemplifican una clase o una relación. Las instancias hacen remisiones a ejemplos reales y su colección constituye la base de datos o base de conocimiento.

- **Reglas de restricción o axiomas:** son expresiones lógicas que se añaden a la ontología y permiten describir relaciones que deben cumplir los elementos de la ontología. Se establecen a partir de valores específicos de las propiedades y se utilizan para especificar sus restricciones. Por ejemplo: «Si A y B son de la clase C, entonces A no es subclase de B», «Para todo A que cumpla la condición B1, A es C», etc. Los axiomas, junto con la herencia de conceptos, permiten inferir conocimiento que no esté codificado explícitamente en la taxonomía de conceptos.

Además de esta técnica de modelado de ontologías propuesta por Gruber (1993), existen otras que complementan o modifican en cierta medida esta. Por ejemplo, la propuesta por Corcho et ál. (2004: 145), que añade los siguientes componentes a la establecida por Gruber: constantes, que hacen referencia a valores numéricos que no cambian en un determinado dominio (por ejemplo, la mayoría de edad se obtiene con 18 años); atributos, que describen las propiedades de los conceptos y de las instancias; y, por último, las reglas, que se utilizan para inferir conocimiento de la ontología.

Por su parte, Contreras y Martínez Comeche (2008) completan la terminología relacionada con los componentes de las ontologías anteriormente mencionados con los siguientes elementos:

- **Clase abstracta:** Clase que no permite que existan instancias de ella. Se usa para agrupar conceptos, introducir cierto orden en la jerarquía, pero suelen ser demasiado generales para admitir instancias.
- **Instancia indirecta:** Cuando una clase es instancia indirecta de otra, quiere decir que es instancia de alguna de sus clases derivadas. En contraposición a instancia directa, donde no existen clases intermedias.
- **Faceta (Restricción sobre las propiedades):** Es alguna propiedad de la propiedad. Por ejemplo, la cardinalidad, si la propiedad es obligatoria o no, etc.
- **Anotación:** Es el proceso de relleno de instancias a partir de texto libre. Existen dos maneras de anotar texto: la más usual es la inclusión de etiquetas semánticas dentro del texto que se está procesando. Esto implica que el formato del texto debe ser editable y procesable. En el caso de no disponer de este formato, la segunda manera consiste en rellenar las instancias directamente en el modelo, dejando el texto original sin modificar.
- **Herencia:** Propiedad de la relación 'es_un' que permite que las clases relacionadas (heredadas) cuenten con los atributos de la clase con la cual se relacionan (clase padre).
- **Herencia múltiple:** Se da cuando una clase dada hereda o cuenta con las propiedades de dos clases padre con las que establece dos relaciones del tipo 'es_un'.
- **Derivación:** Organización de las clases de la ontología en un árbol de jerarquía mediante sucesivas relaciones 'es_un' (también llamadas kind_of, is_a, o herencias) con la propiedad de herencia. Esta organización permite el encadenamiento sucesivo de herencias desde las clases

de nivel superior a las clases situadas en niveles inferiores, llamadas clases derivadas.

En función de estos términos que se utilizan para representar los conocimientos del dominio de especialidad y, por tanto, de su complejidad, se distinguen dos tipos de ontologías: por un lado, las *ontologías ligeras (lightweight ontologies)*, que incluyen conceptos, taxonomías conceptuales, relaciones entre conceptos y propiedades que los describen; y por otro lado, las *ontologías pesadas (heavyweight ontologies)*, que contienen, además de lo anterior, axiomas y restricciones, que aclaran el significado de los conceptos almacenados en ellas (Lozano Tello, 2002: 10).

En resumen, podemos indicar que una ontología es un sistema de representación del conocimiento que resulta de seleccionar un dominio o ámbito del conocimiento, y de aplicar sobre él un método, a fin de obtener una representación formal de los conceptos que contiene y de las relaciones básicas que existen entre dichos conceptos. Además, una ontología se construye normalmente con relación a un contexto de utilización y, por tanto, especifica una conceptualización de un dominio de especialidad por parte de un grupo de expertos del dominio en cuestión. Por último, todas las conceptualizaciones (definiciones, categorizaciones, jerarquías, propiedades, herencia, etc.) de una ontología pueden ser procesables por ordenador y son independientes de la lengua en la que se trabaje.

En general, y después de haber visto los diferentes componentes de las ontologías, podemos apreciar la similitudes entre los componentes de una ontología (términos, relaciones, vocabulario de un ámbito de especialidad, normas para la inclusión de nuevos términos) y los componentes de los conceptos terminológicos. De esta manera, el hecho de contar con un sistema de representación conceptual formalizado, explícito y, en gran medida, estandarizado y consensuado por la comunidad de expertos de un dominio es algo que debería facilitar y mejorar la labor del terminógrafo (Moreno Ortiz, 2004). Por ello, no debe sorprendernos que a partir de la década de los 90, muchos autores hayan destacado las ventajas que para unos y otros, ingenieros y terminólogos, puede tener el trabajo conjunto (Meyer et ál., 1992, Fischer et ál., 1996; Gillam y Ahmad, 1996; Gamper et ál., 1999).

1.4.3.2.1. Criterios generales para la creación de ontologías

En la construcción de cualquier sistema se deben tener en cuenta una serie de criterios generales, especialmente si uno de los objetivos principales consiste en construir representaciones formales del conocimiento susceptibles de intercambiarse independientemente de la plataforma *hardware*, el lenguaje de programación y los protocolos de red empleados y que, además, puedan intercambiar conocimientos entre distintas aplicaciones, lo que se conoce como interoperabilidad (*interoperability*), como se pretende con las ontologías. A continuación, se presentan los cinco criterios básicos propuestos por Gruber, y aceptados por la gran mayoría de expertos en ingeniería del conocimiento, para facilitar el diseño de ontologías compartidas y reutilizables:

- Claridad y objetividad: una ontología debe proveer al usuario con el significado de la clase de forma objetiva, independientemente del contexto social o computacional. Así pues, las definiciones serán lo más objetivas posibles y deben explicarse también en lenguaje natural.
- Coherencia: las definiciones de una ontología deben ser consistentes para asegurar la corrección de las inferencias realizadas de esta.
- Posibilidad de ampliación (*extensibility*): Una ontología debe permitir la inclusión de nuevos términos o la modificación de los ya existentes, de tal manera que se pueda utilizar el vocabulario ya existente para la creación de definiciones nuevas.
- Sesgo mínimo de codificación (*minimal encoding bias*): Una ontología debe especificarse de forma que no dependa de una codificación determinada, a fin de no provocar limitaciones o restricciones en la representación por causa de la anotación o implementación.
- Compromiso ontológico mínimo (*minimal ontological commitment*): Una ontología debe contener clases generales, a fin de ejercer el menor número posible de restricciones sobre la realidad que representa. De esta manera, la relación entre las clases de una ontología y su conceptualización deberá ser mínima. Asimismo, deberá hacerse la menor cantidad de aserciones subjetivas acerca del mundo modelado.

Por su parte, Gómez Pérez et ál. (2004: 43-44) completan estos cinco criterios básicos propuestos por Gruber (1993) con otros tres, a saber: la representación de conocimiento disjunto y exhaustivo, la minimización de la distancia sintáctica entre conceptos emparentados y la normalización de los nombres.

- Representación de conocimiento disjunto y exhaustivo: Si una ontología cuenta con instancias de subclases que forman conjuntos disjuntos, es decir, que no pueden pertenecer a varias subclases de forma simultánea, se puede establecer una descomposición disjunta. Una ontología es exhaustiva si se define completamente el concepto superordinado, es decir, si se contemplan todas las posibilidades.
- Minimización de la distancia sintáctica entre conceptos emparentados: se recomienda emplear el mismo patrón para definir conceptos emparentados y, así, mejorar la comprensión y reutilización de la ontología.¹¹⁴
- Normalización de los nombres, es decir, se trata de nombrar las clases siempre del mismo modo, siempre que esto sea posible.

Al igual que ocurre con la creación de ontologías en general, los lenguajes que se utilizan para crearlas también deben cumplir una serie de requisitos básicos. A la hora de elegir un lenguaje ontológico, deben tenerse en cuenta los siguientes requisitos:

- Debe poseer una sintaxis bien definida para ser capaz de leer las ontologías y de definir las mediante este lenguaje.
- Debe tener semánticas bien definidas, para ser capaz de comprender las ontologías.
- Debe tener suficiente expresividad para poder capturar varias ontologías.
- Debe ser fácilmente modificable de y hacia otros lenguajes ontológicos.
- Debe ser eficiente a la hora de realizar un razonamiento.

En la actualidad, existen muchos lenguajes que van a permitir la creación de ontologías (OWL,¹¹⁵ DAML+OIL,¹¹⁶ SHOE,¹¹⁷ RDF(S),¹¹⁸ KIF,¹¹⁹ Ontilingua¹²⁰), pero no todos

¹¹⁴ Este criterio se podría considerar muy cercano al de coherencia, aunque restringe la aplicación.

ellos manejan el mismo nivel de expresividad a la hora de definir ontologías. Mientras que algunos lenguajes permiten la creación de ontologías tanto ligeras como pesadas, otros solo permiten el diseño de ontologías ligeras al no posibilitar la inclusión de axiomas o reglas. De todos los lenguajes actuales, el lenguaje OWL (*Web Ontology Language*) se ha convertido en el estándar para el desarrollo de ontologías, así como en el punto de referencia para todos los que trabajan con ellas, ya sean usuarios o desarrolladores de tecnología (Moreno Ortiz, 2002). Este uso común del OWL se debe al hecho de que es la propuesta de lenguaje ontológico del W3C (Consortio World Wide Web), un consorcio internacional que trabaja para elaborar estándares Web.¹²¹

En general, podemos decir que con el cumplimiento de todos estos criterios a la hora de crear una ontología, seremos capaces de garantizar la calidad exigible para cualquier aplicación o tarea en la que se emplee dicha ontología. No obstante, estos criterios están enfocados principalmente a los desarrolladores de ontologías, así como a la creación de ontologías genéricas; por este motivo, más adelante veremos cuáles son los criterios específicos para la creación de ontologías para la ontoterminografía.

1.4.3.2.2. Clasificaciones de ontologías

Las ontologías, como sistemas de representación de conocimiento, se pueden clasificar atendiendo a diversos criterios, dependiendo del punto de vista que se considere. A continuación, veremos algunas de las clasificaciones que se han propuesto, aunque existen otras muchas.

¹¹⁵ <<http://www.w3.org/TR/2004/REC-owl-features-20040210/>>

¹¹⁶ <<http://www.daml.org/2001/03/daml+oil-index.html>>

¹¹⁷ <<http://www.cs.umd.edu/projects/plus/SHOE/index.html>>

¹¹⁸ <<http://www.w3.org/RDF/>>

¹¹⁹ <<http://logic.stanford.edu/kif/dpans.html>>

¹²⁰ <<http://ksl.stanford.edu/software/ontolingua/>>

¹²¹ Para profundizar en el tema de los lenguajes de ontologías y sus características, recomendamos la consulta de Gómez Pérez et ál. (2004: 284 y ss.).

Mizoguchi et ál. (1995) clasifican las ontologías desde la perspectiva del tipo de problema que resuelven. Así pues, consideran cuatro categorías diferentes:

- *Ontologías de contenido*, que se elaboran para reutilizar su conocimiento, vistas como cajas blancas que permiten reutilizar su vocabulario.
- *Ontologías de indexación*, que permiten la recuperación de datos cuando los agentes comparten conocimientos a través de bases de datos.
- *Ontologías de comunicación*, que se utilizan por agentes para obtener respuestas a preguntas concretas, vistas como cajas negras de conocimientos.
- *Meta-ontologías*, que se emplean como marcos para representar ontologías.¹²²

Por su parte, van Heist et ál. (1997) atienden al ámbito de conceptualización de la ontología y presentan la siguiente clasificación:

- *Ontologías de representación de conocimiento*, que proporcionan conceptualizaciones subyacentes a los paradigmas o formalismos de representación del conocimiento, es decir, proporcionan el vocabulario necesario para modelar otras ontologías, utilizando un determinado paradigma de representación de conocimiento.
- *Ontologías genéricas o de alto nivel*, que incluyen términos genéricos reutilizables en diferentes dominios. Como, por ejemplo, los términos: estado, evento, acción, componente, etc. Ejemplos de este tipo de ontologías son: Cyc (Lenat, 1995),¹²³ SUMO (Niles y Pease, 2001)¹²⁴ o PROTON (Terziev et ál., 2003)¹²⁵.

¹²² Este tipo de ontologías es lo que se conoce en otras clasificaciones como ontologías genéricas.

¹²³ Cyc, desarrollada por Cycorp, es una de las mayores bases de conocimiento existentes, cuyo objetivo es la construcción de software inteligente gracias al sentido común aportado por una gran base de conocimiento que posee hechos, reglas y heurísticas para el razonamiento sobre los objetos y eventos de la vida diaria. OpenCyc, la versión de código abierto de la base de conocimiento Cyc, representa el conocimiento general humano y contiene unos 500.000 términos agrupados en 43 grupos temáticos (fundamentales, tiempo y fechas, relaciones espaciales, etc.), unas 15.000 tipos de relaciones y unos 5 millones de axiomas relacionadas con dichos términos, además de un motor de inferencia, un navegador de la base de conocimiento y otras herramientas útiles. <<http://www.cyc.com/>>

¹²⁴ La ontología SUMO (Suggested Upper Merged Ontology), desarrollada en el grupo de trabajo IEEE Standard Upper Ontology, es la ontología formal de libre acceso más grande actualmente. Proporciona

- *Ontologías de dominio o de tareas*, que expresan conceptualizaciones específicas para dominios concretos. Estas ontologías proporcionan vocabularios sobre conceptos y sus relaciones específicas, sobre las actividades que se desarrollan y sobre las teorías y principios que rigen en ese dominio determinado. Por ejemplo, se consideran ontologías de dominios las que están formadas por conceptos de un ámbito concreto, como puede ser el derecho, la medicina, el turismo, el comercio electrónico, etc. o de tareas, aquéllas que incluyen los conceptos relativos al diagnóstico, a la realización de una reserva, a la compra-venta de un piso, etc.
- *Ontologías de aplicación*, que contienen todas las definiciones que son necesarias para modelar los conceptos requeridos por una aplicación particular. Incluyen conceptos tomados de ontologías de dominios y genéricas, a menudo definidas utilizando el vocabulario indicado en las ontologías de representación.¹²⁶
- Por su parte, también tenemos ontologías que se distinguen por el tipo de agente al que van destinadas:
 - *Ontologías lingüísticas*, que se vinculan a aspectos lingüísticos, es decir, aspectos gramáticos, semánticos y sintácticos destinados a su utilización por los seres humanos.
 - *Ontologías no lingüísticas*, que están destinadas a ser utilizadas por robots y agentes inteligentes.
 - *Ontologías mixtas*, que combinan las características de las dos anteriores.

Y, según el grado o nivel de abstracción y razonamiento lógico que permitan, las ontologías pueden ser de dos tipos:

definiciones para términos de propósito general y actúa como base para más ontologías de dominio específico. <<http://www.ontologyportal.org/>>.

¹²⁵ PROTON (PROTo ONtology), desarrollado en el seno del proyecto SEKT, contiene alrededor de 300 clases y 100 propiedades, proporciona cobertura de conceptos generales necesarios para una amplia gama de tareas, incluyendo la anotación semántica, indexado y recuperación de documentos. <<http://proton.semanticweb.org/>>.

¹²⁶ Las ontologías de dominio y de aplicación suelen ser menos reutilizables que las anteriores, puesto que son especificaciones concretas de un dominio determinado.

- *Ontologías descriptivas*, que incluyen descripciones, taxonomías de conceptos, relaciones entre los conceptos y propiedades, pero no permiten inferencias lógicas.
- *Ontologías lógicas*, que permiten inferencias lógicas mediante la utilización de una serie de componentes como la inclusión de axiomas, etc.

Como veíamos al principio, existen numerosas clasificaciones de ontologías, pero en este apartado hemos pretendido tratar las más comunes y extendidas en este ámbito. Ahora, y gracias a las clasificaciones que hemos incluido aquí, podemos clasificar la ontología que pretendemos elaborar en nuestro trabajo de investigación según los criterios que se han utilizado en estas diferentes clasificaciones. De esta manera, nuestra ontología se trata de una *ontología de dominio*, ya que está centrada en el dominio del turismo y, concretamente en el segmento de turismo de aventura; una ontología mixta, puesto que podrá ser entendida y utilizada tanto por personas (utilizando un lenguaje natural) como por máquinas (utilizando un lenguaje de programación y unos estándares adecuados); y por último, una ontología descriptiva, puesto que servirá para presentar la descripción y organización del dominio de especialidad. Por todo esto, nuestra ontología será considerada como una ontología ligera, ya que incluirá los conceptos, relaciones entre conceptos y propiedades que los describen, pero carecerá de axiomas y restricciones más complejos.

1.4.3.3. Criterios específicos para el diseño de ontologías en Terminografía

En el apartado 1.4.3.2.1., hemos tratado los criterios generales que hay que seguir a la hora de elaborar una ontología, así como los requisitos que deben cumplir los lenguajes de ontologías para su utilización. Estos criterios sirven de forma global para la elaboración de cualquier ontología y para cualquier objetivo, pero están enfocados principalmente a las ontologías genéricas. En este apartado, nos centraremos en los criterios específicos necesarios para la elaboración de ontologías dentro de un marco terminográfico, los cuales complementarán a los anteriores y centrarán la elaboración de ontologías dentro de la ontoterminografía. En este caso, dejaremos de lado las ontologías genéricas y nos centraremos en las ontologías de dominio, que, como

veíamos anteriormente (cf. 1.4.3.2.2.), se trata de las ontologías que expresan conceptualizaciones específicas para dominios concretos (en el caso de nuestra investigación, el segmento de turismo de aventura).

Los criterios específicos de estas ontologías de dominio para la ontoterminografía están directamente relacionados con las posibilidades que ofrecen las herramientas de elaboración de ontologías, por lo que vamos a tratar en este apartado estos dos aspectos de forma común. De esta manera, los criterios que cumplan las ontologías serán gracias a los criterios que satisfagan sus editores.

En el ámbito terminográfico, las ontologías, además de ser claras, coherentes y objetivas, permitir la ampliación de conceptos, carecer de sesgos, etc., como veíamos anteriormente, deben ser flexibles, compatibles, además de ser fáciles de manejar. Para adquirir esta naturaleza, las ontologías deben ser diseñadas mediante unos editores que cumplan los siguientes criterios, que son los más importantes a la hora de representar un conocimiento especializado:

- Flexibilidad: un editor de ontologías debe permitir que una ontología sea flexible, a fin de que se puedan alcanzar los objetivos propuestos en el trabajo terminográfico y se pueda dar cuenta de las características del lenguaje. La flexibilidad se puede entender en los siguientes puntos principalmente:

- Nivel de complejidad: El usuario es quien decide la complejidad que desea que presente el sistema, tanto con relación a la extensión del dominio que se representa como a la profundidad a la que se quiere llegar. Por ejemplo, en el caso del turismo, nos podríamos centrar en un único segmento, el turismo de aventura, o pretender abarcar más de un segmento o incluso todo el dominio. Por su parte, aunque seleccionemos un único segmento, podríamos profundizar en él tanto como consideráramos oportuno y, en el caso del turismo de aventura, incluir en la representación todas las herramientas necesarias para realizar una actividad, todo el equipamiento, la normativa, etc. o simplemente las actividades de aventura.

- **Multidimensionalidad:** Una ontología debe ser capaz de representar claramente la multidimensionalidad de los conceptos, es decir, debe permitir que cualquier concepto pueda clasificarse de más de una forma en un sistema conceptual, con objeto de representar adecuadamente la realidad. Por ejemplo, un <vehículo> podrá ser situado en uno o en otro lugar del sistema, según los criterios que se elijan para su clasificación.
 - **Dinamismo:** Una ontología debe permitir la actualización de sus conceptos, de sus relaciones y propiedades, para seguir el dinamismo característico de cualquier lenguaje de especialidad. Es decir, una ontología debe posibilitar las modificaciones, eliminaciones y reposiciones de información contenida, así como la inclusión de nueva información.
 - **Relaciones semánticas no jerárquicas:** Una ontología debe estar compuesta tanto por relaciones jerárquicas (genéricas y partitivas) como por relaciones no jerárquicas, las denominadas también asociativas o *ad hoc*, para representar las posibles relaciones que se crean entre los conceptos de un mismo dominio de especialidad. Por tanto, un editor de ontologías debe dar libertad al usuario para establecer cualquier tipo de relación entre los conceptos, ya sean jerárquicas o no jerárquicas, por lo que es necesario disponer de cierta flexibilidad para crear relaciones de este último tipo (actividad-instrumento, actividad-lugar, producto-material, etc.), que no están claramente determinadas. El objetivo final de esta flexibilidad de creación de relaciones consiste en que se puedan representar las diferentes relaciones entre los conceptos que se encuentran (o surjan) en un dominio de especialidad concreto. Además de la inclusión de estas relaciones, también es importante que las relaciones que se establezcan sean bidireccionales, es decir, cuando se establece una relación entre un concepto y otro se debería crear una relación inversa de forma automática.
- **Facilidad de manejo:** Una ontología en el ámbito terminográfico, o mejor dicho, su editor, debe ser fácil de manejar, puesto que sus principales

usuarios serán terminógrafos o traductores (lingüistas en general) que posiblemente no cuenten con un conocimiento profundo en ingeniería del conocimiento. De esta manera, tiene que ser lo más intuitivo posible y lo menos complicado, a la hora de crear la ontología de dominio así como a la hora de manejarla, extraer información o realizar cualquier otra operación, a fin de obtener buenos resultados en el trabajo terminográfico.

- **Compatibilidad:** Una ontología debe estar diseñada en un formato lo más estándar posible para poder ser reutilizada posteriormente, ya sea en el mismo proyecto terminográfico o en otro. De esta manera, el editor de ontologías debe seguir unos estándares apropiados para su creación y su posterior reutilización. Como mencionábamos anteriormente, existen múltiples lenguajes de ontologías, pero el lenguaje OWL es el que se considera como estándar. En el campo de la ontoterminografía, no es suficiente con que una ontología esté diseñada según el estándar del lenguaje de ontologías Web (OWL), sino que también es recomendable que su terminología esté diseñada según el estándar internacional existente para el intercambio de bases de datos terminológicas, el Term Base eXchange (TBX). Este estándar define una serie de lenguajes de marcado terminológico¹²⁷ que facilitan el intercambio de datos terminológicos entre usuarios, ya sean estos traductores y escritores o aplicaciones y sistemas informáticos (como programas de traducción asistida). En este sentido, permiten el flujo de información terminológica mediante un ciclo de producción de información, tanto dentro de una organización como a proveedores y agentes externos (ISO 30042). Este estándar es de código abierto e incluye las categorías de datos recogidos en la norma ISO 12620 (1999) (cf. 2.6.1.4.1.). De esta manera, mediante el empleo de este estándar internacional, conseguimos que nuestros datos terminológicos (ya sean simples bases de datos u ontologías) sean compatibles con otros sistemas de gestión terminológica y que, por tanto, sean reutilizables en otro momento.

¹²⁷ «Terminological markup language» viene definido en la norma ISO 30042 como sigue: «XML application for describing a terminological data collection conforming to the constraints expressed in ISO 16642 (Terminological markup framework)».

Con estos criterios específicos sobre las ontologías y sus editores, los cuales complementan a los criterios generales que vimos en el apartado anterior, se garantiza la creación de ontologías de dominio adecuadas para cualquier trabajo ontoterminográfico. No obstante, en algunos trabajos podrían surgir otros criterios más específicos, pero estos ya estarían más relacionados con el tipo de proyecto que se quisiera realizar que con la creación de ontologías.

A continuación, seguiremos este tema y trataremos algunas de las herramientas más necesarias para la elaboración de ontologías en un marco ontoterminográfico.

1.4.3.4. Herramientas para la elaboración y gestión de ontologías

En los últimos años, el número de aplicaciones y utilidades para la construcción y manejo de ontologías ha crecido exponencialmente debido al gran interés que estos recursos han suscitado en aplicaciones tales como la Web Semántica, el comercio electrónico o la gestión del conocimiento, entre otros. En el campo específico de la Terminografía, sin embargo, existen pocas aplicaciones disponibles que permitan al terminógrafo desempeñar su trabajo plenamente en un marco ontoterminográfico y, por tanto, se ve obligado a utilizar herramientas que no se encuentran diseñadas para este cometido y que no satisfacen todas sus necesidades.

En general, dentro de un marco ontoterminográfico, podemos distinguir tres herramientas esenciales para la consecución de los objetivos: los editores de ontologías, los buscadores y repositorios de ontologías y, por último, los sistemas de gestión ontoterminográficos.¹²⁸ Además de estas, existen otras categorías de herramientas relacionadas con las ontologías, como por ejemplo las herramientas para la evaluación de ontologías, para la fusión y alineación de ontologías, para el aprendizaje de las ontologías, entre otras, pero todas estas requieren de un conocimiento profundo del lenguaje ontológico subyacente así como de aspectos relacionados con la programación, la ingeniería del conocimiento, etc. que escapan normalmente al conocimiento del

¹²⁸ Estos últimos serán tratados en el capítulo 3 (cf. 3.2.1.) en profundidad, junto con otros sistemas de gestión terminológica que ofrecen la posibilidad de añadir cierta información semántica.

terminógrafo o equipo del trabajo terminográfico, además de tratarse de herramientas más enfocadas a aplicaciones diferentes a la ontoterminografía.¹²⁹

En el primer grupo, encontramos los editores de ontologías, también denominados herramientas para el desarrollo de ontologías (Gómez et ál., 2003: 293). Estas herramientas son aplicaciones informáticas diseñadas para ayudar a la creación y edición de ontologías desde cero o reutilizadas. Aparte de esta funcionalidad de edición y navegación, estas herramientas incluyen normalmente documentación, importación y exportación de ontologías, librerías y motores de búsqueda adjuntos.

Por lo general, estos programas están dirigidos a la construcción de ontologías, pero no a la elaboración de recursos terminológicos. De esta manera, permiten introducir cierta información lingüística, como por ejemplo definiciones breves, descripciones, etc., pero no son recomendables para utilizarlos como editores de recursos ontoterminográficos, puesto que su función consiste simplemente en crear representaciones de conocimiento pero no proporcionar información sobre ellas (Durán Muñoz, 2010b: 115). En este sentido, estas herramientas se pueden utilizar como apoyo a la hora de realizar una representación conceptual de un dominio, pero no satisfacen todas las necesidades de un terminógrafo trabajando en un marco ontoterminográfico.

La misma autora (ibíd.) divide estas herramientas para la edición de ontologías en dos grupos diferenciados: por un lado, considera aquellas herramientas que permiten la elaboración de ontologías propiamente dichas; y por otro, aquellas con las que se crean esquemas o mapas conceptuales.¹³⁰ De los dos grupos distinguidos, existen actualmente gran cantidad de programas en el mercado, tanto de software libre como comerciales. Como ejemplo de editores de mapas conceptuales, podemos citar los siguientes: *MindMapper* (<http://www.mindmapper.com/>), un programa comercial muy

¹²⁹ Para profundizar en estas herramientas y sus aplicaciones, se recomienda la consulta de las siguientes fuentes: Duineveld *et ál.* (2000), Gómez Pérez (2002), Denny (2004), Navarro Galindo y Samos Jiménez (2007), que ofrecen una amplia clasificación de estas herramientas, así como una evaluación de sus características y una comparativa entre las diferentes herramientas.

¹³⁰ En algunos casos, estos editores de mapas conceptuales se pueden asemejar a las ontologías (dependiendo de la información que permitan introducir), pero que no llegan a serlo, puesto que la edición de estos mapas es muy superficial y reducida. Por ejemplo, la mayoría no permite crear relaciones diferentes de las relaciones jerárquicas o de las ya preestablecidas, no permiten incluir información lingüística acerca del término, etc. Estos editores de mapas conceptuales se utilizan sobre todo en la enseñanza, con objeto de facilitar el aprendizaje de materias, conceptos nuevos, etc. y se han aplicado en muy pocos proyectos ontoterminográficos.

potente que requiere licencia para la realización de mapas conceptuales; CmapTools (<http://cmap.ihmc.us/>), uno de los programas de software libre más utilizados para crear mapas y esquemas conceptuales, creado por el Institute for Human and Machine Cognition (IHMC); DigiDocMap (<http://www.mapasconceptuales.com/>), programa de la Universidad Pompeu Fabra que permite la creación de mapas conceptuales en línea, sin necesidad de instalarlo en el ordenador; Text2MindMap (<http://www.text2mindmap.com/>), un programa en línea que permite la conversión automática de textos escritos con distintos niveles de tabulación en mapas conceptuales.



Figura 15. Ejemplo de fragmento de mapa conceptual creado con *MindMapper*

Como decíamos, con respecto a los editores de ontologías propiamente dichas, también hay disponible gran número de programas en la actualidad. Así pues, el usuario deberá decidir qué programa utilizar según las características que presente (lenguaje que emplean, granularidad en los esquemas, componentes de las ontologías, opciones de navegación, etc.) y según los objetivos que se haya planteado en su tarea.

De todos los editores de ontologías disponibles en la actualidad, *Protégé* (<http://protege.stanford.edu/>) es el más utilizado, tanto dentro como fuera del ámbito terminográfico, por ser uno de los más completos y fáciles de manejar. Se trata de una herramienta desarrollada por el grupo SMI (Stanford Medical Informatics), de la Universidad de Stanford, gratuito y de código abierto, que cuenta con una sólida base de usuarios y desarrolladores y con una multitud de extensiones (*plug-ins*) y listas de correo de apoyo. Este programa se emplea para construir ontologías de dominio, generar formularios de entrada de datos y efectuar la propia entrada de datos en varios lenguajes estándares (OWL, XML Schema y RDF). 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 mediante un

modelo de conocimiento compatible con el protocolo OKBC (Open Knowledge Base Connectivity).¹³¹ Y el segundo permite desarrollar ontologías utilizando el lenguaje ontológico estándar OWL.¹³²

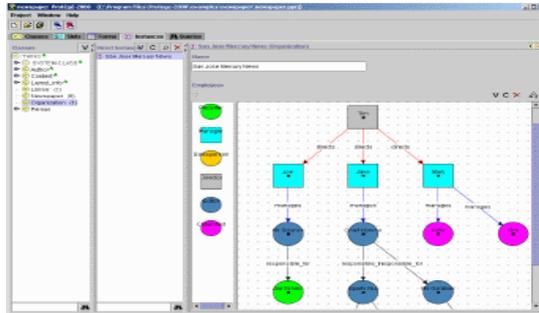


Figura 16. The Protégé Ontology (<http://protege.semanticweb.org>)

Además de este editor, existen otros muchos disponibles en la Web. A continuación, indicamos algunos de ellos:

- KAON (<http://kaon.semanticweb.org/>): un gestor de ontologías de código abierto, desarrollado por la Universidad de Karlsruhe, que permite la creación de ontologías mediante estructura en forma de árbol y con una extensión del lenguaje RDF(S). Incluye un conjunto de herramientas para crear y gestionar ontologías de forma flexible y otras herramientas para construir aplicaciones basadas en ontologías.
- OntoEdit (<http://www.ontoknowledge.org/tools/ontoedit.shtml>): un editor de ontologías desarrollado inicialmente por el Instituto AIFB en la Universidad de Karlsruhe, pero que actualmente está comercializada por Ontoprise GmbH. Existe una versión libre y otra profesional. Esta última incluye muchas más funcionalidades que la de libre acceso, como un motor de inferencia, instrumentos de consulta gráficas, editores de reglas, etc., lo que la ha convertido en la versión más completa del mercado.

¹³¹ Este protocolo permite acceder a bases de conocimiento almacenadas en diferentes sistemas de representación de conocimiento y las cuales pueden estar basadas en diferentes métodos de representación de conocimiento. <<http://www.ai.sri.com/~okbc/>>

¹³² En Maroto García (2007: 263-271) se puede consultar una descripción más detallada de la arquitectura, características y operabilidad del programa.

- SWOOP (<http://code.google.com/p/swoop/>): un editor de ontologías inspirado en hipertexto que sirve para crear, editar y modificar ontologías en OWL. Fue desarrollado por el laboratorio MIND de la Universidad de Maryland, pero actualmente es un recurso de código abierto.
- WebODE (<http://webode.dia.fi.upm.es/WebODEWeb/index.html>): una herramienta desarrollada por el grupo de investigación sobre ontologías de la Universidad Politécnica de Madrid. No se trata de una herramienta únicamente diseñada para el desarrollo de ontologías, sino que se trata de una plataforma de trabajo que permite llevar a cabo la mayoría de las actividades implicadas en la construcción de una ontología (conceptualización, razonamiento e intercambio). Este editor permite exportar las ontologías en XML y otros lenguajes ontológicos como RDF(S), OIL, DAML+OIL, OWL, entre otros, por lo que flexibiliza el intercambio de ontologías entre diferentes editores, así como su reutilización y desarrollo de otras aplicaciones.
- PROTON Ontology (PROTo ONtology) (<http://proton.semanticweb.org/>): una plataforma que sirve para construir ontologías, la anotación semántica, la indexación y la recuperación de información. Esta herramienta ha sido desarrollada dentro del programa SEKT (Semantically-Enabled Knowledge Technologies) de la Unión Europea
- Ontolingua (<http://www.ksl.stanford.edu/software/ontolingua/>): un editor de ontologías desarrollado por el Knowledge Systems Laboratory (KSL) de la Universidad de Stanford, cuyo principal objetivo es facilitar el desarrollo colaborativo de ontologías y proporcionar un repositorio de las mismas, así como permitir la creación, edición, modificación, navegación y utilización de ontologías mediante la Web.

En el segundo grupo de herramientas, encontramos las aplicaciones que permiten la búsqueda, consulta y recuperación de ontologías en la Web, a fin de facilitar la reutilización de las ontologías que se van creando en diferentes proyectos. Entre otras, podemos citar las siguientes herramientas:

- Swoogle (<http://swoogle.umbc.edu/>): un buscador de ontologías que permite la indexación, recuperación y organización de la información para documentos de la Web Semántica.
- ONKI2 (<http://www.yso.fi/onki2/?l=en>): un buscador de ontologías formado por un repositorio de ontologías y vocabularios. El objetivo de este buscador consiste en proporcionar los medios para encontrar y utilizar ontologías para indexadores de contenido, buscadores de información y desarrolladores de ontologías. Actualmente permite el acceso a 78 ontologías de diferentes dominios.
- Ontosearch (<http://www.ontosearch.org/>): un buscador de ontologías diseñado para ayudar a los usuarios a encontrar información ontológica basada en RDF. Es compatible con varios algoritmos de visualización y de representación, lo que permite al usuario realizar evaluaciones rápidas sobre los archivos recuperados.
- KPOntology
(http://www.isoco.com/innovacion_aplicaciones_kpontology.htm): una librería para gestionar ontologías, basada en un interfaz con un alto nivel de abstracción, que permite el uso de diferentes tecnologías como JENA, Sesame o WebODE.
- SchemaWeb (<http://www.schemaweb.info/>): un repositorio de esquemas RDF expresado en lenguajes RDFS, OWL y DAML+OIL con numerosos recursos y herramientas sobre ontologías. Se trata del mayor repositorio de ontologías y recursos sobre ontologías, cuyo objetivo es recopilar y ordenar la información sobre esquemas RDF que se han publicado en la Web y dar acceso a ellos.

Este segundo grupo es bastante útil a la hora de trabajar con ontologías, puesto que permite realizar búsquedas y consultas de ontologías ya elaboradas en proyectos anteriores y, por tanto, la reutilización de estas en proyectos nuevos. En este sentido, reduce en gran medida el trabajo de la construcción de ontologías. No obstante, los terminógrafos se verán obligados normalmente a realizar modificaciones y ajustes a las

ontologías encontradas para ajustarlas a los proyectos en los que estén trabajando, a fin de que satisfagan las necesidades del momento.

Además de estas herramientas diseñadas para la construcción o búsqueda de ontologías también existen otras herramientas que asisten al terminógrafo a la hora de realizar tareas relacionadas con la creación de estos recursos. Por ejemplo, encontramos *TerminoWeb* (http://terminoweb.iit.nrc.ca/terminoWeb-v2_e.html), una aplicación informática, desarrollada por C. Barrière, cuyo principal objetivo consiste en detectar las relaciones semánticas entre los términos de un corpus textual. Esta herramienta es un prototipo que permite la compilación de un corpus, tanto a través de la Web como mediante la subida de un corpus ya compilado, así como la extracción de términos y de relaciones semánticas a partir de dicho corpus textual. Otras herramientas son los extractores semánticos, aunque su aplicación no está demasiado extendida por los terminógrafos, puesto que requiere de un trabajo previo de anotación semántica muy laborioso y las aplicaciones actuales no dan muy buenos resultados. Un ejemplo de este tipo de herramientas es GATE (<http://gate.ac.uk/ie/>), de la Universidad de Sheffield, una aplicación que ofrece un entorno de desarrollo de componentes y recursos de PLN, a la vez que ofrece algunas aplicaciones como la extracción de información semántica sobre tipos de eventos, entidades o relaciones previamente especificadas a través de anotación semántica. Por último, podemos incluir aquí los repositorios de patrones léxico-semánticos multilingües que representan relaciones semánticas, con objeto de facilitar la elaboración de ontologías de nueva creación a través de la reutilización de patrones previamente incorporados al repositorio. El funcionamiento de estos estos repositorios es el siguiente: cuando los usuarios introducen una frase en lenguaje natural, es decir, sin ninguna marca de procesamiento, el sistema les ofrece su correspondiente patrón léxico-semántico y su respectivo patrón ontológico. De esta manera, el repositorio se puede considerar como un medio para adquirir conocimientos y elaborar ontologías de nueva planta a partir de oraciones en lenguaje natural (Montiel Ponsoda, 2011). Como ejemplo de estos repositorios, encontramos actualmente los siguientes: LoaWiki:CPRepository, uno de los repositorios más antiguos desarrollado por el Laboratorio de ontología aplicada del Consejo de Investigación nacional italiano,¹³³ el

¹³³ <<http://wiki.loa-cnr.it/index.php/LoaWiki:CPRepository>>

Ontology Design Patterns (ODPS) Public Catalogue¹³⁴ para la elaboración de ontologías en el dominio de la biología, creado por investigadores de la Universidad de Manchester en el marco del proyecto GENE ONTOLOGY; y, por último, el Ontology Design Patterns Portal,¹³⁵ un portal de la Web semántica Web desarrollado por investigadores que participan en el proyecto europeo para apoyar el desarrollo de aplicaciones semánticas denominado NeOn.¹³⁶ Todos estos repositorios son públicos y se encuentran disponibles en la Red, lo que contribuye a la idea de reusabilidad de patrones léxico-semánticos.¹³⁷

Una vez que hemos tratado ampliamente las ontologías, sus características y aplicaciones, vamos a finalizar con las ventajas e inconvenientes que presentan estos recursos a la hora de utilizarlos en general, y particularmente en un marco ontoterminográfico.

1.4.3.5. Ventajas e inconvenientes del empleo de ontologías en Terminografía

En general, podemos decir que las ontologías favorecen la comunicación entre personas y organizaciones; logran la interoperabilidad entre sistemas informáticos, y favorecen el razonamiento automático, la reutilización del conocimiento, etc. como veremos a continuación (Abián, 2005; Lorente Casafont, 2005):

- Las ontologías favorecen la comunicación entre personas y empresas, al proporcionar una forma de representar y compartir el conocimiento de un dominio mediante un vocabulario común y controlado, de modo que se eliminan confusiones y ambigüedades conceptuales y terminológicas. En otras palabras, se alcanza un acuerdo a través de la representación conceptual del dominio de especialidad en cuestión que permite la homogeneización de

¹³⁴ <<http://www.gong.manchester.ac.uk/odp/html/index.html>>

¹³⁵ <<http://www.ontologydesignpatterns.org>>

¹³⁶ <http://www.neon-project.org/nw/Welcome_to_the_NeOn_Project>

¹³⁷ Para profundizar en el uso y elaboración de estos repositorios, remitimos a la tesis doctoral de Montiel Ponsoda (2011), donde se presenta la metodología seguida para la elaboración de este tipo de repositorios, además de su puesta en práctica para la elaboración de ontologías.

realidades y situaciones diferentes creadas a través de una terminología sin consenso.

- Las ontologías sirven para conseguir que los sistemas o aplicaciones interoperen, es decir, que trabajen conjuntamente de una forma automática, sin esfuerzo por parte del usuario. Para ello, es necesario que ambas aplicaciones utilicen la misma terminología o que cuenten con una ontología, que ejerce de traductora entre ambas. La ontología actuaría como puente entre ambas, como una especie de lengua común.

- Las ontologías resultan muy útiles para facilitar el razonamiento automático, es decir, sin intervención humana. Partiendo de unas reglas de inferencia, un motor de razonamiento puede usar los datos de las ontologías para inferir conclusiones de ellos.

- Las ontologías permiten la reutilización del conocimiento, ya que pueden ser importadas y exportadas a diferentes sistemas de gestión o editores de ontologías, siempre y cuando estén desarrolladas en un formato de intercambio de conocimiento estándar y hayan sido previamente consensuadas por la comunidad experta.

- Las ontologías posibilitan la ejemplificación de las representaciones conceptuales creadas mediante la inclusión de instancias, que aportan una información esencial para la extracción de información (por ejemplo, las personas concretas que realizan una acción, con su lugar y fecha de realización).

- Las ontologías ofrecen una granularidad muy elevada, es decir, proporcionan un desarrollo semántico profundo mediante la introducción de conceptos, atributos (o propiedades), relaciones jerárquicas y no jerárquicas, lo que permite que un concepto esté bien definido dentro del dominio.

Más concretamente, dentro de un marco ontoterminográfico, las ontologías ofrecen además las siguientes ventajas, las cuales se corresponden con las funciones que deben cumplir los sistemas de conceptos según la norma ISO 704 (2000: 12):

- Las ontologías permiten la representación conceptual a partir de los conocimientos propios de un dominio de especialidad de forma ordenada y sistemática, lo que ejemplifica y aclara el conocimiento de dicho dominio, tanto para consensuar la terminología entre expertos como para facilitar la comprensión a los no especialistas y evitar, así, la polisemia.
- Las ontologías posibilitan la clarificación de las relaciones entre conceptos, ya sean estas jerárquicas o no jerárquicas.
- Las ontologías sirven de base para una terminología uniforme y estandarizada, es decir, para conseguir una comunicación controlada y consensuada.
- Las ontologías facilitan el análisis comparativo de los conceptos y los términos en diferentes lenguas, al representar el sistema conceptual de un dominio especializado mediante sus conceptos, relaciones, propiedades, instancias, etc.
- Las ontologías facilitan la elaboración de definiciones de forma sistemática, puesto que establecen una homogeneización y orden a la hora de definir los conceptos.

Además de estas ventajas relacionadas con esta norma ISO, las ontologías presentan a su vez otras ventajas, más relacionadas con la concepción de la Terminología moderna. Entre ellas, podemos destacar la representación de la multidimensionalidad de los términos mediante la estructuración de conceptos en diferentes sistemas conceptuales y con diversas relaciones semánticas, así como la representación del dinamismo de los conceptos dentro de un mismo campo de especialidad mediante la modificación de sus propiedades o relaciones, la inclusión de nuevos conceptos o la eliminación de otros.

No obstante, a pesar de todas las ventajas expuestas hasta el momento, debemos tener en cuenta también los inconvenientes que presentan estas herramientas de representación del conocimiento (Herrero Jiménez, 2004: 18-19), especialmente en lo referente al desarrollo de ontologías desde cero. Los resultados pueden ir mejorando con el desarrollo de herramientas, uso de lenguajes ontológicos estándar, ontologías más

específicas, etc., pero existen una serie de limitaciones de fondo que no se pueden obviar y a los que vamos a referirnos a continuación:

- Cada sistema tiene expresado los conocimientos en un lenguaje ontológico o en una herramienta, lo que provoca la falta de reutilización e intercambio. Este inconveniente se podría subsanar si todos los desarrolladores de ontologías o de editores ontológicos utilizasen lenguajes ontológicos estándar como OWL, aunque no siempre es posible o utilizado.
- Existe una problemática con la representación de los sinónimos al no existir un consenso en cuanto a la representación de esta relación. Algunos pueden considerar los sinónimos con la relación IS-A; otros, como variantes del término, etc., lo que provoca una desorganización en la representación del conocimiento. Este inconveniente se podría subsanar si se llega a un consenso general para la representación de esta relación que se siguiera por todos los editores de ontologías.
- El traslado del conocimiento desde los textos especializados y las mentes de los especialistas de un dominio a una representación abstracta y efectiva de dicho dominio (la conceptualización) resulta a menudo muy complicada, por lo que la comunicación de desarrolladores de ontologías (ya sean ingenieros o terminógrafos) y especialistas del dominio de trabajo en cuestión debe ser constante y bidireccional durante todo el proceso de trabajo.

En síntesis, las ontologías presentan una serie de ventajas que las convierten en unas herramientas muy útiles y relevantes para numerosas disciplinas, entre las que se encuentra la Terminografía, permitiendo crear recursos más completos y de forma más sistemática. Por lo general, y gracias a las ventajas que traen consigo, las ontologías permiten la consecución de objetivos y de mejores resultados en los trabajos en los que se emplean; no obstante, es importante tener en cuenta los inconvenientes o limitaciones que presentan e intentar subsanarlos en la medida de lo posible, ya sea utilizando lenguajes ontológicos estándar o acuerdos previos de representación del conocimiento en la comunidad especialista.

1.5. PROYECTOS BASADOS EN LA REPRESENTACIÓN DEL CONOCIMIENTO ESPECIALIZADO MEDIANTE CORPUS ESPECIALIZADOS

Por último, y para cerrar este repaso de las nuevas líneas de investigación en Terminografía, trataremos algunos de los proyectos que han utilizado ontologías en sus trabajos terminográficos basados en corpus textuales, es decir, dentro de un marco ontoterminográfico. Empezaremos por el proyecto COGNITERM (cf. 1.5.1.), que fue el pionero en introducir el uso de las ontologías a la práctica terminográfica para la representación del conocimiento, y seguiremos con el proyecto ONCOTERM (cf. 1.5.2.), el proyecto GENOMA-KB (cf. 1.5.3.) y, por último, el proyecto PROLEX (cf. 1.5.4.). Estos proyectos tienen como objetivo la elaboración de una base de conocimiento terminológica, pero su aproximación es ligeramente diferente y sus objetos de estudio son campos de especialidad diversos. No obstante, a pesar de no compartir los mismos objetivos, todos los resultados de los proyectos están basados en corpus textuales y, por tanto, en el estudio de los términos *in vivo*, es decir, en la lengua real y de forma contextualizada en la situación comunicativa.

Antes de comenzar, consideramos relevante clarificar qué se entiende por «base de conocimiento terminológica», con objeto de diferenciarlas de las bases de conocimiento léxicas como WordNet,¹³⁸ puesto que la mayoría de los trabajos ontoterminográficos que se han llevado a cabo y se están realizando tienen como objetivo la creación de este tipo de recursos. Asimismo, realizaremos una breve distinción entre «base de datos» y «base de conocimiento».

¹³⁸ WordNet® es una base de datos léxica en inglés, desarrollada bajo la dirección de G. A. Miller, de libre acceso. En esta base de datos, los nombres, verbos, adjetivos y adverbios están agrupados en conjuntos de sinónimos cognitivos, denominados synsets, cada uno de los cuales identifica un concepto diferente. Los synsets están vinculados entre sí mediante relaciones semántico-conceptuales y léxicas, lo que ha dado lugar a una red de conceptos que se puede navegar a través de un navegador interno. A partir de esta base de datos monolingüe, se han ido desarrollando otras bases de conocimiento multilingüe como EuroWordNet (<http://www.illc.uva.nl/EuroWordNet/>), MultiWordNet (<http://multiwordnet.fbk.eu/english/home.php>), Spanish WordNet (http://www.lsi.upc.edu/~nlp/web/index.php?Itemid=57&id=31&option=com_content&task=view), etc., que han mantenido la misma estructura de synsets de WordNet y la han completado con otros idiomas. URL: <http://wordnet.princeton.edu/>.

La distinción entre «base de datos» y «base de conocimiento»¹³⁹ ha sido objeto de discusión durante mucho tiempo entre estas comunidades de científicos, pero actualmente se han conseguido fijar una serie de diferencias concretas en cuanto a metodologías, técnicas de implementación, usos, contenidos y objetivos entre ambos recursos. En general, podemos decir que un modelo de datos se asocia con una base de datos, mientras que su correspondiente en una base de conocimiento es el denominado esquema de representación. Además de esta, existen otras distinciones importantes que mencionamos a continuación.

Las bases de datos son unos recursos previos, y por tanto antecedentes, de las bases de conocimiento. Desde su aparición a mediados de la década de los cincuenta han ido pasando por diferentes fases de evolución: primero, las bases de datos estructuradas jerárquicamente; después, las bases de datos de red; y posteriormente, las bases de datos relacionales. El objetivo principal de estos recursos es, como su nombre indica, almacenar grandes cantidades de datos organizados siguiendo un determinado esquema o modelo de datos que facilite su almacenamiento, recuperación y modificación.

Por su parte, las bases de conocimiento (*Knowledge Base*, KB) surgieron en una etapa muy posterior a partir de la investigación en Inteligencia Artificial y como respuesta a las necesidades de las aplicaciones en esta disciplina. Las bases de conocimiento son una forma avanzada de base de datos que no solo pretenden almacenar, recuperar y modificar grandes cantidades de información, sino también plasmar elementos de conocimiento (generalmente en forma de hechos y reglas de inferencia), así como la manera en que este conocimiento ha de ser utilizado. Este recurso lingüístico es un modelo de un mundo/empresa/sección de la realidad, como indican Mylopoulos y Levesque (1984),¹⁴⁰ en el que se considera el mundo/universo como una colección de individuos o entidades y el conjunto de relaciones que existen entre esos individuos. La colección de individuos, más las relaciones entre estos, constituye un estado, cuyos cambios causan la creación o modificación de individuos o

¹³⁹ Moreno Ortiz (2000) ofrece una explicación muy detallada de las características que diferencian las bases de datos de las bases de conocimiento.

¹⁴⁰ Citado por Pérez Hernández (2002).

de las relaciones entre ellos (Arano, 2003). Para representar dicho conocimiento, este recurso requiere de una notación precisa, que se denomina esquema de representación. El tipo de esquema de representación de mayor difusión son las redes semánticas que estructuran sus datos en nodos que simbolizan a los conceptos unidos por arcos que representan las relaciones conceptuales. Como veíamos en el apartado 1.4.3.2.2., los tipos de esquemas pueden ser los grafos conceptuales o las redes de marcos (*frames*), que es el esquema más explorado y utilizado por los investigadores.

En Lingüística, estas bases de conocimiento han dado lugar a las bases de conocimiento léxicas y a las bases de conocimiento terminológicas o especializadas, que han tenido como precedentes a las bases de datos léxicas y terminológicas. Estas bases de conocimiento, como veíamos anteriormente, permiten acceder no solamente a la información almacenada, sino que también dan la posibilidad de acceder y recuperar información semántica mediante la consulta de una ontología. La diferencia entre la base de conocimiento léxica y la terminológica consiste en el tipo de información que contienen: la primera está enfocada a la lengua general y, por tanto, contendrá unidades léxicas, es decir, sin carga de especificidad de dominio; y la segunda está dirigida a la terminología de uno o varios campos de especialidad y, por tanto, está formada por unidades terminológicas (unidades léxicas con carga de especificidad). En este sentido, encontramos de nuevo la contraposición de Lexicografía y Terminografía que comentábamos anteriormente.

En concreto, las bases de conocimiento terminológico se conciben como un recurso que está orientado al concepto y no al término, como las bases de datos terminológicas, y, además, permiten aglutinar información gramatical, textual, terminológica, documental y semántica (Cabré Castellví, 2006: 100). De esta manera, la información que ofrecen es mucho mayor y más completa que cualquier base de datos terminológica, por lo que los resultados de la búsqueda son más satisfactorios, al permitir el acceso a más información y más variada. Por ejemplo, estos recursos dan la posibilidad de recuperar los contextos en los que aparecen las unidades terminológicas, permiten la recuperación de información semántica (relaciones entre conceptos), etc., lo que posibilita la consulta de los términos de forma integrada en un contexto y en una

situación comunicativa y no, como ocurre en las bases de datos, de forma aislada y descontextualizada.

A continuación, trataremos los proyectos por separado y comentaremos sus características principales y las contribuciones aportadas a la ontoterminografía.

1.5.1. PROYECTO COGNITERM

El proyecto COGNITERM, dirigido por el profesor Skuce de la Universidad de Ottawa (<http://aix1.uottawa.ca/~imeyer/research.htm>), es uno de los primeros proyectos que ha aplicado la representación del conocimiento mediante ontologías a la Terminografía, puesto que estos autores han sido pioneros en el desarrollo e implementación de un prototipo de base de conocimiento terminológico desarrollado en el marco del proyecto COGNITERM. Este proyecto se desarrolló en el laboratorio de Inteligencia Artificial de dicha universidad y consistió principalmente en la construcción de una base de conocimiento terminológica bilingüe (francés-inglés) en el dominio de las tecnologías de almacenamiento ópticos, como discos ópticos, unidades de almacenamiento, etc.

De acuerdo con los autores (Meyer et ál.: 1992: 958), COGNITERM consiste básicamente en una combinación de una base de datos terminológica convencional y una base de conocimiento. Cada concepto está representado en una estructura en forma de marco denominada descriptor del concepto (*concept descriptor*, CD), que contiene dos tipos de categorías de información: por un lado, se encuentra la categoría de información conceptual, que corresponde con el componente de la base de conocimiento y está formado por las propiedades conceptuales, sus valores y las relaciones; y por otro, la categoría de información lingüística, que se trata del componente de la base de datos terminológica y que proporciona toda la información lingüística contenida normalmente en una base de datos (definición, equivalentes, ejemplos, etc.). La base de conocimiento se puede visualizar gráficamente en diferentes tipos de relaciones semánticas, tanto jerárquicas como no jerárquicas.

Para llevar a cabo este trabajo de construcción de la base de conocimiento, desarrollaron el sistema de gestión de conocimiento denominado CODE (*Conceptually Oriented Design Environment*) dentro del mismo proyecto. Este sistema se utilizó en la

ejecución del proyecto COGNITERM como herramienta de gestión del conocimiento, pero desde entonces ha sufrido una evolución hasta llegar a la aplicación actual llamada DOCKMAN, que además de la gestión del conocimiento también incorpora el análisis de corpus. Esta aplicación tiene las siguientes características: i) ofrece una representación jerárquica de los sistemas de conceptos, basados en la tecnología de ingeniería del conocimiento; ii) es una aplicación basada en hipertexto que permite la navegación por la base de conocimiento; iii) implementa la herencia automática de las características conceptuales desde los conceptos más genéricos a los más específicos; iv) ofrece la posibilidad de comparar y contrastar conceptos relacionados; y v) tiene la capacidad de mostrar la multidimensionalidad de las estructuras conceptuales (Skuce y Lethbridge, 1995). Además, esta última aplicación permite la identificación de términos, así como la identificación de los contextos ricos en conocimientos en los textos que forman el corpus.

De entre los objetivos que se plantearon en este proyecto, la representación de la multidimensionalidad, el dinamismo y las relaciones semánticas entre conceptos, la construcción de definiciones, entre otros, fueron unos de los puntos más estudiados. Las investigaciones llevadas a cabo en el seno de este proyecto dieron lugar a una serie de aspectos básicos a la hora de trabajar en un marco ontoterminográfico y que, a pesar de haber pasado tanto tiempo desde la finalización de este proyecto, se siguen teniendo en cuenta, como es el caso de la búsqueda de contextos ricos en conocimiento (*knowledge-rich contexts*) a la hora de buscar y extraer patrones lingüísticos (Meyer, 2001: 281),¹⁴¹ la relevancia de la multidimensionalidad en las representaciones conceptuales y sus dificultades a la hora de estructurar el conocimiento, la utilización de ontologías como base para el trabajo terminográfico, entre otros.

1.5.2. PROYECTO ONCOTERM

El proyecto ONCOTERM («Oncoterm: Sistema bilingüe de información y recursos oncológicos»), desarrollado en la Universidad de Granada bajo la dirección de la profesora P. Faber (<http://www.ugr.es/~oncoterm/>) trabajó en la misma línea de

¹⁴¹ «By *knowledge-rich context*, we designate a context indicating at least one item of domain knowledge that could be useful for conceptual analysis. In other words, the context should indicate at least one conceptual characteristic, whether it be an attribute or a relation» (Meyer, 2001: 281).

investigación que el proyecto anterior, aunque su finalidad y su aproximación fueron un poco diferentes. El objetivo global del proyecto consistió en la elaboración de una base de conocimiento terminológica bilingüe (inglés-español) del subdominio médico de la oncología con objeto de proporcionar recursos de información para médicos, pacientes, traductores y redactores de textos especializados. Esta base de conocimiento contaba en el momento de cierre del proyecto (año 2001) con 1.896 conceptos relacionados con el cáncer y 4.033 términos en inglés y español más generales (formas completas, sinónimos, siglas, etc.). En Faber y Jiménez (2002), los investigadores de este proyecto hacen un repaso a los objetivos principales de la investigación y presentan los logros alcanzados.

Los objetivos operativos de este proyecto consistieron en la creación de un corpus de textos médicos tanto en inglés como en español para la extracción de información terminológica y semántica; en la elaboración de un inventario de relaciones conceptuales específicas del evento médico, en general, y del evento médico oncológico, en particular; en la elaboración de una base de datos terminológica basada en una ontología que representa el subdominio médico de la oncología, entre otros. Debido a la carencia de herramientas en ese momento para alcanzar estos objetivos, se decidió desarrollar una herramienta, el programa Ontoterm®, en el seno de este proyecto. Como veremos en el capítulo 3 (cf. 3.2.1.), esta aplicación informática permite la estructuración y el almacenamiento de grandes cantidades de conceptos, relaciones y propiedades, así como la introducción de la información terminológica necesaria, mediante una ontología vinculada a una base de datos terminológica.

Los fundamentos teóricos de este proyecto para la codificación del conocimiento especializado se basan principalmente en el modelo lexemático funcional (MLF), de M. Mingorance y la Semántica de Marcos, de Fillmore (López Rodríguez et ál., 2006: 229)¹⁴², cuya combinación aporta óptimos resultados a la hora de representar el

¹⁴² El MLF es una teoría de base léxica que facilita la representación de relaciones paradigmáticas (las que establece una unidad con otras de la misma categoría gramatical capaces de intercambiarse en un contexto lingüístico concreto) y sintagmáticas (las que establece una unidad con otras que aparecen en la misma oración) en el lenguaje especializado. Este modelo ofrece una metodología para la representación de conceptos a partir del análisis de definiciones terminográficas estructuradas sistemáticamente. En el eje paradigmático, los conceptos son estructurados en una jerarquía de dominios y subdominios, partiendo de la premisa de que los miembros de un mismo dominio conceptual tienen una serie de propiedades en común. Por su parte, en el eje sintagmático se codifica el potencial combinatorio del concepto, es decir, se

conocimiento del dominio de la oncología de forma clara, concisa y coherente y, por tanto, al momento de elaborar definiciones terminográficas consistentes de forma sistemática.

Este proyecto ha realizado varias contribuciones que han adquirido gran importancia en el ámbito de la ontoterminografía.¹⁴³ Entre ellas, podemos destacar la especificación de esquemas categoriales (Faber et ál. 2001: 169) en eventos concretos (en este caso, el evento médico), que reflejan la organización interna de cada una de las categorías del evento en cuestión, a fin de garantizar la consistencia de las definiciones terminográficas, de la construcción de la ontología y de la base de datos terminográficas. Asimismo, se consideran también importantes otras contribuciones como el hecho de incluir de forma sistemática información pragmática de utilidad para el médico y el paciente en las entradas terminológicas; la elaboración de un inventario de relaciones conceptuales específicas del evento médico en general y del evento médico oncológico en particular; la introducción de imágenes y dibujos para mejorar la comprensión del conocimiento, entre otros.

Esta misma metodología se ha aplicado también en los proyectos PUERTOTERM: *Ingeniería de puertos y costas: estructuración del conocimiento y generación de recursos terminológicos* (<http://manila.ugr.es/puertoterm/>), cuyo objetivo principal es la representación de la estructura de conocimiento que subyace al dominio de la ingeniería de costas mediante una organización dinámica de sus conceptos; MARCOCOSTA: *Marcos de conocimiento multilingüe para la gestión integrada de las zonas costeras*, que se trata de una continuación del proyecto anterior y que tiene por objetivos generar herramientas y recursos terminológicos reutilizable e intercambiables en este dominio e incrementar el conocimiento sobre la dinámica de las zonas costeras, entre otros, y,

especifican las concordancias y colocaciones que pueden tener los términos seleccionados, así como las relaciones que permiten codificar grados, relación instrumental, causal, entre otras. Según esta teoría, la base de una estructura conceptual es la convergencia de estos dos ejes, por lo que el lexicón mental se concibe como una red compleja en la que cada nodo es un concepto, y los conceptos están interconectados por diferentes tipos de relaciones. Estas redes semánticas formadas por las interconexiones entre conceptos se dividen en dominios; estos, en marcos, y los marcos, a su vez, en conceptos interrelacionados. Este es el motivo por el que en el proyecto ONCOTERM se combina el modelo MLF con los principios de la Semántica de Marcos (cf. 1.5.2.).

¹⁴³ Los propios investigadores denominan su metodología como *gestión de terminología orientada a procesos* (Faber et ál., 2005).

finalmente, el proyecto ECOSISTEMA: *Espacio Único de Sistemas de Información Ontológica y Tesauro sobre el Medio Ambiente* Finalmente (en fase de ejecución), en coordinación con el grupo de investigación HUM-106 «Lexicografía y Traducción» (dirigido por la Dra. G. Corpas Pastor), que tiene como objetivos la elaboración de tesauros y recursos ontoterminográficos multilingües sobre el medio ambiente. Estos tres proyectos se encuentran bajo la dirección de la profesora P. Faber en la Universidad de Granada, a excepción del último que es coordinado con la Universidad de Málaga, y todos siguen la misma metodología del proyecto ONCOTERM.

1.5.3. PROYECTO GENOMA-KB

El proyecto GENOMA-KB, desarrollado en el Instituto de Lingüística Aplicada (IULA) de la Universidad Pompeu Fabra bajo la dirección de la profesora M. T. Cabré (<http://genoma.iula.upf.edu:8080/genoma/index.jsp>), se trata de otro proyecto que ha llevado a cabo la elaboración de una base de conocimiento terminológico trilingüe (inglés-español-catalán), en este caso sobre el genoma humano. Esta base de datos es de libre acceso a través de Internet y está constituida por una plataforma modular que integra a la vez unidades terminológicas, su estructuración conceptual, los textos reales de los que proceden y las referencias de dichas fuentes. Según la investigadora principal del proyecto (Cabré Castellví, 2006: 100), esta base de conocimiento está construida sobre la base de los cuatro módulos siguientes, que son independientes pero están interrelacionados:

- Banco textual, donde se alberga el corpus textual formado por los textos seleccionados sobre el tema del genoma humano;
- Banco documental, donde se recogen las referencias de las fuentes originales de donde se han recopilado los textos del banco textual, así como los diccionarios u otros recursos utilizados (en su caso) y otra información de interés (número de palabras, número de textos, clasificación de los textos, etc.);
- Banco terminológico, que contiene las fichas terminológicas de los términos seleccionados para el recurso; y

- Ontología, donde se representa la estructura conceptual del dominio

Los dos últimos módulos, es decir, la ontología y la base de datos terminológica se construyeron en paralelo con el programa Ontoterm®, por lo que se partió de la ontología como base para ir introduciendo paulatinamente la información terminológica a los términos extraídos del corpus textual. Por el contrario, los otros dos módulos no se pudieron trabajar con esta herramienta y, por ello, se tuvieron que emplear otras aplicaciones. En el caso del módulo documental, se utilizó MS-Access 2000; y en el caso del corpus, se empleó un programa informático, BwanaNet tool, disponible en IULA, para llevar a cabo el análisis del corpus textual compilado y realizar un etiquetado morfosintáctico a los textos.

Este proyecto aporta una nueva concepción de las bases de conocimiento, ya que integra en una misma aplicación la información terminológica junto con la información conceptual, bibliográfica y textual. Con ello, se consigue un recurso ontoterminográfico muy completo y versátil, que puede cubrir las necesidades de una amplia variedad de usuarios (traductores, redactores, terminógrafos, especialistas, etc.).

1.5.4. PROYECTO PROLEX

El proyecto PROLEX (Maurel, 2008), realizado por el laboratorio de informática de la Universidad François-Rabelais de Tours y dirigido por el profesor Maurel, tiene como objetivo la creación de una base de conocimiento multilingüe de nombres propios, denominada *Prolexbase* (<http://www.cnrtl.fr/lexiques/prolex/>) para el Procesamiento del Lenguaje Natural (PLN) en traducción automática, traducción asistida por ordenador, recuperación de información y diccionarios de pronunciación. Esta herramienta está formada por dos módulos interrelacionados: por un lado, el módulo que contiene información lingüística sobre los nombres propios y sus variantes; y por otro, la ontología que incluye diferentes relaciones jerárquicas y no jerárquicas entre los nombres propios.

Este proyecto se inició en la década de los noventa para procesar nombres propios, primero con el estudio de topónimos en francés y, posteriormente, con la implementación de una versión en serbio. A continuación, se creó el diccionario multilingüe relacional de nombres propios, *Prolexbase*, desarrollado en una base de

datos relacional, que, a partir de mediados de 2007, se encuentra disponible con libre acceso en formato .XML. Actualmente, se está trabajando en una nueva versión de este recurso utilizando los estándares ISO para facilitar la reutilización y el intercambio de contenido terminológico y ontológico.

El modelo Prolex está basado en dos aspectos principales, relacionados con los dos módulos en los que está dividida la base de conocimiento: por un lado, un idioma pivote independiente, para realizar la ontología en el nivel conceptual; y por otro, el prolexema (*prolexeme*), la proyección del idioma pivote en una lengua particular. La variación lingüística es un aspecto muy importante para el PLN, puesto que un mismo nombre propio puede escribirse de diferentes maneras, puede reemplazarse por otro nombre, etc. En el primer caso, el pivote es un punto de vista diacrónico, diafásico y diastrático del referente, es decir, una variación independiente de la lengua de un nombre. Estos conceptos pivotes están relacionados entre sí por tres tipos de relaciones semánticas: cuasisinónimos, partitivas y asociativas. Por su parte, el prolexema es un conjunto de variantes (forma completa, forma corta, acrónimo, variante ortográfico, transcripción, etc.), cuasisinónimos y derivados morfosintácticos que están relacionados entre sí (nombre relacionado, adjetivo relacionado, etc.).

En la actualidad, *Prolexbase* es una base de conocimiento que contiene básicamente nombres propios en francés y derivados, aunque también algunas traducciones a otros idiomas, principalmente serbio. La parte francesa de la base de datos contiene 54.774 lemas, 730 variantes y 20.614 derivados. Además, dispone de 50.567 relaciones, de las que 2.249 son asociativas, 47.670 son partitivas y 648 son cuasisinónimos.

En general, este proyecto es muy importante, puesto que los nombres propios suelen ser objeto de controversia a la hora de traducirlos o de utilizarlos, debido a la falta de reglas para su traducción. Por ello, esta base de conocimiento es un paso adelante para homogeneizar el uso de estos nombres, tanto entre personas como entre aplicaciones del PLN. El inconveniente que presenta actualmente es la carencia de traducciones a más idiomas, puesto que la mayor parte de estos nombres propios se encuentran en francés.

CAPÍTULO 2.

EL TRABAJO BASADO EN CORPUS DIRIGIDO A TRADUCTORES

ÍNDICE DE CONTENIDOS

- 2.1. Introducción
- 2.2. La traducción de textos especializados y la competencia traductora
- 2.3. La competencia documental en traducción
 - 2.3.1. Las fuentes de información para traductores
- 2.4. Las fuentes de información terminológicas
 - 2.4.1. Clasificación de los recursos terminológicos
 - 2.4.1.1. Según la presentación
 - 2.4.1.2. Según el número de lenguas
 - 2.4.1.3. Según el tipo de soporte
 - 2.4.1.4. Según la ordenación de sus contenidos
 - 2.4.1.5. Según la normalización
 - 2.4.1.6. Según el grupo de destinatarios
 - 2.4.1.7. Según el tipo de unidades terminológicas
 - 2.4.1.8. Según la extensión en la especialidad
 - 2.4.1.9. Según la fuente de información utilizada
 - 2.4.1.10. Según la finalidad
 - 2.4.2. Evaluación de fuentes de información terminológica
 - 2.4.2.1. Plantilla de evaluación. Parámetros e indicadores
- 2.5. Expectativas y necesidades de consulta de traductores profesionales
 - 2.5.1. Estudios previos realizados
 - 2.5.2. Presentación de la encuesta y su elaboración
 - 2.5.3. Resultados de la encuesta y conclusiones
 - 2.5.4. Algunos trabajos terminográficos dirigidos a traductores
- 2.6. Metodologías para el trabajo ontoterminográfico sistemático basado en corpus
 - 2.6.1. Propuesta de metodología para el trabajo ontoterminográfico sistemático basado en corpus dirigido a traductores
 - 2.6.1.1. Fase I: Delimitación del trabajo y análisis del dominio
 - 2.6.1.2. Fase II: Compilación del corpus especializado
 - 2.6.1.2.1. Diseño del corpus
 - 2.6.1.2.2. Búsqueda documental
 - 2.6.1.2.3. Organización de textos
 - 2.6.1.3. Fase III: Gestión ontoterminográfica
 - 2.6.1.3.1. Extracción de candidatos a término y validación
 - 2.6.1.3.2. Búsqueda de equivalentes de traducción
 - 2.6.1.3.3. Extracción de relaciones semánticas
 - 2.6.1.3.4. Elaboración de la ontología de dominio
 - 2.6.1.4. Fase IV: Elaboración de una base de datos ontoterminográfica
 - 2.6.1.4.1. Selección de campos
 - 2.6.1.4.2. Elaboración de definiciones
 - 2.6.1.4.3. Selección de ejemplos contextuales
 - 2.6.1.5. Fase V: Revisión y validación
 - 2.6.1.6. Fase VI: Exportación y edición

Ante el alud creciente de diccionarios de toda clase que se van publicando, unos excelentes o buenos, otros mediocres y hasta malos, una actitud crítica es más necesaria que nunca (Haensch, 1997: 237).

2.1. INTRODUCCIÓN

Como hemos visto en el capítulo anterior, la Terminología moderna se aleja de la corriente terminológica tradicional, tanto en lo que respecta a sus fundamentos teóricos como a su metodología, y se aproxima cada vez más a lo que se ha venido denominando la Lexicografía especializada. Asimismo, defiende la necesidad de orientar los trabajos terminográficos al usuario y a sus necesidades, a fin de satisfacer los requisitos específicos de cada grupo de destinatarios. De esta manera, los resultados del trabajo terminográfico dependerán siempre de los usuarios a los que vaya dirigido el recurso terminológico que se elabore.

En este contexto, queda patente que los recursos terminográficos que se empleen por traductores profesionales deberían ir precedidos de un estudio de necesidades o, al menos, de una consulta a estudios previos realizados sobre estas necesidades. Esta realidad, que en la mayoría de los trabajos terminográficos se asume de forma automática, no se cumple, sin embargo, en lo referente a los traductores profesionales;

de esta manera, encontramos multitud de recursos dirigidos a estudiantes de lenguas, a especialistas, a redactores, etc. pero no a traductores profesionales. Por este motivo, dedicamos este capítulo a realizar un seguimiento del proceso traductor y de las competencias traductoras de los profesionales, de las diferentes fuentes de información disponibles, haciendo mayor hincapié en las fuentes terminológicas, para pasar a tratar la necesidad de evaluar y seleccionar adecuadamente los recursos terminográficos, así como los diferentes recursos existentes dirigidos en parte a los traductores profesionales y que cumplen, en cierta manera, las necesidades y expectativas de este grupo de destinatarios según las conclusiones extraídas del estudio realizado mediante una encuesta. Por último, presentamos nuestra propuesta metodológica para elaborar recursos terminográficos dirigidos a traductores según las nuevas líneas de investigación de la Terminología moderna.

Debemos destacar que en este capítulo se entrelazan la Traductología, que nos permite establecer las características teóricas y prácticas de la traducción de textos especializados, así como las principales competencias traductoras de los profesionales, con la Documentación, que nos proporciona los conocimientos necesarios para realizar las búsquedas de estos recursos y la clasificación de las fuentes de información en general, así como con la *metaterminografía*,²⁷⁷ que nos permite realizar una investigación empírica sobre los diferentes tipos de recursos terminográficos existentes y elaborar una plantilla de evaluación para dichos recursos. Finalmente, la Traductología y la *metaterminografía* se vuelven a combinar al realizar el estudio mediante una encuesta sobre las necesidades de los traductores profesionales con

²⁷⁷ El término «metaterminografía» está aún por acuñar por los expertos en Terminografía. Hasta el momento se ha empleado en algunos trabajos (Gómez González-Jover, 2006: 691; Casas Gómez, 2006: 27), aunque siempre con referencia a las decisiones previas que determinar el modelo de recurso terminográfico que se va a compilar. En nuestro caso, hemos considerado necesario utilizar este término como equivalente en Terminografía al término «metalexicografía», que hace referencia a una rama de la Lexicografía que tiene como objeto de estudio el diccionario y su objetivo es investigar sobre él y no producirlos. En la metalexicografía, se diferencian cinco campos diferentes de actuación según Hartmann (2001:30), a partir de los propuestos por Wiegand (1998: 79-80), a saber: a) historia de los diccionarios (*Dictionary History*), b) crítica de los diccionarios (*Dictionary Criticism*), c) estructura de los diccionarios (*Dictionary Structure*), d) tipología de los diccionarios (*Dictionary Tipology*) y e) uso de los diccionarios (*Dictionary Use*). En nuestro trabajo, defendemos que esta clasificación establecida en Lexicografía se extrapole a la Terminografía y que, por tanto, se emplee el término *Metaterminografía* por analogía a *Metalexicografía* y se determinen estos mismos campos con relación a los recursos terminográficos, es decir, a los diccionarios especializados en general, incluyendo vocabularios, léxicos, bases de datos, etc.

respecto a los recursos terminográficos, siguiendo la rama aplicada de la traductología que, según Holmes (1972 [1988]),²⁷⁸ consiste en el estudio de herramientas que utiliza el traductor (entre otras cosas).

Mediante la aplicación de la *metaterminografía* en nuestra investigación, y siguiendo los cinco campos en los que se debería dividir esta rama de la Terminografía (cf. Nota a pie 1), realizaremos el siguiente estudio: en primer lugar, estableceremos una tipología de fuentes de información, así como una clasificación más específica de fuentes terminológicas, que incluirá diccionarios especializados, bases de datos, corpus especializados, etc. (cf. 2.4.1.); en segundo lugar, nos situaremos en el campo de la crítica de los diccionarios a la hora de establecer una plantilla de evaluación y una metodología para elaborar reseñas de estos recursos (cf. 2.4.2.); en tercer lugar, nos centraremos en la estructura de diccionarios, para analizar los criterios y parámetros pertinentes para la plantilla de evaluación (cf. 2.4.2.1.); y por último, estudiaremos el uso de los recursos terminográficos (incluyendo los diccionarios) durante el estudio mediante encuesta realizado a los traductores profesionales (cf. 2.5.).

Con todo esto, pretendemos principalmente conocer los tipos de fuentes de información que existen en la actualidad, así como las necesidades que presentan los traductores profesionales en general, con objeto de alcanzar un mayor conocimiento sobre cómo debería ser una herramienta adecuada para que cubra los vacíos que existen actualmente de recursos para estos traductores y proponer una metodología de trabajo para la elaboración de recursos que satisfagan sus necesidades. Así pues, en este capítulo, nos dedicaremos a tratar el análisis del proceso de traducción, los recursos disponibles, las necesidades de los traductores profesionales de forma general y diferentes metodologías propuestas hasta el momento, a fin de obtener una visión generalista y fuera de cualquier influencia de especialidad, para centrarnos en un dominio de especialidad, más concretamente el turismo en general y el turismo de aventura en particular, y en unas lenguas de trabajo determinadas (alemán, inglés y español) en el siguiente capítulo.

²⁷⁸ Este trabajo, «The Name and Nature of Translation Studies», fue publicado por primera vez en el Third International Congress of Applied Linguistics, celebrado en Copenhague en 1972, pero aquí citamos la edición posterior recogida en Holmes (1988).

2.2. LA TRADUCCIÓN DE TEXTOS ESPECIALIZADOS Y LA COMPETENCIA TRADUCTORA

La comunicación especializada y, por ende, los lenguajes de especialidad²⁷⁹ se materializan expresamente en unas producciones lingüísticas que se denominan «textos especializados». Por texto especializado se entiende, pues, la producción comunicativa que se ha generado dentro de alguna disciplina científica (experimental, exacta, humanística, etc.), de alguna materia técnica (ingeniería, construcción, comunicación, etc.) o de algún campo específico de actividad (deporte, comercio, finanza, etc.) (Recorder y Cid, 2004: 79) y que presenta elementos diferenciadores respecto de los textos generados dentro de la lengua común, es decir, textos producidos en situaciones no profesionales y dirigidos a un público más amplio.

Estos textos especializados son el objeto de trabajo de la traducción de textos especializados,²⁸⁰ que según Hurtado Albir (2001: 59) se definiría como «la traducción de textos dirigidos a especialistas y pertenecientes a los llamados lenguajes de especialidad». Por su parte, para Cabré Castellví (1993: 139) la traducción de textos especializados hace referencia a «cualquier comunicación realizada en el seno de los lenguajes de especialidad». En este sentido, esta última autora amplía el rango de aplicación de este tipo de traducción con respecto a la primera autora, puesto que no la limita a la comunicación entre especialistas, que sería una comunicación altamente especializada, sino que también incluye la comunicación entre especialistas y no especialistas, que hace referencia a la comunicación semi-especializada o divulgativa (cf. 1.2.2.2.). En nuestro caso, seguiremos la definición realizada por Cabré Castellví, puesto que consideramos que la que nos ofrece Hurtado Albir es en cierta manera reduccionista y no se ajusta totalmente a las pretensiones de la Terminología moderna.

²⁷⁹ Para cualquier ampliación o concreción acerca de los lenguajes de especialidad y sus características, remitimos al primer capítulo (cf. 1.2.), donde se tratan en profundidad.

²⁸⁰ Hurtado Albir (2001: 59-62) defiende el término «traducción de textos especializados» frente a «traducción especializada», ya que considera que cualquier tipo de traducción es en cierta manera especializada y precisará también, por tanto, de habilidades y conocimientos específicos. En nuestro caso, emplearemos también esta terminología siguiendo la recomendación de esta autora.

Como decíamos, la traducción de textos especializados está relacionada con los lenguajes especializados y, por ello, está determinada por las características propias de este tipo de traducción. Así pues, se pueden distinguir varias particularidades (Merlo Vega, 2004: 310) comunes a ambas realidades, que exponemos a continuación:

En primer lugar, es un lenguaje mayoritariamente exento de figuras retóricas, ya que utiliza un estilo directo y preciso, sin apenas recursos estilísticos,²⁸¹ por lo que la traducción debe seguir los mismos parámetros y mantener este tipo de estilo. En segundo lugar, el contenido temático pertenece a una disciplina concreta, lo que obliga al traductor a manejar o, por lo menos, conocer dicha disciplina. En tercer lugar, cuenta con una terminología concreta, la cual se debe conocer para realizar una buena traducción. En cuarto lugar, los destinatarios suelen ser doctos en la materia, profesionales o personas interesadas en el tema, por lo que las traducciones deben ser precisas y ajustarse a la terminología que se utiliza en el texto origen. Y, por último, la tipología textual puede variar dentro de un mismo ámbito de conocimiento, por lo que se debe prestar gran atención al género textual que se está traduciendo en cada momento.

En esta línea, podemos determinar que el traductor actúa como mediador en una comunicación especializada entre dos interlocutores, de los cuales el emisor, al menos, es especialista en la materia que se está tratando. En este contexto, el traductor se pone en el lugar de quien emite el mensaje, es decir, se convierte en un segundo especialista, y asume las mismas competencias que el emisor inicial, tanto en lo que se refiere a conocimientos sobre el tema de especialidad que se trata y su especificidad cognitiva, como a habilidades adecuadas y precisas de expresión, de convenciones textuales, de terminología, etc. Así pues, para conseguir que el texto traducido sea, en relación con el original, literal en cuanto a contenido, gramaticalmente correcto, adecuado en sus convenciones y ajustado estilísticamente, el mediador deberá acercarse lo máximo posible a la terminología que habría seleccionado el autor del texto origen si se hubiese expresado en la lengua meta (Cabré Castellví, 2000b: 2).

²⁸¹ En el lenguaje turístico, esta peculiaridad no se cumple siempre, ya que es un lenguaje que, por lo general, busca la belleza estilística y utiliza abundantes recursos retóricos, como veremos en el siguiente capítulo (cf. 3.3.4.1.).

Teniendo en cuenta la postura mediadora y de segundo autor especializado que ostenta la figura del traductor, podemos distinguir los conocimientos imprescindibles con los que debe contar a la hora de enfrentarse a un texto especializado: por un lado, debe tener conocimientos temáticos sobre la materia que esté traduciendo para ser capaz de comprender y de trasladar los conceptos del texto original al texto meta; y por otro lado, el traductor debe disponer también de conocimientos terminológicos, es decir, de la terminología propia del campo de especialidad en el que esté traduciendo, y de los géneros característicos del ámbito que está traduciendo, con objeto de conocer el funcionamiento de cada uno de los géneros en cuanto a convenciones lingüísticas y textuales (Hurtado Albir, 2001: 61). Además de estos conocimientos fundamentales, la capacidad para documentarse ocupa un lugar primordial en el proceso de traducción, ya que le permite al traductor adquirir conocimientos sobre el campo temático, la terminología y los géneros textuales y, así, salvar los posibles problemas de traducción en ese aspecto (Gamero, 1998: 100). De esta manera, podemos concluir que la traducción de textos especializados en general requiere del traductor profesional habilidades y destrezas particulares que se concretan en una serie de competencias específicas, como veremos a continuación, a fin de obtener óptimos resultados en la traducción de este tipo de textos.

Estas competencias específicas se pueden englobar en lo que se denomina «competencia traductora», que se puede definir de la siguiente manera:

[...] Conjunto de capacidades, destrezas, conocimientos e incluso actitudes que reúnen los traductores profesionales y que intervienen en la traducción como actividad experta, es decir, que, en su conjunto, distinguen al profesional del no profesional, al experto del no experto, o del «native translator» [traductor nativo] de Toury (Kelly, 2002: 9).

Los estudios que han tratado esta competencia traductora en un ámbito profesional se han ido sucediendo a lo largo de las últimas décadas (Wills, 1976; Delisle, 1980; Roberts, 1984; Nord, 1991; Pym, 1992; Gile, 1995; Hurtado Albir, 1996; Hatim y Mason, 1997; Neubert, 2000; Kelly, 2002; PACTE 2003, 2005, entre otros)²⁸² y han ido

²⁸² En Kelly (2002), se realiza un estudio sobre las diferentes propuestas realizadas hasta ese momento y una comparación entre ellas como preámbulo para proponer una clasificación nueva y diferente.

aportando diferentes propuestas, así como diferentes denominaciones para referirse a las subcompetencias que integran esta competencia traductora.

Actualmente, disponemos además de una norma oficial para tratar de regular y otorgar calidad a los servicios de traducción, la *Norma de Calidad Europea para los Servicios de Traducción* (UNE EN-15038: 2006),²⁸³ publicada en mayo de 2006 por el Comité Europeo de Normalización (CEN), que recoge una serie de competencias con las que debe contar el traductor para ser considerado profesional y, por tanto, para alcanzar unos buenos resultados finales en su traducción.

Estas competencias recogidas en la *Norma* son las siguientes: competencia traductológica, competencia lingüístico-textual en la lengua origen y en la lengua meta, competencia heurístico-documental, competencia cultural y competencia técnica.²⁸⁴

Por competencia traductológica, entendemos la capacidad que tiene el traductor de trasponer un texto de una lengua origen a un texto de una lengua meta en el registro necesario y de acuerdo con los criterios acordados en el encargo del cliente. En otras palabras, se trata de la capacidad de comprender el texto original y reexpresarlo en la lengua de llegada según la finalidad de la traducción y las características del destinatario (HurTado Albir, 2001: 395), manteniendo las especificaciones realizadas por el cliente.

La competencia lingüístico-textual consiste en el conocimiento de los dos sistemas lingüísticos en que se está trabajando, es decir, la lengua origen y la lengua meta, además de conocer el funcionamiento y las diferencias existentes en el discurso entre ambas lenguas y sus convenciones textuales, así como la terminología específica del ámbito de especialidad. Según Hurtado Albir (2001: 395), está compuesta por los siguientes conocimientos y habilidades: 1) competencia gramatical, es decir, el conocimiento morfosintáctico y léxico de la lengua meta; 2) competencia textual, que consiste en el conocimiento de las convenciones textuales de los diferentes géneros así

²⁸³ <<http://www.statsaut-translator.no/Files/Standard-15038-final-draft-en.pdf>>

²⁸⁴ La traducción al español de las competencias se han sacado de Corpas Pastor (2004: 138). Los nombres originales en inglés de dichas competencias son los siguientes: «translating competence», «linguistic and textual competence in the source language and the target language», «research competence», «information acquisition and processing», «cultural competence» y «technical competence», respectivamente.

como la coherencia y cohesión del texto; 3) competencia sociolingüística o pragmática, relacionada con el conocimiento de las reglas socioculturales de uso como normas de cortesía, relevancia o propiedad en distintos repertorios relacionada con la producción y comprensión apropiados en diversos contextos sociolingüísticos.

La competencia heurístico-documental es una de las competencias más importantes de la labor traductológica, ya que a través de ella el traductor es capaz de adquirir los conocimientos necesarios para llevar a cabo la traducción y, así, resolver cualquier problema conceptual o terminológico que se le presente de forma eficaz y rápida durante la comprensión del texto origen y la reproducción del texto meta. El traductor debe conocer las fuentes de información que tiene a su disposición y, a la vez, saber utilizarlas para seleccionar la información pertinente en cada caso. Según Hurtado Albir (2001: 62), «la capacidad para documentarse ocupa un lugar central en el conjunto de competencias, ya que permite al traductor adquirir conocimientos sobre el campo temático, sobre la terminología y sobre las normas de funcionamiento textual del género en cuestión».

La siguiente competencia que recoge la *Norma* es la competencia cultural, que hace referencia a los conocimientos culturales que debe poseer el traductor para comprender un texto origen y poder adaptarlo correctamente a la cultura meta. El traductor debe ser consciente de las divergencias que se pueden dar con respecto a la estructura de los tipos textuales concretos dentro de un mismo género y de las normas que gobiernan la producción de textos científicos tanto en la comunidad lingüístico-cultural de origen como en la de llegada. De la misma manera, debe realizar los cambios y reajustes necesarios en el contenido y en el formato del texto original para conseguir un texto meta que funcione adecuadamente en la cultura meta, siempre con respecto a la finalidad de la traducción. Asimismo, es imprescindible que el traductor conozca las posibles divergencias culturales que pueden existir entre una cultura y otra, ya sea con relación a las connotaciones que pueden llevar asociadas algunas unidades, la necesidad de ampliar la información que aparece en el texto origen con respecto a alguna realidad que no exista o sea diferente en la cultura meta, etc. En definitiva, un conocimiento

amplio sobre las dos culturas implicadas en la traducción, a modo de evitar cualquier incompreensión, falta de entendimiento o de comunicación.

Por último, nos encontramos con la competencia técnica. La *Norma* considera esta competencia como la capacidad de preparar y realizar traducciones de forma profesional, lo cual supone, por ejemplo, el manejo de entornos informáticos y de fuentes de información terminográfica (Corpas Pastor, 2004: 138). Esta competencia sería lo que Hurtado Albir denomina competencia profesional o instrumental (Hurtado Albir, 2001: 396), que incluye los conocimientos y habilidades relacionados con el ejercicio de la traducción profesional, esto es, el conocimiento y uso de las fuentes documentales de todo tipo, de las nuevas tecnologías aplicadas a la traducción así como del funcionamiento del mercado laboral (búsqueda de trabajo, contacto con el cliente, *marketing*, etc.).

Como veíamos anteriormente, además de esta propuesta de subcompetencias traductoras, que se puede considerar actualmente como la oficial y, por tanto, determinante, existen otras propuestas similares que incluyen más o menos competencias, así como denominaciones diferentes. A continuación, trataremos también la propuesta de competencias traductoras que ha realizado el grupo PACTE, de la Universidad Autónoma de Barcelona, la cual ha sido considerada como una de las propuestas más completas y de las más utilizadas.

A partir de los resultados obtenidos en una investigación empírico-experimental, el grupo PACTE²⁸⁵ ofrece una definición de lo que consideran la «competencia traductora», así como una propuesta de las competencias de esta competencia traductora (Hurtado Albir, 2001: 394-400; Beeby et ál. 2001: 39-45). Para ellos, la «competencia traductora» es el «sistema subyacente de conocimientos, habilidades y aptitudes necesarios para traducir» y es «cualitativamente distinta a la competencia bilingüe» (Hurtado Albir, 2001: 394). En una definición posterior (PACTE, 2003: 48), este grupo de investigación completa la definición de la siguiente manera: «Translation

²⁸⁵ El grupo de investigación PACTE está integrado en la actualidad por A.Beeby, L.Berenguer, D.Ensinguer, O.Fox, A. Hurtado Albir, N.Martínez Melis, W.Neunzig, M.Orozco y M.Presas; la investigadora principal es A. Hurtado Albir, y pertenecen al Departamento de Traducción e Interpretación de la Universitat Autònoma de Barcelona.

competence is the ability to carry out the transfer process from the comprehension of the source text to the reexpression of the target text, taking into account the purpose of the translation and the characteristics of the target-text readers». De esta forma, y a diferencia de lo que se hace en la *Norma*, consideran la competencia traductora como un todo que engloba diferentes aspectos, denominados subcompetencias, a saber: subcompetencia lingüística en las dos lenguas, subcompetencia extralingüística, subcompetencia profesional e instrumental, subcompetencia de transferencia, subcompetencia psicofisiológica y subcompetencia estratégica.

Las subcompetencias comunicativa, profesional y de transferencia equivaldrían a las competencias lingüístico-textual, técnica y traductológica de la *Norma*, respectivamente. Sin embargo, el grupo PACTE habla de la subcompetencia extralingüística y de la competencia estratégica, que no tendrían equivalentes en la *Norma*, si no que formarían parte de otras competencias. En su caso, la competencia extralingüística se podría incluir en la competencia cultural y la competencia estratégica, en la competencia traductológica o profesional. Por último, este grupo también incluye una subcompetencia que no se podría incluir en ninguna de las competencias de la *Norma*, que es la subcompetencia psicofisiológica, la cual hace referencia a la habilidad de aplicación de mecanismos psicomotores, cognitivos y actitudinales de todo tipo, como son las facultades cognitivas de atención y memoria, la curiosidad intelectual, la perseverancia, rigor, espíritu crítico y habilidades para desarrollar procesos de creatividad, razonamiento lógico, etc.

En definitiva, y sea cual sea la propuesta científica para clasificar las competencias traductorales, todos los modelos pretenden detectar los aspectos que influyen en el proceso traductor y dar una respuesta a cada uno de ellos mediante las habilidades y capacidades traductorales necesarias a la hora de traducir.

Después de haber expuesto las competencias traductorales, podemos decir que los pilares fundamentales de la traducción de textos especializados son la competencia lingüístico-textual, especialmente en lo que respecta a la terminología, y la competencia documental. La primera nos permite comprender un texto y transponerlo correctamente a la lengua meta y la segunda nos permite alcanzar el conocimiento necesario para

poder llevar a cabo la transposición de manera correcta y eficaz. No obstante, el resto de competencias se deben considerar esenciales también a la hora de proporcionar buenos resultados en la traducción profesional.

Como hemos indicado anteriormente, las competencias traductoras nos permiten hacer frente a los problemas de traducción y, así, evitar posibles errores traductológicos. Según Hurtado Albir (2001: 280), no existe todavía una definición de problema de traducción consensuado por los traductólogos y, por ello, podemos encontrar diferentes definiciones de este concepto. Asimismo, se produce a menudo una confusión entre «problema» y «dificultad», y se observa un empleo arbitrario generalizado de estos conceptos. En nuestra opinión, se debe realizar una distinción clara entre estos dos conceptos, ya que se tratan de dos realidades próximas pero diferentes entre sí. Siguiendo a Nord (1991: 151), una de los autores que aborda explícitamente la cuestión de los problemas de traducción, consideramos que un problema de traducción es un problema objetivo que todo traductor (independientemente de su nivel de competencia y de las condiciones técnicas de su trabajo) debe resolver en el transcurso de una tarea de traducción determinada, mientras que una dificultad de traducción es subjetiva y tiene que ver con el propio traductor y sus condiciones de trabajo particulares.²⁸⁶ De este modo, cuanto más preparado se encuentre un traductor, menos dificultades encontrará a la hora de traducir un texto. Sin embargo, un traductor muy especializado encontrará en un texto los mismos problemas (ya sean textuales, pragmáticos, culturales o lingüísticos) que un traductor lego o un estudiante.

Con respecto a la cuestión de problema de traducción, Hurtado Albir (2001: 288), nos ofrece una clasificación de estos problemas y los agrupa en cuatro categorías: problemas lingüísticos, problemas extralingüísticos, problemas instrumentales y problemas pragmáticos. Estos problemas de traducción se pueden encontrar en todo tipo

²⁸⁶ La definición original de «problema de traducción» que Nord propone en su libro (1991: 51) es la siguiente: «A translation problem is an objective problem which every translator (irrespective of his level of competence and of the technical conditions of his work) has to solve during a particular translation task». Para la definición de «dificultad de traducción», indica lo siguiente: «Translation difficulties, on the other hand, are subjective and have to do with the translator himself and his specific working conditions».

de traducción, ya sea general o especializada; no obstante, en la traducción de textos especializados, estos problemas se hacen más acuciantes y su presencia es aún mayor.

Según Hurtado Albir (ibíd.), los problemas lingüísticos son problemas de carácter normativo que recogen discrepancias en las dos lenguas en sus diferentes planos: léxico, morfosintáctico, estilístico y textual. Por su parte, los problemas extralingüísticos hacen referencia a cuestiones de tipo temático, cultural o enciclopédico. Los problemas instrumentales derivan de la dificultad en la documentación o en el uso de herramientas informáticas. Y, por último, los problemas pragmáticos están relacionados con los actos comunicativos presentes en el texto original, la intencionalidad del autor, las presuposiciones y las implicaturas, así como las cuestiones del encargo de traducción, las características de los destinatarios y del contexto en que se efectúa la traducción.

Como indicábamos al principio, estos problemas permanecerán inamovibles con independencia de la profesionalidad del traductor o de sus conocimientos, por tratarse de problemas objetivos y propios del texto objeto de traducción.

Por su parte, las dificultades de traducción también se pueden englobar en cuatro grupos diferentes, a saber: las dificultades específicas que presenta el texto, como puede ser su coherencia, cohesión, la claridad en la exposición (la estructura del tema y el rema), etc.; las dificultades relacionadas con el traductor, debidas principalmente a la falta de conocimientos culturales y lingüísticos de sus lenguas de trabajo; las dificultades pragmáticas, que están relacionadas con el objetivo de la traducción, como por ejemplo los receptores específicos, la función comunicativa, el canal de comunicación, etc.; y, por último, las dificultades técnicas, que vienen provocadas por la carencia de materiales adecuados de documentación, terminologías, programas de traducción, etc. y por la falta de conocimientos en el manejo de dichos materiales (Nord, 1991: 152 y ss.).

Relacionado con los problemas y las dificultades de traducción, nos encontramos con los errores de traducción, que Hurtado Albir (ibíd.: 289) los define como «una equivalencia inadecuada para la tarea traductora encomendada», es decir, como una ruptura de las reglas de coherencia de un texto meta, sean estas de tipo lingüístico

(gramatical, terminológico, de congruencia y precisión semántica, etc.) o extralingüísticas (cultural, funcional y de registro).

Actualmente se están realizando muchos estudios sobre la evaluación y la corrección de traducciones, tanto desde un punto de vista profesional como pedagógico, debido en parte a la aplicación de la *Norma de Calidad Europea para los Servicios de Traducción* (EN-15038).

A continuación, y después de haber tratado los aspectos más relevantes de la traducción de textos especializados, la competencia traductora necesaria para llevar a cabo una traducción con resultados favorables, así como los problemas y dificultades de la traducción en general, nos disponemos a profundizar en una de las competencias más importantes de la labor traductora, la competencia documental, y seguir, así, con el tema principal de nuestra investigación: el estudio y análisis de los recursos terminográficos para traductores y sus necesidades.

2.3. LA COMPETENCIA DOCUMENTAL EN TRADUCCIÓN

Hacer traducción de textos especializados supone conocer los elementos metodológicos y los recursos para resolver problemas de terminología planteados en la traducción (Cabré Castellví, 2004: 105).

La actividad traductora en el siglo XXI no se puede concebir sin el apoyo de la tecnología y, por tanto, la pericia del traductor se vuelve imprescindible para el manejo de las nuevas tecnologías aplicadas a la traducción. El trabajo artesanal que realizaba el traductor de antaño ha ido dando paso a un traductor profesional consciente de que el éxito de su trabajo radica, además de en el conocimiento profundo de sus lenguas de trabajo y de las culturas de estas, en su conocimiento tecnológico y documental.

En su trabajo, «Information Literacy: The Web is not an Encyclopaedia», Larsen (2001) nos introduce el concepto de «alfabetización informativa» (*information literacy*), que define como la capacidad de saber buscar la información necesaria, encontrarla,

evaluarla, procesarla y utilizarla para tomar decisiones adecuadas en la vida de cada uno. Precisamente, los traductores se enfrentan diariamente a estas tareas en su vida profesional y, por ello, esta «alfabetización informativa» (o competencia documental, según la norma EN-15038) se considera una competencia imprescindible y fundamental para un traductor profesional a la hora de traducir un texto de cualquier naturaleza. Esta competencia está presente durante todo el proceso traductor y puede llegar a ocupar, según algunos autores, el ochenta por ciento de este proceso (Königs, 1989: 126), siendo la traducción de textos especializados la que más recursos documentales requiere.²⁸⁷

En el apartado anterior, veíamos cómo la competencia documental es una de las competencias que incluye la nueva *Norma de Calidad Europea para los Servicios de Traducción* (EN-15038), la cual hace referencia a los conocimientos que posee el traductor para buscar información sobre la traducción que está realizando de forma eficaz y rápida y, así, adquirir los conocimientos necesarios en caso de que se encuentre con alguna dificultad conceptual o terminológica durante la comprensión del texto origen y la reproducción del texto meta. Para el traductor profesional de textos especializados es imprescindible esta competencia, ya que a menudo se encuentra con dificultades de traducción.

Según Palomares Perraut (2000: 15-16), el traductor acude a las fuentes documentales para adquirir básicamente los tres tipos de información siguientes:

- Información sobre el contenido del texto de origen que le permita comprender el asunto del que trata el texto objeto de traducción, así como su contexto y especificidad cultural, social o ideológica.

²⁸⁷ La tarea de documentación en la traducción de textos especializados es mucho mayor que en la traducción general debido principalmente a la gran cantidad de terminología que aparece en el texto que se debe traducir. Esto se debe principalmente al hecho de que la terminología juega un papel determinante en la comunicación especializada por tres razones: a) porque los especialistas utilizan la terminología para transferir el conocimiento de una materia en una o más lenguas; b) porque los términos concentran con mayor intensidad el conocimiento especializado; y c) porque la calidad de una traducción de textos especializados depende en gran medida de la correcta y adecuada utilización de la terminología en la lengua meta (Recorder y Cid, 2004: 82).

- Información terminológica que le permita utilizar con precisión el vocabulario especializado de esa materia.
- Información fraseológica que le permita usar el estilo propio del texto y transmitir su forma en la lengua de llegada.

Es decir, la competencia documental es una competencia transversal que se encuentra en todas las fases de traducción de un texto. En este sentido, nos permite preparar y comprender el texto origen (fase de pretraducción), elaborar las diferentes versiones del texto meta (fase de traducción) y redactar la versión definitiva del texto meta después de su correspondiente revisión y corrección lingüística, científica y traductológica (fase de revisión y control de calidad) (Corpas Pastor, 2001: 164).

Además de mantener esta estrecha relación entre el traductor y la documentación durante todo el proceso de traducción, el traductor desarrolla tres papeles diferentes: como usuario, como procesador y como documentalista (Gonzalo García, 2004: 275-276):

- En primer lugar, el traductor actúa como *usuario* de la documentación, experto en la búsqueda, selección y evaluación de la información que utiliza para poder solventar los diversos tipos de problemas durante el proceso de traducción. Para llevar a cabo esta tarea de documentación, el traductor debe contar con una competencia documental desarrollada, que le permita identificar, evaluar, manejar y hacer lo más rentable posible todas las fuentes de información de las que dispone.
- En segundo lugar, el traductor trabaja como *procesador* y *productor* de información al realizar sus traducciones en memorias de traducción, donde se guarda su trabajo de forma estructurada y organizada para sus futuros trabajos. Además, puede ir confeccionando sus propios repertorios terminográficos a la par que realiza las búsquedas correspondientes a los problemas terminológicos que se encuentra durante su trabajo.
- Y, por último, el traductor se convierte en *documentalista*, al tener que realizar el proceso de documentación. Este trabajo se ha visto asistido por la

revolución que se ha producido en la Red y, de esta manera, el traductor dispone a su alcance de comunidades virtuales, foros de discusión, listas de distribución o aulas virtuales, por citar solo algunos recursos disponibles, además de toda la información que tiene a su disposición a través de buscadores, metabuscadores, directorios, etc.²⁸⁸ En este sentido, el traductor no solo debe aprender el lenguaje de interrogación de Internet, sino que además tiene que conocer sus parámetros e indicadores de evaluación, a la hora de seleccionar la información fiable y adecuada para sus necesidades.

En el proceso de documentación, el traductor puede consultar diferentes tipos de fuentes de información, entendiendo por «fuente de información» «los recursos documentales creados para facilitar datos sobre una persona, institución, documento o asunto» (Merlo Vega, 2004: 312). A pesar del gran uso que se hace de la expresión «fuente de información», no existe a día de hoy una clasificación definitiva que compartan todos los especialistas en documentación,²⁸⁹ lo que hace difícil el tratamiento de estas fuentes.

Tradicionalmente, las fuentes de información han estado disponibles básicamente en bibliotecas especializadas. Sin embargo, gracias al gran desarrollo de Internet y a la Era de la Información, hoy en día se puede encontrar gran cantidad y variedad de fuentes de información en la Red y se pueden consultar de forma rápida y sencilla y con acceso inmediato. Este desarrollo que ha sufrido Internet en los últimos años ha motivado la aparición de nuevas fuentes de información y la continua revisión y clasificación de este tipo de fuentes, lo que ha provocado la falta de esta clasificación homogénea que aúne todas las clasificaciones realizadas hasta el día de hoy.

²⁸⁸ Cf. 2.3.1.

²⁸⁹ Algunas de las propuestas de clasificaciones que se pueden consultar actualmente se encuentran en Seghiri Domínguez (2006), Gonzalo García (2004 [2003] y 2004), Palomares Perraut (1999 y 2000), Pinto Molina y Cordón García (1998), Codina (2000a), Austermühl (2001), Martínez López *et ál.* (2001 [1997]), Benvenuti (2002a), Rotta (2002), Corpas Pastor (2003b), Palomares Perraut y Gómez Camarero (2004), Gonzalo García y Fraile Vicente (2004) y Merlo Vega (2004).

2.3.1. LAS FUENTES DE INFORMACIÓN PARA TRADUCTORES

Siguiendo las clasificaciones de fuentes de información propuestas por Gonzalo García (2004 [2003]) y Seghiri Domínguez (2006), podemos dividir las fuentes de información pertinentes para un traductor especializado en siete categorías diferentes, las cuales se subdividen a su vez en otras tantas. El objetivo de estas propuestas pretende clasificar únicamente recursos electrónicos; sin embargo, también se puede utilizar esta taxonomía para los recursos en papel o en otro tipo de formato, por ejemplo en CD-ROM.²⁹⁰

En un primer bloque, nos encontramos las fuentes telemáticas, que consisten en sistemas de búsqueda de información o localización de recursos en la red (Gonzalo García, 2004 [2003]). Estas fuentes están adquiriendo cada vez mayor peso, debido a la abundancia de información que se puede encontrar en Internet actualmente y al interés por agilizar las búsquedas realizadas en este medio. Dentro de esta categoría de fuentes, podemos distinguir básicamente tres tipos diferentes: los sistemas de búsqueda generales, los sistemas de búsqueda temáticos y otros sistemas de búsqueda. Para cada categoría, además, podemos encontrar varios tipos de sistemas de búsqueda, como veremos a continuación.

El primer grupo hace referencia a los sistemas de búsqueda que no están especializados en ningún campo determinado y son el primer intento para facilitar un proceso de búsqueda documental de forma rápida y sencilla. Dentro de este grupo, distinguimos los motores de búsqueda, los compiladores de buscadores y los metabuscadores.

Los motores de búsqueda, o también llamados buscadores, se consideran sistemas informáticos que indexan archivos almacenados en servidores web. Actualmente, hay gran variedad de buscadores disponibles para los usuarios de Internet

²⁹⁰ En nuestra investigación, trataremos con recursos electrónicos así como impresos y en CD-ROM, por lo que consideramos importante esta ampliación del uso de estas propuestas de clasificación.

que ofrecen sus resultados basados en diferentes criterios de búsqueda,²⁹¹ aunque todos pretenden el mismo fin, a saber: obtener el mayor número de resultados relacionados con la palabra clave o el tema de la búsqueda inicial. Como ejemplos de motores de búsquedas actuales, podemos hablar de *Google*,²⁹² el buscador más utilizado actualmente; *Alltheweb*,²⁹³ *Mooter*,²⁹⁴ *Yahoo*,²⁹⁵ *Fleiro*,²⁹⁶ etc. Otro tipo de motor de búsqueda que está adquiriendo importancia en la actualidad son los buscadores verticales, es decir, motores de búsqueda especializados en un determinado campo. Para nuestra investigación basada en el dominio del turismo,²⁹⁷ encontramos numerosos motores de búsquedas especializados en este sector, como por ejemplo *Masactividades.com*,²⁹⁸ *Toprural.com*,²⁹⁹ *Buscadorturismo.com*,³⁰⁰ *Spa Vacation*,³⁰¹ *Sportsuchmaschine*,³⁰² etc.

Dentro de los sistemas de búsqueda generales, existen también los denominados compiladores de buscadores, cuyo cometido consiste en agrupar y clasificar los motores de búsqueda recopilados, facilitando de este modo la localización de la información. Entre los compiladores más prominentes en la actualidad, encontramos *Search Engines Worldwide*,³⁰³ *Search Engine Colossus*³⁰⁴ y *Search Engine Watch*,³⁰⁵ entre otros.

²⁹¹ Véase Torres Pombert (2003) para una mayor profundización en el tema.

²⁹² <<http://www.google.com>>

²⁹³ <<http://www.alltheweb.com/>>

²⁹⁴ <<http://mooter.com/>>

²⁹⁵ <<http://www.search.yahoo.com>>

²⁹⁶ <<http://www.fleiro.de/>>

²⁹⁷ A partir del capítulo siguiente, se concretará el turismo, en general, y el turismo de aventura, en particular, como el dominio de especialidad en el que vamos a llevar a cabo nuestro estudio de campo. Por ello, en este apartado, ejemplificaremos algunas fuentes de información de este dominio en concreto.

²⁹⁸ <<http://www.masactividades.com/>>

²⁹⁹ <<http://www.toprural.com/activo/>>

³⁰⁰ <<http://www.buscadorturismo.com/>>

³⁰¹ <<http://www.spafinders.com/>>

³⁰² <<http://www.sportsuchmaschine.de>>

³⁰³ <<http://home.inter.net/takakuwa/search/search.html>>

³⁰⁴ <<http://www.searchenginecolossus.com/>>

³⁰⁵ <<http://searchenginewatch.com/links/>>

Otro tipo de sistemas de búsqueda general son los llamados metabuscadores, que son herramientas de búsqueda que actúan sobre los buscadores de información generales más conocidos. Hoy en día, existen numerosos metabuscadores como *Kartoo*,³⁰⁶ *Turbo10*,³⁰⁷ *Webcrawler*,³⁰⁸ *Dogpile*,³⁰⁹ *Yippy*,³¹⁰ citando solo unos cuantos.

En el segundo grupo de sistemas de búsqueda, nos encontramos con los sistemas de búsqueda temáticos, es decir, sistemas de búsqueda que organizan la información agrupándola en grupos o categorías que el usuario debe seleccionar para acceder a ella. Dentro de este grupo, se distinguen varios tipos de sistemas: los directorios, los anillos temáticos, las listas de distribución y foros de discusión y las redes temáticas.

Los directorios están formados por bases de datos estructuradas manualmente en áreas y subáreas temáticas y consisten en listas de enlaces a recursos de Internet relacionados con cada una de las áreas en las que están divididos. Como ejemplo de directorios generales, tenemos *Yahoo!*,³¹¹ *Flesko*,³¹² *Open Directory Project (ODP)*,³¹³ *Adirlink*,³¹⁴ entre otros.

Además de estos directorios generales, podemos encontrar directorios especializados según materias. En nuestro campo de especialidad, el turismo, se dispone de numerosos directorios de esta naturaleza en la actualidad, tanto de entidades privadas (*WorldTourismDirectory.com*,³¹⁵ *Directorio de Turismo en España*,³¹⁶ *Tourism Offices Worldwide Directory*³¹⁷, *Tourismus-Webkatalog.com*,³¹⁸ etc.) como de entidades

³⁰⁶ <<http://www.kartoo.com/>>

³⁰⁷ <<http://turbo10.com/>>

³⁰⁸ <<http://www.webcrawler.com/>>

³⁰⁹ <<http://www.dogpile.com/>>

³¹⁰ <<http://search.yippy.com/>>

³¹¹ <<http://es.search.yahoo.com/>>

³¹² <<http://www.flesko.es/>>

³¹³ <<http://dmoz.org>>

³¹⁴ <<http://www.adirlink.com/>>

³¹⁵ <<http://www.worldtourismdirectory.com/>>

³¹⁶ <http://www.sitiosespana.com/paginas/tur_dir.htm>

³¹⁷ <<http://www.towd.com/>>

públicas (*Utilidades*,³¹⁹ *Recursos en Turismo*,³²⁰ *Turismonet*,³²¹ etc.). Asimismo, se pueden consultar numerosos directorios dirigidos a profesionales de la lengua (traductores, intérpretes, terminólogos, etc.), que contienen gran cantidad de enlaces a recursos de diferente naturaleza (lingüísticos, técnicos, financieros, etc.). Como ejemplo de estos directorios, podemos citar los siguientes: *Gotranslator.com*,³²² *El Cuaderno de Bitácora*,³²³ *Übersetzer Verzeichnis*,³²⁴ etc.

Siguiendo con los sistemas de búsqueda temáticos, hallamos los anillos temáticos, o simplemente anillos, como se denomina a las colecciones de páginas o sitios web enlazadas entre sí y que tratan sobre un mismo tema, como puede ser el turismo, el arte, el deporte, etc. Actualmente, se pueden encontrar variedad de anillos sobre temas diferentes que, además, crecen día a día. Como ejemplos, podemos nombrar *Ringsurf*,³²⁵ uno de los principales portales sobre agrupación de enlaces; *El anillo*,³²⁶ una de las páginas de anillos web en español más visitadas; *WebRing*,³²⁷ que alberga más cantidad de información, entre otros. Dentro de este último ejemplo de anillo, podemos consultar otros anillos que tratan sobre ámbitos de la traducción y la lingüística, como *Translation on the Web*,³²⁸ *A Ring of Languages and Linguistics*,³²⁹ *The Foreign Language Teachers Web Ring*,³³⁰ *Languages*,³³¹ *Translators Webring*,³³²

³¹⁸ <<http://www.tourismus-webkatalog.com/>>

³¹⁹ <<http://www.aloj.us.es/gist/utilidades.htm>>

³²⁰ <http://biblioteca.uca.es/sbuca/bibcjer/recursos_tur.asp?capbd=4>

³²¹ <<http://bibemp2.us.es/turismobibemp/>>

³²² <<http://www.gotranslators.com/>>

³²³ <<http://www.elcuadernodebitacora.org/>>

³²⁴ <<http://www.uebersetzer-verzeichnis.com/>>

³²⁵ <<http://www.ringsurf.com>>

³²⁶ <<http://www.elanillo.com>>

³²⁷ <<http://dir.webring.com/rw>>

³²⁸ <<http://www.webring.com/t/Translation-on-the-Web>>

³²⁹ <<http://www.webring.com/hub?ring=lang>>

³³⁰ <<http://www.webring.com/hub?ring=flteachers>>

³³¹ <<http://www.webring.com/hub?ring=searchlanguage>>

³³² <<http://www.webring.com/hub?ring=translatorswebri>>

Translators and Interpreters,³³³ que pueden resultar realmente útiles a la hora de realizar cualquier tipo de traducción, ya sea general o especializada. Asimismo, fuera de este gran anillo temático, podemos encontrar otros sistemas de búsqueda similares que tratan sobre la traducción, como es el caso de *The Ring of the Translation Research*.³³⁴

Unos sistemas de búsqueda temática que están adquiriendo mayor importancia en la actualidad son las listas de distribución y los foros de discusión. En el campo de la traducción, especialmente entre los traductores profesionales, estos recursos se están convirtiendo en herramientas muy valiosas a la hora de consultar dudas de diferente naturaleza, buscar fuentes de referencias, preguntar por cursos de formación, ofrecer ofertas de trabajo, etc. y, poco a poco, van consiguiendo un puesto preferente entre estos profesionales de la lengua. Actualmente, se pueden encontrar numerosos foros y listas sobre cualquier tema, y lo mismo ocurre en el ámbito de la traducción. Con respecto a los foros de discusión, podemos consultar el foro *English Spanish Translators Org*,³³⁵ el foro de debates para alumnos de traducción del Centro Virtual Cervantes,³³⁶ *Terminologieforum*,³³⁷ *Proz.com*,³³⁸ etc. Y, con respecto a las listas de distribución, *Rediris*,³³⁹ *Yahoo*³⁴⁰ y *Google*³⁴¹ ofrecen listas variadas, sobre turismo, traducción, deportes, etc. cuya inscripción y consulta son completamente gratuitas. Asimismo, encontramos otras listas de distribución más específicas, como *Tradares*,³⁴² que se trata de una lista para traductores de alemán-español; *U-Forum*,³⁴³ que es una lista genérica

³³³ <<http://www.webring.com/hub?ring=translatorsandin>>

³³⁴ <<http://www.translationresearch.com/webring/>>

³³⁵ <<http://www.english-spanish-translator.org/foros-traductores-foro-traduccion.php>>

³³⁶ <http://cvc.cervantes.es/aula/el_atril/fet/>

³³⁷ <<http://www.terminologieforum.de/index.php>>

³³⁸ <<http://www.proz.com/forum>>

³³⁹ <<http://www.rediris.es/>>

³⁴⁰ <<http://groups.yahoo.com>>

³⁴¹ <<http://groups.google.com>>

³⁴² <www.elistas.net/lista/tradales>

³⁴³ <www.techwriter.de/thema/u-forum.htm>

para traductores de alemán; *Infotrad*,³⁴⁴ que informa sobre ofertas de trabajo, convocatorias, reuniones, etc. relativas al mundo de la traducción, entre otras.

Para terminar con los sistemas de búsqueda temática, debemos hablar de las redes temáticas. Estos sistemas de búsqueda son relativamente recientes y su objetivo consiste básicamente en proporcionar una plataforma donde los profesionales de un mismo campo puedan interactuar e interrelacionarse a través de Internet. Se pueden encontrar ejemplos de redes temáticas en todos los campos del saber: *Retics.net*,³⁴⁵ en el ámbito de la investigación cooperativa sanitaria; Red Temática de Acuicultura; Red de conservación de tierras y aguas; Red de coordinación de las actividades de investigación en Astrofísica, Red de volcanología, Red virtual de formación y difusión del CSIC; Red del patrimonio histórico y cultural,³⁴⁶ la red *Turismo 2.0*,³⁴⁷ etc. Asimismo, en el ámbito de la traducción y de la lingüística también nos encontramos con redes temáticas, cuyo objetivo es permitir la conexión y la interacción de los profesionales de la traducción. Como ejemplos, podemos nombrar algunas de las redes temáticas financiadas por el Ministerio de Innovación y Ciencia español relacionadas con el campo de la traducción y la lingüística: *DWeb*,³⁴⁸ de la Universidad de Deusto; la red temática *Observatorio de Tecnologías de la Traducción*,³⁴⁹ del grupo de investigación TecnoLeTTra, de la Universidad de Jaume I (Castellón); la red temática *Lengua y Ciencia*,³⁵⁰ de la Universidad de Salamanca; la *Red Temática Española de Morfología*,³⁵¹ de la Universidad Autónoma de Madrid; la *Red Temática en Tecnologías del Habla*, de la Universidad de Sevilla, etc. Fuera de España, también nos encontramos con redes

³⁴⁴ <<http://lites.uji.es/mailman/listinfo/infotrad>>

³⁴⁵ <<http://www.retics.net/>>

³⁴⁶ Estas redes temáticas forman parte del Centro Superior de Investigaciones Científicas de España. URL: <http://www.csic.es/redes_tematicas.do>

³⁴⁷ <<http://www.turismo20.com/>>

³⁴⁸ <http://www.deli.deusto.es/News/1080209131/index_html>

³⁴⁹ <http://tecnolettra.uji.es/es/?page_id=44>

³⁵⁰ <<http://seneca.uab.es/files/neolcyt/traci/>>

³⁵¹ <http://www.uam.es/gruposinv/retem/default_archivos/Page3634.html>

temáticas de este campo de especialidad: *Community Translation Interface*,³⁵² *Hatim-Online Community*,³⁵³ *LingoZ*,³⁵⁴ y *Molan, Lanqua, Enlu, TNP Languages*,³⁵⁵ entre otras muchas.

Para finalizar con las fuentes telemáticas, trataremos ahora los portales y las bibliotecas de referencia en línea, dos recursos que no se pueden encuadrar en ninguna de las dos opciones de sistemas de búsqueda anteriores.

Los portales son sitios web cuyo objetivo es ofrecer al usuario, de forma fácil e integrada, el acceso a una serie de recursos y de servicios, entre los que suelen encontrarse buscadores, foros, documentos, aplicaciones, compra electrónica, etc. Se trata de unos sistemas de búsqueda muy populares en la actualidad que forman la mayor parte de las páginas disponibles en Internet.³⁵⁶ En este caso, nos limitaremos a nombrar algunos de los portales dedicados a la traducción y de turismo que destacan en la actualidad: en primer lugar, para la traducción podemos consultar *Spanish Translator Services*,³⁵⁷ *Foreignword*,³⁵⁸ *CVC. Recursos para traductores*,³⁵⁹ *TRADUguide*,³⁶⁰ *Portal für Rechtsterminologie und Übersetzen im Internet*,³⁶¹ entre otros; y para turismo, encontramos todas las páginas de regiones turísticas del mundo (España,³⁶² Andalucía,³⁶³ Extremadura,³⁶⁴ Málaga,³⁶⁵ Alemania,³⁶⁶ Berlín,³⁶⁷ etc.), así como de

³⁵² <<https://cti.sunvirtuallab.com/community/index.jsp?token=496264528>>

³⁵³ <<http://www.hatim-online.com/de/index.php>>

³⁵⁴ <<http://www.lingoz.com/de/Home>>

³⁵⁵ Estas cuatro últimas redes temáticas están creadas dentro del European Language Council de Berlín. URL: <<http://www.celelc.org/>>

³⁵⁶ En la página Web de la compañía líder en estudios sobre Internet y sus recursos, *Alexa Internet* (<http://www.alexa.com/>), se pueden consultar ranking de los sitios más visitados, según el país, la categoría, etc.

³⁵⁷ <<http://www.spanish-translator-services.com/>>

³⁵⁸ <<http://www.foreignword.com/>>

³⁵⁹ <<http://cvc.cervantes.es/traductores/>>

³⁶⁰ <<http://deutsch.traduguide.com/>>

³⁶¹ <<http://rechtsinformatik.jura.uni-sb.de/portal/>>

³⁶² <<http://www.spain.info/Tourspain>>

³⁶³ <<http://www.andalucia.org/>>

³⁶⁴ <<http://www.turismoextremadura.com/>>

revistas turísticas (*Turismo.com*),³⁶⁸ hoteles (*AC Hoteles*),³⁶⁹ agencias de viajes (*Halcón Viajes*),³⁷⁰ transporte (*Renfe*),³⁷¹ etc.

Para finalizar con los sistemas de búsqueda tenemos que hablar de las bibliotecas de referencia en línea. Estos sistemas ofrecen al usuario una selección de obras de referencia disponibles en Internet, generalmente catalogada por temas y realizada por profesionales de documentación. De las bibliotecas de referencia que destacan actualmente, podemos indicar las *Bibliotecas de Referencia de los Centros de Información de las Naciones Unidas (CINU)*,³⁷² la *Biblioteca de referencia en línea de la UNED*,³⁷³ la *Biblioteca de la Universidad Complutense*,³⁷⁴ *Europeana*,³⁷⁵ *The European Library*,³⁷⁶ *Bibliothek*,³⁷⁷ etc. En el ámbito de la traducción, destaca la biblioteca de la Universidad Pompeu Fabra,³⁷⁸ de Barcelona, donde se recogen numerosos recursos dirigidos a traductores e intérpretes.

El segundo grupo de fuentes de información se corresponde con las fuentes institucionales, que aportan información sobre instituciones, organismos, empresas, entidades o colectivos en todos sus aspectos. En las propuestas de clasificación de Seghiri Domínguez (2006) y de Gonzalo García (2004 [2003]), solo se tienen en cuenta los recursos electrónicos relacionados con estas fuentes institucionales, es decir, los directorios y los sitios web que ofrecen información sobre instituciones. Sin embargo, si

³⁶⁵ <<http://www.malagaturismo.com>>

³⁶⁶ <<http://www.deutschland-tourismus.de/>>

³⁶⁷ <<http://www.visitberlin.de/index.es.html>>

³⁶⁸ <<http://www.turismo.com/>>

³⁶⁹ <<http://www.ac-hotels.com/>>

³⁷⁰ <<http://www.halconviajes.com/>>

³⁷¹ <<http://www.renfe.es>>

³⁷² <<http://unic.un.org/aroundworld/unic/es/whatWeDo/productsAndServices/libraries/index.asp>>

³⁷³ <<http://www.uned.es/biblioteca/referencia/PAGINAINICIO.HTML>>

³⁷⁴ <<http://www.ucm.es/BUCM/>>

³⁷⁵ <<http://www.europeana.eu/portal/>>

³⁷⁶ <<http://search.theeuropeanlibrary.org/portal/en/index.html>>

³⁷⁷ <<http://www.bibliothek.de/>>

³⁷⁸ <<http://www.upf.edu/bibtic/es/traduc/traduc.html>>

ampliamos el campo de visión e incluimos las fuentes en formato papel, también podríamos añadir en este apartado todos los folletos informativos, manuales, catálogos, etc. que fueran publicados por instituciones, tanto nacionales como internacionales. Como ejemplo de fuentes institucionales en la Red, podemos nombrar el sitio web del Ayuntamiento de Málaga,³⁷⁹ del Ministerio de Asuntos Exteriores,³⁸⁰ de la Unión Europea,³⁸¹ de la Organización de las Naciones Unidas,³⁸² etc. En el ámbito de la traducción, también encontramos numerosas fuentes institucionales de asociaciones de traductores e intérpretes, como la Asociación Ibérica de Estudios de Traducción e Interpretación (AIETI),³⁸³ Asociación Española de Traductores, Correctores e Intérpretes (ASETRAD),³⁸⁴ Bundesverband der Dolmetscher und Übersetzer e.V. (BDÜ),³⁸⁵ Institute of Translation and Interpretation (ITI),³⁸⁶ etc.

En tercer lugar, nos encontramos las fuentes personales, que hacen referencia a las fuentes que aportan información sobre personas o grupos entre las que existe normalmente una relación profesional. Dentro de este apartado, se incluyen las Guías de expertos y los Quién es Quién. Ambos recursos nos facilitan, aunque con una metodología diferente, la localización de expertos en diferentes materias para poder realizar consultas sobre cualquier temática especializada sobre la que estemos traduciendo (en nuestro caso, sobre la traducción turística).

Como ejemplos de Guía de expertos, citaremos las páginas web *Alle Experten*³⁸⁷ o *Todoexpertos.com*,³⁸⁸ que se tratan de guías de expertos de distintas especialidades;

³⁷⁹ <<http://www.malaga.eu/opencms/opencms/aytomalaga/portal/portada/portada>>

³⁸⁰ <<http://www.maec.es/es/home/Paginas/HomeEs.aspx>>

³⁸¹ <http://europa.eu/index_es.htm>

³⁸² <<http://www.un.org/>>

³⁸³ <<http://www.aieti.net>>

³⁸⁴ <<http://www.asetrad.org>>

³⁸⁵ <<http://www.bdue.de>>

³⁸⁶ <<http://www.iti.org.uk>>

³⁸⁷ <<http://www.du-machst.de/index.php?id=26>>

³⁸⁸ <<http://www.todoexpertos.com/>>

guías de expertos de universidades, como la de la Universidad de Navarra,³⁸⁹ la Universidad de La Laguna,³⁹⁰ la Universidad de Kiel,³⁹¹ la Universidad de Bielefeld,³⁹² etc. Con relación al campo de la traducción, se pueden considerar guías de expertos los actuales buscadores de traductores disponibles en Internet, como *Das Übersetzung-Verzeichnis*,³⁹³ *Standard Linguistic*,³⁹⁴ *Translation Town*,³⁹⁵ la lista de miembros del *Institute of Translation and Interpreting (ITI)*,³⁹⁶ *Dolmetscher und Übersetzer Suche*,³⁹⁷ etc. Con estos buscadores, se localizan de forma ágil y rápida traductores, que se han inscrito previamente en la base de datos, especializados en una materia o en una combinación de lenguas determinada.

Además de las guías de expertos, en las fuentes personales también encontramos varios Quién es Quién disponibles, como el de la Unión Europea (*EU Whoiswho*),³⁹⁸ *Quién es quién en España*, *Who's Who Online*,³⁹⁹ *Wer ist wer. Das deutsche Who is Who*,⁴⁰⁰ etc., que nos aportan información sobre los expertos de una empresa, una institución, un campo de especialidad, etc.

Dentro de las fuentes personales, debemos hablar también las páginas personales, o los actuales *blogs*, que están teniendo cada vez más presencia en Internet. En el campo de la traducción, encontramos numerosas páginas personales de traductores profesionales que ofrecen información personal y laboral, a la vez que proporcionan recopilatorios de enlaces a recursos útiles para traductores (diccionarios, bases de datos,

³⁸⁹ <<http://www.unav.es/comunicacioninstitucional/expertos.html>>

³⁹⁰ <<http://www2.ull.es/xpert.htm>>

³⁹¹ <<http://www.uni-kiel.de/ps/cgi-bin/expfg.php>>

³⁹² <<http://www.uni-bielefeld.de/Universitaet/Einrichtungen/Pressestelle/expertenthemen.html>>

³⁹³ <<http://www.uebersetzer-suche.eu/>>

³⁹⁴ <<http://standling.ods.org/search/trads/>>

³⁹⁵ <<http://www.translationtown.com/>>

³⁹⁶ <<http://www.iti-lrg.org.uk/>>

³⁹⁷ <<http://www.experten-branchenbuch.de/uebersetzer/>>

³⁹⁸ <<http://europa.eu/whoiswho/public/index.cfm?lang=es>>

³⁹⁹ <<http://www.whoswho-online.com/>>

⁴⁰⁰ <http://www.beleke.de/shop/RID/185/wer_ist_wer_das_deutsche_whos_who>

programas informáticos, asociaciones, revistas de traducción, etc.).⁴⁰¹ Incluso en este apartado podríamos incluir los perfiles de profesionales de la traducción y académicos en las redes sociales, como *Facebook* o *Twitter*, que al igual que los *blogs* están teniendo éxito en la actualidad.

Las fuentes normativas son el siguiente grupo de fuentes de información que se incluye en estas propuestas. Este tipo de fuente nos ofrece información sobre las normas o especificaciones técnicas que se deben cumplir para garantizar la seguridad, eficacia y calidad en una actividad determinada, como puede ser la traducción de textos especializados. Para los traductores, las fuentes normativas más relevantes y a las que hay que prestar especial atención son las organizaciones de normalización, como son la *International Organization for Standardization (ISO)*,⁴⁰² *European Committee for Standardization (CEN)*⁴⁰³ y la *Asociación Española de Normalización y Certificación (AENOR)*,⁴⁰⁴ las cuales crean las normas y las especificaciones, anteriormente mencionadas, para establecer una normalización y una unificación sobre cualquier ámbito.

El siguiente grupo de fuentes de información se trata de las fuentes legislativas, que hace referencia a toda la información sobre los órganos legislativos y a toda documentación legislativa que producen, así como a las páginas web que se dedican a solventar dudas legislativas. Dentro de este grupo, encontramos gran cantidad de páginas web dedicadas a este tema. Por lo que respecta al alemán, podemos citar las siguientes como ejemplo: *Gesetze und Verordnungen*,⁴⁰⁵ página web de la Deutsche Gesellschaft für Reiserecht e. V.; *Reiserecht für Reisende*,⁴⁰⁶ página del bufete de

⁴⁰¹ En la página personal de la traductora Diana Márquez (<http://www.webpersonal.net/dmarques/resources/trad3.htm>), se recogen numerosos enlaces a páginas personales de traductores profesionales, lo que se considera un correcto directorio para el campo de la traducción.

⁴⁰² <<http://www.iso.ch/iso/en/ISOOnline.openerpage>>

⁴⁰³ <<http://www.cenorm.be/>>

⁴⁰⁴ <<http://www.aenor.es/desarrollo/inicio/home/home.asp> >

⁴⁰⁵ <<http://www.dgff.de/>>

⁴⁰⁶ <<http://www.recht-im-tourismus.de/>>

abogados Andersch & Smettan-Öztürk, Gesetze in WWW,⁴⁰⁷; *Gesetze im Internet*, del Ministerio de Justicia alemán;⁴⁰⁸ *dejure.org*, directorio jurídico de la empresa Rechtsinformationssysteme GmbH,⁴⁰⁹ *Justiz online*,⁴¹⁰ base de datos del gobierno regional de Renania del Norte-Westfalia que ofrece legislación estatal y regional, etc. Para la legislación española, podemos consultar la página web del Ministerio de Industria, Turismo y Comercio,⁴¹¹ la página de la Mesa del Turismo. Legislación General Turística,⁴¹² *El Web Jurídico*, un portal jurídico;⁴¹³ *Pórtico Legal*, un directorio con información jurídica;⁴¹⁴ *ARTI*,⁴¹⁵ la base de datos oficial del Senado; *Iberlex*,⁴¹⁶ base de datos del Ministerio de Presidencia que almacena legislación estatal, autonómica y comunitaria; *Colex-Data*,⁴¹⁷ bases de datos españolas para la consulta de legislación y jurisprudencia, es altamente recomendado para los profesionales del Derecho; *Legistur*,⁴¹⁸ la base de datos del Instituto de Estudios Turísticos del Ministerio de Turismo, Industria y Comercio, etc. Y para la legislación inglesa, encontramos la *Office of Public Sector Information* (OPSI),⁴¹⁹ donde se aloja *Her Majesty's Stationery Office*, editorial del gobierno británico donde se pone a disposición del público numerosas publicaciones oficiales (reglamentos, órdenes, etc.) en línea y de forma gratuita; *DirectGov*,⁴²⁰ página web del gobierno del Reino Unido que ofrece enlaces a diferentes instituciones británicas y amplia información sobre diversos temas de interés

⁴⁰⁷ <<http://www.rechtliches.de/>>

⁴⁰⁸ <<http://www.gesetze-im-internet.de/>>

⁴⁰⁹ <<http://dejure.org/>>

⁴¹⁰ <http://www.lexsoft.de/lexisnexis/justizportal_nrw.cgi>

⁴¹¹ <<http://www.mityc.es/es-ES/Servicios/Legislacion/>>

⁴¹² <<http://www.mesadelturismo.com>>

⁴¹³ <<http://www.intercom.es/webjur/>>

⁴¹⁴ <<http://www.porticolegal.com>>

⁴¹⁵ <<http://www.senado.es/basesdedatos/index.html>>

⁴¹⁶ <<http://www.boe.es/g/es/iberlex/>>

⁴¹⁷ <<http://www.colex-data.com/modules/common/homepage.jsp>>

⁴¹⁸ <http://www.iet.tourspain.es/paginas/cdte/legistur.aspx?option=doc&option2=BD_ES&idioma=es-ES>

⁴¹⁹ <<http://www.opsi.gov.uk/>>

⁴²⁰ <<http://www.direct.gov.uk/en/index.htm>>

(educación, transporte, viajes, medio ambiente, impuestos, etc.); o el portal de *BAILII* (*Bristish and Irish Legal Information*),⁴²¹ que permite el acceso de forma gratuita a la legislación del Reino Unido e Irlanda, así como a la legislación comunitaria. En el ámbito europeo o internacional, destacamos la base de datos *Lextour*,⁴²² de la Organización Mundial del Turismo; *CELEX*⁴²³ y *EUR-Lex*,⁴²⁴ bases de datos de legislación y documental de la Unión Europea.

Por último, nos encontramos con las fuentes lingüísticas, que hacen referencia a las fuentes que emplea el traductor para extraer la información que necesita sobre unidades léxicas, terminológicas o fraseológicas, ya sea para encontrar su definición, un equivalente, información gramatical o semántica o cualquier tipo de información lingüística. Dentro de este grupo, podemos distinguir tres tipos, a saber: fuentes lexicográficas, constituidas fundamentalmente por diccionarios generales monolingües, bilingües o multilingües (*Diccionario de la Real Academia Española*, como ejemplo de fuente monolingüe); fuentes gramaticales, formadas por las gramáticas, las ortografías y los manuales de estilo de las lenguas de trabajo (*Manual de estilo de El País*, del diario El País; *Manual de dudas y dificultades*, de M. Seco; *Ortografía de la Real Academia Española*, de la Real Academia Española, como ejemplos de fuentes gramaticales); y fuentes terminológicas, que ofrecen información sobre las unidades terminológicas y fraseológicas del ámbito de especialidad (glosarios, vocabularios, bases de datos y diccionarios especializados).⁴²⁵ Este tipo de fuentes de información se puede encontrar tanto en formato electrónico como en papel, aunque su presencia en Internet va en aumento. Cada vez más se prefiere publicar estas obras de referencia en formato electrónico, ya sea en la Red o en CD-ROM, ya que mediante este medio se obtienen unas ventajas que no se pueden conseguir en papel, como son la actualización, la ampliación, la reducción de costes, la inmediatez, la facilidad de movimiento e incluso la mejora a través de *feedback* de los usuarios.

⁴²¹ <<http://www.bailii.org/>>

⁴²² <<http://www.world-tourism.org/doc/S/lextour.htm>>

⁴²³ <<http://europa.eu.int/eur-lex/lex/>>

⁴²⁴ <<http://europa.eu.int/celex/>>

⁴²⁵ Estas fuentes terminológicas serán abordadas extensamente en el siguiente apartado.

Además de estas fuentes lingüísticas, también se utilizan los servicios de consultas lingüísticas que ponen a disposición de los usuarios diferentes instituciones y organizaciones. Como ejemplo de estos servicios, podemos citar el que ofrecen en sus páginas web la Real Academia Española,⁴²⁶ la Real Academia Canaria,⁴²⁷ la empresa de traducción Hermes Traducciones y Servicios Lingüísticos, S.A.,⁴²⁸ entre otros.

Como hemos visto con estas propuestas de clasificación de los recursos electrónicos (aunque también se pueden utilizar para los de otro formato), las posibilidades que ofrece Internet son muy amplias y aportan gran facilidad a la hora de documentarse durante el proceso traductor, a la vez que ha aumentado la calidad de las traducciones especializadas gracias a las consultas de todo tipo de información a través de la Red. No obstante, el volumen ingente de información disponible en Internet puede llevar al traductor a la confusión y al error si los recursos que selecciona no disponen de fiabilidad o rigor científico. A este fenómeno se le conoce actualmente como «infoxicación», un neologismo que ha introducido Cornellà (2000) para designar el exceso de información que puede llegar a desbordar la capacidad de asimilación humana. A partir de su aparición, se han empezado a realizar trabajos en torno a este término y, de forma curiosa, ahora se pueden consultar en Internet más de 21.000 páginas relacionadas con ese término. De nuevo, nos encontramos con la omnipresente «infoxicación».

Por este motivo, uno de los aspectos más importantes en la competencia documental del traductor profesional es la capacidad de filtro y de evaluación de estos recursos documentales. En otras palabras, el traductor no solo debe ser capaz de buscar y de consultar los recursos documentales disponibles en la Red, sino que además debe contar con conocimientos de evaluación de recursos electrónicos para poder filtrar toda la información presente en Internet y, así, seleccionar la información que considere

⁴²⁶ <http://cronos.rae.es/cgi-bin/consulta_form.pl>

⁴²⁷ <<http://www.academiacanarialengua.org/consultas.php>>

⁴²⁸ <http://www.hermestrans.com/es/index.html>

pertinente y fiable. De este modo, podrá determinar si es adecuado utilizar el recurso encontrado o no.⁴²⁹

2.4. LAS FUENTES DE INFORMACIÓN TERMINOLÓGICA

Los diccionarios y las enciclopedias terminológicos son unos auxiliares de los que en muchos casos no se puede prescindir si uno quiere realizar un trabajo científico y responsable (Martínez de Sousa, 2004: 173).

Como hemos indicado con anterioridad, la terminología es un elemento clave en la traducción de textos especializados y, por ello, los traductores necesitan contar con una competencia documental bien desarrollada que les permita alcanzar los mejores resultados en su traducción. En este contexto, un traductor profesional se puede enfrentar a problemas terminológicos de distinta naturaleza, como se muestra a continuación (Cabré Castellí, 2000: 216):

1. Problemas relacionados con la lengua de origen:
 - Conocer o precisar el significado de una unidad terminológica del texto de origen.
 - Confirmar el carácter especializado de una unidad.
 - Conocer las denominaciones alternativas y las condiciones relativas de uso en los textos, situaciones comunicativas, etc.
2. Problemas relacionados con la lengua de llegada:
 - Conocer la existencia de una unidad terminológica equivalente en la lengua de llegada.

⁴²⁹ En el apartado 3.3.6.3. de este capítulo, profundizaremos en la evaluación de los recursos documentales, haciendo hincapié en los recursos terminográficos, y presentaremos una plantilla de evaluación de elaboración propia.

- En caso de que no exista ninguna unidad, conocer los instrumentos adecuados para adaptar o crear una denominación correcta.
- Conocer la denominación más apropiada, teniendo en cuenta el tema, los objetivos del trabajo y los destinatarios.
- Conocer las posibilidades combinatorias de las unidades.
- Conocer la fraseología del ámbito de especialidad.
- Conocer las denominaciones alternativas para un mismo concepto y las condiciones pragmáticas relativas.
- Asegurarse de que la denominación seleccionada corresponda exactamente con el concepto que debe designar.

Algunos de estos problemas terminológicos se podrían subsanar con una formación específica en terminología a los traductores como en el caso de la creación de neologismos o en la elaboración de terminología *en la traducción (terminología puntual o ad hoc)* (cf. Cabré Castellví, 2004: 117). No obstante, los traductores pretenden solucionar sus dudas terminológicas mediante la consulta de fuentes de información terminográfica, ya sean estos en papel o en formato electrónico, es decir, mediante recursos terminográficos *para* el traductor (ibíd.).

La gestión terminológica puntual o *ad hoc* que realiza el traductor durante el proceso de análisis y traducción de un texto es a menudo consecuencia de la necesidad de enfrentarse a una terminología que no se recoge adecuadamente en las fuentes de información terminográfica o que aparece de forma aislada. Este tipo de terminología hace referencia a glosarios, vocabularios, diccionarios, bases de datos, etc. que va elaborando el propio traductor y que recogen sus soluciones y resultados obtenidos durante sus traducciones, de manera puntual y no sistemática. En este grupo también se deben tener en cuenta los corpus textuales, paralelos o comparables, que el traductor compila para conseguir un mayor acercamiento conceptual a la materia, así como para observar los términos en un contexto real, para adquirir un mayor conocimiento del lenguaje que se utiliza en el dominio de especialidad, para encontrar ejemplos de uso,

para extraer términos que le permitan crear sus propios recursos, entre otras razones.⁴³⁰ Se trata, por tanto, de recursos muy vinculados a cada traducción y a cada traductor, donde la información terminológica recogida (equivalencias, colocaciones, fraseología, observaciones de uso, falsos amigos, etc.) tiene que ver con las preocupaciones de resolver de manera adecuada y coherente los problemas que plantea la traducción del texto (Gómez González-Jover y Vargas Sierra, 2004: 367).

A pesar de que este tipo de gestión terminológica es una realidad en la vida diaria del traductor, debemos realizar dos puntualizaciones al respecto con relación a las respuestas obtenidas en la encuesta que se presentará más adelante (cf. 2.5.2.): por un lado, los traductores prefieren utilizar los recursos terminográficos disponibles (como diccionarios, bases de datos, etc.), cuya opción obtuvo un 49,29%, frente a crear recursos por sí mismos (20,33%); y por otro lado, el 34,78% de los traductores profesionales contactados contestaron que no a la pregunta: *¿Compila su propio corpus (conjunto de textos electrónicos) cuando traduce?*, y la razón principal por la que no lo hacían era por falta de tiempo (60%).

Esta postura permite concluir que la terminología puntual o *ad hoc* en la práctica traductológica ocupa un lugar de preferencia secundario frente a la terminología sistemática. Esto se debe principalmente a las desventajas que pueden encontrar los traductores-terminólogos durante este trabajo de terminología *ad hoc*, como nos indican Wright y Wright (1997: 148):

⁴³⁰ Estas razones por las que los traductores profesionales compilan y utilizan los corpus en sus traducciones han sido extraídas de una encuesta realizada como parte de este trabajo de investigación, la cual se presentará en el apartado 2.5.2. de este capítulo y que se encuentra como Anexo I. Exactamente, la pregunta que se realizó en la encuesta y sus diferentes opciones (con los porcentajes obtenidos) se muestra a continuación:

18.1. ¿Por qué compila su propio corpus?

- Los recursos terminográficos de mi dominio no son fiables (2,46%).
- Puedo comprobar los términos en su contexto (27,10%).
- Puedo comprobar el uso real del lenguaje de un dominio (12,46%).
- Puedo buscar ejemplos (19,86%).
- Puedo entender mejor los significados (14,93%).
- Puedo extraer los términos para hacer mi propio banco terminológico (20,58%).
- Otro (2,61%).

- Por lo general, no son expertos en el dominio de especialidad en el que están trabajando.
- Pueden encontrar dificultades a la hora de determinar el dominio de especialidad al que pertenece el texto en el que trabajan.
- Los materiales de investigación disponibles tanto en la lengua de llegada como de partida son inadecuados.
- Carecen de acceso a especialistas del dominio en cuestión.
- Debido a la falta de tiempo y a los plazos restrictivos, no pueden realizar un trabajo de investigación extenso (y profundo).
- Incluso cuando la documentación se encuentra disponible, carecen del tiempo necesario para crear entradas terminológicas amplias y bien documentadas.

No obstante, a veces ocupa un lugar primordial en la práctica traductora, especialmente cuando los recursos terminográficos existentes y disponibles no satisfacen las necesidades de los traductores, ya sea porque no están actualizados, les falta información necesaria (pragmática, de contexto, lingüística), etc., y, por tanto, no ayudan a solucionar los problemas terminológicos que veíamos al comienzo de este apartado.

La terminología sistemática, es decir, la terminología *para* el traductor, hace referencia a los repertorios léxicos y terminológicos (diccionarios, enciclopedias, vocabularios, léxicos, bases de datos, etc.) elaborados según una metodología sistemática, es decir, a partir de un estudio exhaustivo de las unidades terminológicas de un dominio de especialidad concreto (áreas o subáreas) y de la comprensión total de la organización de dicho dominio conceptual. En un marco traductológico, estos recursos deberían contener información útil para ayudar a la toma de decisiones en el proceso de la traducción, pero, como veremos más adelante, esto no siempre se puede afirmar.

En nuestro trabajo, nos limitaremos a tratar los recursos que se elaboran de forma sistemática, es decir, los recursos *para* el traductor,⁴³¹ y presentaremos a continuación una clasificación de estos recursos, así como una propuesta de evaluación para conocer la fiabilidad y adecuación de estos para su utilización durante la traducción de un texto especializado.

2.4.1. CLASIFICACIÓN DE LOS RECURSOS TERMINOLÓGICOS

Como veíamos en el apartado anterior, las fuentes de información terminográfica proporcionan la información sobre las unidades terminológicas o fraseológicas de un ámbito de especialidad. Según la norma ISO 12620: 2, un «recurso terminológico» se definiría de la siguiente manera:

Database collection consisting of individual data entries each of which documents a concept and provides data pertinent to the terms associated with that concept in one or more language.

La tipología de fuentes de esta naturaleza es muy amplia y los contenidos que nos ofrecen, muy diferentes. Además, como ocurre en otros casos que ya hemos discutido, no existe todavía una clasificación única sobre los tipos de fuentes terminológicas, ya que los especialistas no se han puesto aún de acuerdo y muchos han propuesto sus propias clasificaciones. En este trabajo de investigación, hemos partido de las clasificaciones realizadas por Rey (1970), Zgusta (1971), Malkiel (1975), Alvar Ezquerro (1980, 1996), McArthur (1986), Hausmann (1989, vol. 1), Gak (1991), Bergenholtz y Tarp (1995), Roberts (1997), Haensch (1997), Bajo Pérez (2000), García Palacios (2003) y Martínez de Sousa (2004), a fin de confeccionar una clasificación de fuentes terminológicas, completa e integradora, ajustada a los recursos que se pueden consultar actualmente. Estas clasificaciones, al igual que la gran mayoría de las

⁴³¹ Para profundizar en la terminología puntual por parte de los traductores, recomendamos la consulta de la encuesta *LISA Terminology Management Survey: Terminology Management Practices and Trends* (2005), que proporciona un informe sobre las prácticas terminológicas relacionadas con la industria de la localización, y el estudio realizado por Zielinski y Ramírez Safar (2005), sobre la gestión y la extracción terminológica, entre otros.

clasificaciones realizadas hasta el momento, se basan en la macroestructura de la fuente y no toman como valor de clasificación las diferentes microestructuras. En nuestro caso, también hemos tomado la macroestructura de las fuentes terminológicas como objetivo para su clasificación, ya que la microestructura de cada fuente puede variar según los criterios de publicación, del autor, etc. y no se puede concretar un único proceder.

Nuestra clasificación recoge los siguientes criterios distintivos de las fuentes terminológicas:⁴³²

1. Según la presentación.
2. Según el número de lenguas
3. Según el tipo de soporte
4. Según la ordenación de sus contenidos
5. Según la normalización
6. Según el grupo de destinatarios
7. Según el tipo de unidades terminológicas
8. Según la extensión en la especialidad
9. Según la fuente de información utilizada
10. Según la finalidad

2.4.1.1. Según la presentación

Según este criterio, las fuentes terminológicas se pueden dividir en ocho grupos diferentes: diccionarios especializados (o terminológicos), diccionarios visuales, enciclopedias, vocabularios, glosarios, léxicos, tesauros, siglas y acrónimos, abreviaturas y símbolos, bases de datos y corpus textuales.

⁴³² A lo largo de la presente clasificación, presentaremos ejemplos que ilustren cada tipo de obra de referencia.

Los diccionarios especializados incluyen información sobre las unidades terminológicas de una materia de especialidad determinada y son muy precisos en la información que proporcionan. En algunos casos, estos diccionarios pueden ser tratados como diccionarios enciclopédicos, pero siempre que estos añadan a la definición una explicación más o menos extensa para describir la materia o el aspecto concreto que se define, incluyendo sus fundamentos científicos, historia, características técnicas, etc. Actualmente, se dispone de un gran abanico de estos diccionarios especializados en diferentes materias, como por ejemplo el *Diccionario de turismo*, de Montaner Montejano et ál. (1998), el *Diccionario de terminología técnica multilingüe*, de la Universidad Politécnica de Catalunya, etc.

Los diccionarios visuales especializados permiten buscar las denominaciones de una temática a partir del concepto, que se encuentra representado a partir de una ilustración. Mediante este proceso, se pueden mostrar claramente los objetos, las máquinas, los procesos, etc. así como cada una de sus partes. Como ejemplos, podemos nombrar los siguientes: *The Culturally Authentic Pictorial Lexicon*,⁴³³ en inglés y alemán; el *Glosario de Iconos turísticos*,⁴³⁴ en español, etc.

Las enciclopedias ofrecen gran cantidad de información sobre las diferentes disciplinas, puesto que, además de incluir terminología, presentan datos conceptuales mediante clasificaciones, esquemas, ilustraciones, etc. Se diferencian con los diccionarios enciclopédicos en que no siempre proporcionan una definición del término que presentan. Asimismo, se pueden distinguir dos tipos de enciclopedias: enciclopedias sistemáticas, también llamadas temáticas o metódicas (Martínez de Sousa, 2004: 177) y enciclopedias alfabéticas. Estas enciclopedias se distinguen entre sí por el tipo de ordenamiento que siguen, ya que uno es de tipo sistemático y otro de tipo alfabético. Como enciclopedias, podemos citar la enciclopedia ilustrada de Salud de A.D.A.M., de *MedlinePlus*,⁴³⁵ la enciclopedia colaborativa *Wikipedia*,⁴³⁶ la *Brockhaus Enzyklopädie*; la *Enciclopedia Libre Universal en Español*,⁴³⁷ etc.

⁴³³ <<http://www.washjeff.edu/CAPL/>>

⁴³⁴ <<http://www.iberiarural.es/glosario.htm>>

⁴³⁵ <<http://www.nlm.nih.gov/medlineplus/spanish/encyclopedia.html>>

Las siguientes fuentes, vocabulario, léxico y glosario, presentan cierta controversia en su definición, ya que los tres son, en cierta manera, hipónimos de los diccionarios (Elena Bajo, 2000: 14) y, además, la línea de separación entre ellos es a menudo poco clara. Por ello, hemos utilizado las definiciones que nos ofrece el *Diccionario de la Real Academia de la Lengua* (DRAE, 2008, versión en línea), para aclarar sus diferencias, así como otras fuentes de referencia en los casos en que podía haber algún tipo de duda en las definiciones.

Según el DRAE, la definición de «vocabulario» es, en su tercera acepción, el «conjunto de palabras de un idioma pertenecientes al uso de una región, a una actividad determinada, a un campo semántico dado, etc.» y, añade, en la quinta, que es un «catálogo o lista de palabras, ordenadas con arreglo a un sistema, y con definiciones o explicaciones sucintas». Asimismo, la Norma ISO 1087 (1990) *Vocabulary of Terminology* lo define de la siguiente manera: «A terminological dictionary which contains designations and definitions from one or more specific subject fields. NOTE: The vocabulary may be monolingual, bilingual or multilingual». De esta forma, y siguiendo las dos definiciones dadas, consideramos vocabulario como la fuente terminológica que ofrece definiciones de las unidades terminológicas que incluye de un ámbito restringido del campo de conocimiento y que, además, puede proporcionar equivalentes en otras lenguas. En el caso del sector turístico, un vocabulario trataría sobre un segmento, una actividad concreta, unos agentes determinados, etc. e incluiría definiciones de sus términos incluidos, como por ejemplo *The Climbing Dictionary*,⁴³⁸ el *Diccionario turístico*,⁴³⁹ de la página web *Mallorca web site*, etc.

La definición que ofrece el DRAE (2008, versión en línea) sobre «léxico» no aclara realmente la diferencia que existe con «vocabulario», ya que lo define «vocabulario, conjunto de las palabras de un idioma, o de las que pertenecen al uso de una región, a una actividad determinada, a un campo semántico dado, etc.». Sin

⁴³⁶ <<http://es.wikipedia.org/wiki/Wikipedia:Portada>>

⁴³⁷ <http://enciclopedia.us.es/index.php/Enciclopedia_Libre_Universal_en_Espa%C3%B1ol>

⁴³⁸ <http://home.tiscalinet.de/ockier/climbing_dict.htm>

⁴³⁹ <<http://www.mallorcawebsite.com/balearik/4diccionario.htm>>

embargo, después de consultar otras fuentes de referencia (Cabré, 1999: 33), hemos determinado distinguir «léxico» de «vocabulario» en que el léxico es un inventario de términos, acompañados de sus equivalentes en una o más lenguas, que no ofrece definición, como por ejemplo *Foreign language for travelers*,⁴⁴⁰ el *Diccionario multilingüe de buceo y actividades subacuáticas*,⁴⁴¹ el *Windsurf-Dictionary*,⁴⁴² el *Gastronomisches Wörterbuch*,⁴⁴³ etc.

La definición del DRAE de «glosario» también lleva a confusión, ya que lo asemeja a lo que se ha denominado anteriormente como «vocabulario». En su primera acepción, lo define como «Catálogo de palabras oscuras o desusadas, con definición o explicación de cada una de ellas» y, en su segunda acepción, como «Catálogo de palabras de una misma disciplina, de un mismo campo de estudio, etc., definidas o comentadas». Por ello, preferimos manejar la definición de «glosario» de Haensch (1997: 46-47), que lo define como sigue: «un repertorio o listado, generalmente no muy extenso, de palabras que pertenecen a un subconjunto del léxico, por ejemplo terminología técnica, palabras coloquiales o jergas». Como ejemplos de glosarios, podemos citar el *Glosario de biotecnología para la agricultura y la alimentación*,⁴⁴⁴ de la Food and Agricultural Organisation (FAO); el *Glosario de deportes*,⁴⁴⁵ el *Klettern Glosar*,⁴⁴⁶ el *Glosario de términos* de la Organización Mundial del Turismo,⁴⁴⁷ etc.

Los tesauros, según la norma *UNE 50-106-90* de AENOR (1990), se definen como un «vocabulario de un lenguaje de indización controlado organizado formalmente con objeto de hacer explícitas las relaciones, *a priori*, entre conceptos». Es decir, un tesoro trata de plasmar el sistema conceptual de una materia determinada y consiste en

⁴⁴⁰ <<http://www.travlang.com/languages/search.html>>

⁴⁴¹ <<http://www.foreignword.com/glossary/plongeon/spn/defd.htm>>

⁴⁴² <<http://www.home.umk.pl/~skoter/diction.htm>>

⁴⁴³ <<http://www.todine.net/dictionary.html>>

⁴⁴⁴ <<http://www.fao.org/docrep/004/y2775s/y2775s00.htm#Contents>>

⁴⁴⁵ <http://wikilengua.fundeu.es/index.php/Glosario_de_deportes#Deportes_de_riesgo>

⁴⁴⁶ <<http://www.klettern-md.de/home/klettern/glossar?pn=2>>

⁴⁴⁷ <http://www.wto.org/spanish/thewto_s/glossary_s/glossary_s.htm>

un conjunto de términos estructurados jerárquicamente según la generalidad o especificidad de sus significados y relacionados por vínculos asociativos y de equivalencia. Los tesauros muestran la equivalencia entre las unidades terminológicas del lenguaje natural y los términos normalizados y preferentes del lenguaje documental (Amat Noguera, 1978:179). Tradicionalmente los tesauros no incluyen definiciones de términos, sino solamente en alguno de ellos se añade lo que se conoce como «nota de alcance», que consiste en una aclaración de cómo debe utilizarse el término dentro del sistema documental dado (Maldonado Martínez et ál., 2004). Como ejemplo de tesauros, podemos citar el tesauro plurilingüe de la Agencia Europea de Medio Ambiente, *Eionet*,⁴⁴⁸ el tesauro sobre urbanismo de CINDOC,⁴⁴⁹ el tesauro de matemáticas,⁴⁵⁰ de la Universidad de Cambridge; *Eurovoc*, el tesauro de la Unión Europea,⁴⁵¹ etc.

Los repertorios de siglas y acrónimos están formados por siglas y acrónimos de la disciplina en la que se ubiquen; y los repertorios de abreviaturas y símbolos,⁴⁵² están formados igualmente por las abreviaturas y los símbolos de su campo de especialidad. Ambos repertorios son muy abundantes en el lenguaje especializado, especialmente en el lenguaje científico-técnico, como por ejemplo *Acronym Finder*,⁴⁵³ *Abkuerzungen, die*

⁴⁴⁸ <http://www.eionet.europa.eu/gemet/index_html?langcode=es>

⁴⁴⁹ <http://thes.cindoc.csic.es/index_URBA_esp.html>

⁴⁵⁰ <<http://thesaurus.maths.org/mmkb/view.html?resource=index>>

⁴⁵¹ <<http://europa.eu/eurovoc/>>

⁴⁵² A modo de recordatorio, proporcionamos las definiciones que ofrece el DRAE para «sigla», «acrónimo», «abreviatura» y «símbolo», ya que la diferencia entre ellos no queda siempre clara. Según el DRAE (2008, versión en línea), «sigla» es el «Palabra formada por el conjunto de letras iniciales de una expresión compleja; p. ej. O(rganización de) N(aciones) U(nidas), o(bjeto) v(olante) n(o) i(dentificado), Í(ndice de) P(recios al) C(onsumo)» y «acrónimo» se define como «Tipo de sigla que se pronuncia como una palabra». Normalmente, ambas unidades dan origen a nuevas formaciones léxicas y neologismos.

Según el DRAE (2008, versión en línea), «abreviatura» se refiere a «Tipo de abreviación que consiste en la representación gráfica reducida de una palabra mediante la supresión de letras finales o centrales, y que suele cerrarse con punto» y «símbolo» consiste en «Tipo de abreviación de carácter científico o técnico, constituida por signos no alfabetizables o por letras, y que difiere de la abreviatura en carecer de punto». Estas formaciones suponen la reducción gráfica de una unidad o grupo de unidades con la finalidad de representarlas.

⁴⁵³ <<http://www.acronymfinder.com/>>

Maschine für Abkürzungen,⁴⁵⁴ *Spanisch-Deutsch Abkürzungen*,⁴⁵⁵ *Abreviaturas en Español*,⁴⁵⁶ etc.

Las bases de datos terminológicas⁴⁵⁷ son unos instrumentos relativamente nuevos como fuentes de información terminográfica y han tenido, y siguen teniendo, mucho auge debido principalmente a la incorporación y desarrollo de herramientas informáticas en su elaboración. Estos recursos ofrecen la información más completa y actualizada sobre las unidades terminológicas y fraseológicas de una determinada disciplina. Es el tipo de fuente más utilizado por los traductores a la hora de traducir, aunque cada vez más viene utilizada también por terminólogos, lingüistas, editores, etc., ya que facilita la actualización de la información con su soporte electrónico, permite almacenar grandes cantidades de unidades con mucha información y también permite recuperar la información fácilmente. Como ejemplo, podemos citar las siguientes: *IATE*,⁴⁵⁸ la base de datos de la Unión Europea; *TourisTerm*⁴⁵⁹ y *WTOTERM*, las bases de datos de la Organización Mundial del Turismo; *Terminologiedatenbank*, la base de datos del Institut für Translationswissenschaft Innsbruck;⁴⁶⁰ *WebTerm*,⁴⁶¹ la base de datos sobre terminología turística de la Universidad de Colonia, etc.

Por último, encontramos los corpus textuales. La mayoría de los autores consultados no hacen referencia a estos recursos, pero consideramos necesario incluirlos

⁴⁵⁴ <<http://www.abkuerzungen.de/>>

⁴⁵⁵ <http://www.anabell.de/spanien/spanische_deutsch_abkuerzungen.html>

⁴⁵⁶ <http://www.arcom.net/belca/como_esc/c_abreviat.html>

⁴⁵⁷ En Mejía García *et ál.* (1998) se realiza un estudio muy completo sobre las bases de datos terminológicos, que trata la evolución desde sus inicios, las ventajas que presentan estos recursos frente a otros, la tipología que se puede encontrar, los diferentes usuarios que los utilizan, entre otros aspectos. También es importante realizar una distinción entre una base de datos terminológicos y un banco de datos terminológicos, ya que a veces se tratan como sinónimos pero realmente no lo son: por un lado, la base de datos terminológicos hace referencia a la recopilación estructurada y automatizada de unidades terminológicas, acompañadas de información sobre ellas, destinada a responder a las necesidades de un grupo definido de usuarios; y por otro lado, los bancos de datos se refieren al conjunto estructurado de varias bases de datos interrelacionadas de forma lógica, accesibles mediante un lenguaje de acceso, consulta y explotación de datos (Pavel y Nolet, 2002: 116).

⁴⁵⁸ <<http://iate.europa.eu/>>

⁴⁵⁹ <<http://www.unwto.org/trad/index.php>>

⁴⁶⁰ <<http://starwww.uibk.ac.at/dolm/termdb.html>>

⁴⁶¹ <http://www.iim.fh-koeln.de/webterm/deuterm/tourismus/tourismus_d.htmlm/termdb.html>

aquí como recursos terminográficos debido al gran uso que se está haciendo de ellos en la actualidad. Como vimos en el capítulo anterior, estas fuentes consisten en una recopilación estructurada de textos sobre una materia determinada y con unas características comunes y su objetivo principal consiste en mostrar al usuario el lenguaje en uso, es decir, mostrar el contexto de uso de las unidades terminológicas y fraseológicas. Los corpus textuales actuales son fuentes electrónicas, es decir, se pueden consultar a través del ordenador, ya sea a través de Internet o en CD-ROM. Podemos distinguir también corpus públicos, que se pueden consultar libremente, como los corpus *CREA* y *CORDE*, de la Real Academia de la Lengua;⁴⁶² el corpus *COSMAS I*,⁴⁶³ del Institut für Deutsche Sprache, de Mannheim (Alemania), o el British National Corpus;⁴⁶⁴ y los corpus privados, para los que se requiere una clave o una suscripción para su consulta, como el corpus *TURICOR*⁴⁶⁵ y corpus *ad hoc*, creados por traductores, terminólogos, etc. con un objetivo concreto de utilización y en un determinado momento.

2.4.1.2. Según el número de lenguas

Este criterio está presente en todas las clasificaciones consultadas, ya que es el criterio más claro para la distinción de fuentes terminológicas. Así, podemos distinguir fuentes monolingües, bilingües o plurilingües.⁴⁶⁶

La gran mayoría de las fuentes disponibles son monolingües, ya que ofrecen tanto para la entrada como para la definición del término un único idioma. Ejemplos de

⁴⁶² <<http://www.rae.es/rae/gestores/gespub000019.nsf/voTodosporId/D55F5BFB05D63980C1257164003F02E5?OpenDocument&i=1>>

⁴⁶³ <<http://corpora.ids-mannheim.de/ccdb/>>

⁴⁶⁴ <<http://www.natcorp.ox.ac.uk/>>

⁴⁶⁵ TURICOR: Compilación de un corpus de contratos turísticos (alemán, español, inglés, italiano) para la generación textual multilingüe y la traducción jurídica (MCYT, ref.: BFF2003-04616, 2003-2006). <<http://www.turicor.net>>

⁴⁶⁶ No debemos confundir el número de lenguas que incluye un trabajo con el tipo de trabajo en relación a las lenguas. Por el criterio del número de lenguas que un trabajo incluye (criterio que hemos utilizado en nuestra clasificación), podemos distinguir trabajos monolingües, bilingües, trilingües, etc. Por el criterio del tipo de trabajo, distinguimos entre trabajos monolingües, monolingües con equivalencias, bilingües o multilingües, según si todas las fases del trabajo se han realizado en una lengua o en más de una.

estas fuentes son el *Deutsches Universalwörterbuch*, de la editorial Duden (2003), el *Diccionario del uso del español*, de María Moliner, o los diccionarios de Oxford en línea.⁴⁶⁷

Las fuentes bilingües son las que trabajan con dos idiomas, por ejemplo alemán-español o inglés-español. Podemos encontrar varios tipos de fuentes bilingües, pero los más comunes son los siguientes: las fuentes con la entrada en el idioma en cuestión y su equivalente en otro idioma y la entrada con su definición y la equivalencia en otro idioma, como por ejemplo el *Diccionario Pons español-alemán*.⁴⁶⁸

El tercer tipo de fuentes son las fuentes plurilingües, en las que se utilizan más de dos idiomas, por ejemplo alemán, español e inglés o alemán, español, inglés y francés. No habría límite de idiomas en este grupo, aunque cuanto mayor número de idiomas estén involucrados, mayor dificultad habrá a la hora de encontrar equivalencias en todos los idiomas. En general, estas fuentes no ofrecen la definición de la entrada, sino únicamente una lista con las entradas en un idioma y sus equivalentes en los demás idiomas. Un ejemplo de estas fuentes puede ser el *Lèxic de transport turístic*,⁴⁶⁹ de TERMCAT, que ofrece el léxico relacionado con el transporte turístico en español, francés, inglés, alemán, catalán, gallego, entre otras lenguas.

2.4.1.3. Según el tipo de soporte

En las clasificaciones más antiguas (Rey, 1970; Zgusta, 1971; Malkiel, 1975), no aparece este criterio, ya que la posibilidad de consultar fuentes terminológicas en otros formatos distintos del papel ha surgido hace apenas unos años. Actualmente, podemos distinguir dos tipos de soporte básicos: papel y electrónico (DVD, CD-ROM y en línea), el cual es cada vez más frecuente gracias a la difusión de las nuevas tecnologías a todos los usuarios. Anteriormente, también se podían encontrar obras de referencia en formato de microfichas, aunque en la actualidad este formato está prácticamente en desuso.

⁴⁶⁷ <<http://oxforddictionaries.com/?attempted=true>>

⁴⁶⁸ <<http://www.pons.de/>>

⁴⁶⁹ <<http://www.termcat.cat/dicci/transport/index.html>>

Las fuentes en papel son las fuentes tradicionales, que se publican por editoriales y normalmente aparecen los nombres de los autores. Actualmente, se tiende a realizar versiones electrónicas de las fuentes en papel o a publicar a la vez un mismo trabajo en ambos soportes, para aprovechar así las ventajas de los dos soportes. Por ejemplo, los diccionarios de la editorial Oxford se están publicando de esta manera.

Las fuentes electrónicas son relativamente recientes, pero están prosperando rápidamente gracias a las ventajas que presentan con respecto al soporte papel. En este soporte, se puede encontrar todo tipo de fuente terminológica y de cualquier especialidad.

Básicamente, las diferencias entre las fuentes en papel y las electrónicas se pueden resumir en tres: cantidad de información, organización de la información y acceso a la información.

Con relación a la cantidad, la capacidad de almacenamiento de una fuente electrónica es mucho mayor que la de un diccionario en papel. De la misma manera, el contenido de las fuentes electrónicas es mucho más variado, tanto en la macroestructura como en la microestructura (Sánchez Ramos, 2005a: 210).

La organización también es muy diferente entre estos dos soportes, ya que en el soporte papel se puede organizar de las tres formas clásicas que veremos a continuación (alfabética, sistemática e ideológica) mientras que en el soporte electrónico se puede dar la organización de hipertexto además de esas tres modalidades anteriores.

No obstante, el acceso a la información es realmente lo que diferencia un soporte de otro, ya que el formato papel está limitado básicamente a la búsqueda alfabética mientras que el soporte electrónico puede ofrecer varios tipos de búsquedas y combinaciones.

En las fuentes electrónicas, las consultas se pueden realizar mediante diferentes formas de búsqueda: búsqueda por término, que consiste en introducir la unidad que se desea consultar en la ventana de entrada de datos; búsquedas en lista de entradas (alfabética e inversa), que consiste en buscar en una lista de palabras ordenadas alfabéticamente en orden creciente o decreciente; búsqueda asistida, que permite

encontrar las entradas escritas incorrectamente ofreciendo una lista de palabras alternativas; búsqueda múltiple, que permite al usuario realizar búsquedas a partir del contenido de los artículos (definición, ejemplo, etc.); búsqueda mediante palabras relacionadas, que permite utilizar el diccionario como un tesoro y recorrer relaciones léxicas como la sinonimia, la antonimia o familia de palabras; búsqueda de anagramas, que permite buscar todas las palabras que contienen las mismas letras que la palabra dada; búsqueda con comodines, que permite utilizar símbolos (*, #) para reemplazar los caracteres de las palabras; búsqueda por palabras clave de la definición, que permite introducir diferentes palabras que definen la unidad buscada y partir de ellas para buscar el término en cuestión, y búsqueda utilizando abreviaturas y marcas, que consiste en recuperar formas teniendo en cuenta las distintas abreviaturas y marcas e uso que muestran en sus entradas (Sánchez Ramos, 2005a: 213).

Como se puede deducir ante esta variedad de posibles búsquedas en fuentes con formato electrónico, su consulta se hace más cómoda y fácil que las posibles búsquedas en formato papel, ya que estas últimas se reducen básicamente a las búsquedas alfabéticas o semasiológicas.

2.4.1.4. Según la ordenación de sus contenidos

Según las clasificaciones generales, las fuentes terminológicas pueden ofrecer ordenar su contenido de dos formas diferentes: semasiológica o alfabética y onomasiológica.

La ordenación semasiológica o alfabética consiste en ordenar las unidades terminológicas según el orden alfabético. Se trata de una ordenación semasiológica, es decir, va del término al concepto, por lo que se retira claramente del cientifismo que se busca en Terminología o en otras disciplinas (Martínez de Sousa, 2004: 178). Sin embargo, la gran mayoría de las fuentes terminológicas organizan sus contenidos con este orden, ya que son las más aceptadas por los usuarios en general debido a que proporcionan una forma de búsqueda eminentemente práctica.

Por su parte, la ordenación onomasiológica es propia de los tesauros y de algunas enciclopedias, aunque también los diccionarios y otras fuentes terminológicas pueden hacer uso de esta ordenación. Se trata de una ordenación que va del concepto al término y, por tanto, es mucho más científica que la alfabética (Martínez de Sousa, 2004: 178). Esta ordenación está basada en una ordenación del lenguaje en campos léxicos y, así, permite buscar los conceptos y las materias donde se halla recogido el léxico correspondiente. Asimismo, facilita la detección de vacíos y faltas de términos correspondientes a una parte de un todo, así como la colocación de cada término en su categoría, hallar el hiperónimo y sus correspondientes hipónimos y definirlos o tratarlos según su diferenciación semántica. Además de esta parte onomasiológica, normalmente estos recursos incluyen también un índice temático amplio y bien realizado, así como un índice alfabético lo más amplio posible, que permita hallar todos y cada uno de los términos tratados. Como ejemplos de este tipo de ordenación, podemos citar la obra de Roget, *Thesaurus of English Words and Phrases* (1927),⁴⁷⁰ o el *Diccionario temático de la lengua española VOX* (1975).

Como decíamos, estas dos formas son la manera clásica de ordenar el contenido de una fuente de información normalmente en formato papel, pero a día de hoy deberíamos incluir un tercer tipo de ordenación presente en los recursos electrónicos: el hipertexto. El hipertexto consiste en «non-sequentially linked pieces of text or other information» (Nielsen, 1995), es decir, el enlace de información o de un texto de forma no lineal. Mediante el hipertexto, se van creando enlaces dentro de un mismo texto, en este caso una entrada de una fuente terminológica, para crear una estructura asociativa del contenido. Al final, el resultado es una red de enlaces no lineales que permiten acceder a la información que indican de forma rápida y sencilla. Esta ordenación del contenido solo se puede realizar en los recursos electrónicos, ya que en papel el hipertexto no sería posible.

⁴⁷⁰ Desde entonces se han realizado multitud de reediciones. La última tuvo lugar en 1987. No obstante, el diccionario está disponible de forma electrónica en el siguiente enlace: <<http://www.gutenberg.org/etext/10681>>.

Por último, debemos destacar aquí que actualmente la mayoría de los recursos electrónicos, ya sean en línea o en otro formato, no ofrece directamente ninguno de estos tipos de ordenación externa de su contenido, sino que normalmente realiza búsquedas de términos a través de los buscadores disponibles en el recurso. En otras palabras, la ordenación de los contenidos se realiza en la estructura del recurso pero en pocas ocasiones se pone a disposición del usuario, que se limita a introducir el término de búsqueda en el campo dispuesto para ello, a no ser que se trate de recursos que ofrecen una lista de unidades incluidas a las que se accede a través de hipervínculo. Asimismo, además de este tipo de búsqueda por término, se están realizando investigaciones actualmente para ofrecer otros tipos de búsquedas más completas y eficientes, como por ejemplo la búsqueda textual, que ofrecería un acceso a la información terminológica directamente desde el texto de trabajo (cf. 4.3.1.).

En definitiva, nos parece conveniente incluir este criterio de catalogación, puesto que aún nos encontramos con numerosos recursos ordenados por orden alfabético o semasiológico, pero debemos tener en cuenta que existen otras formas de ordenación relacionadas con el tipo de búsqueda que ofrecen los recursos en cuestión.

2.4.1.5. Según la normalización

Dentro de este criterio, podemos distinguir dos tipos de fuentes terminológicas: las fuentes normativas (o prescriptivas) y las fuentes descriptivas.

Las fuentes normativas pretenden fijar un determinado uso de un idioma. Para ello, recogen las unidades léxicas y acepciones reconocidas por las autoridades académicas, normalmente las Academias de la Lengua. De esta forma, parten de un determinado ideal de lengua considerado como norma y son muy reacios ante incorporaciones nuevas de vocabulario como neologismos y extranjerismos. Estas fuentes son consideradas oficiales del idioma en cuestión y, por tanto, son el referente de todas las instancias oficiales y profesionales. Por lo general, estas fuentes normativas son fuentes lexicográficas, es decir, recogen las unidades léxicas de la lengua común (por ejemplo, el *Diccionario de la Real Academia de la Lengua*, en España). No

obstante, también podemos encontrarnos algunas fuentes normativas y terminológicas dentro de algunos campos de especialidad, por ejemplo el DSM-IV (*Diagnostic and Statistical Manual of Mental Disorders*) en la medicina psiquiátrica, las publicaciones realizadas por la Real Academia de Ciencias Exactas, Físicas y Naturales⁴⁷¹ o las publicadas por la International Electrotechnical Commission (IEC).⁴⁷²

Las fuentes descriptivas, por el contrario, pretenden recoger y reflejar la realidad del lenguaje de especialidad con sus neologismos, extranjerismos, tecnicismos, etc., es decir, describen el uso real del idioma. Estas fuentes recogen una selección representativa del léxico utilizado realmente y sin restricciones. En general, las fuentes que no son académicas son descriptivas, ya que la categoría de normativo se la otorga la institución que lo edita. Por ejemplo, en España, el único diccionario normativo sería el DRAE y el resto de fuentes serían descriptivas solamente. En algunos países como en Alemania o Reino Unido no existe una Academia oficial de la Lengua, por lo que no se podría considerar en rigor ninguna fuente como normativa. Sin embargo, en el caso de Alemania, el diccionario *Duden* sería el principal referente de norma; y en el caso del inglés británico, los diccionarios publicados por la Oxford University Press, en concreto el *Oxford English Dictionary* (OED), son reconocidos como autoridad.

Además de las fuentes normativas nombradas con anterioridad, podemos incluir como referentes también en este grupo a las normas internacionales publicadas por organizaciones de normalización, como *ISO* (*Internacional Organization for Standardisation*),⁴⁷³ *CEN* (*The European Committee for Standardization*),⁴⁷⁴ *AENOR* (*Asociación Española de Normalización y Certificación*),⁴⁷⁵ *DIN* (*Deutsches Institut für Normung*),⁴⁷⁶ entre otras.

⁴⁷¹ <http://www.rac.es/0/0_1.asp>

⁴⁷² <<http://www.iec.ch/>>

⁴⁷³ <<http://www.iso.ch/iso/en/ISOOnline.frontpage>>

⁴⁷⁴ <<http://www.cenorm.be/cenorm/index.htm>>

⁴⁷⁵ <<http://www.aenor.es>>

⁴⁷⁶ <<http://www.din.de/>>

2.4.1.6. Según el grupo de destinatarios

Este criterio se considera a menudo difícil de determinar, especialmente una vez que el diccionario está elaborado. En general, se pueden distinguir tres tipos de usuarios de una fuente de información terminológica, según su nivel de conocimiento: especialistas, estudiantes y legos.

Al igual que ocurría con los diferentes niveles de abstracción en los textos especializados, las fuentes terminológicas adaptarán la utilización de su lenguaje a sus destinatarios. En este sentido, las obras que vayan dirigidas a legos en la materia proporcionarán una definición más extensa y con un nivel semántico inferior, como el *Deutsches Rechtswörterbuch (Drw)*⁴⁷⁷ o el *Diccionario jurídico básico*, de Martínez de Navarrete (2008); mientras que si la obra va dirigida a especialistas, el lenguaje utilizado será más escueto, conciso y preciso y no se requerirán definiciones extensas, como por ejemplo el *Diccionario Jurídico*, de la editorial Aranzadi (2001).

A pesar de incluir este criterio en nuestra clasificación, consideramos que no proporciona información pertinente y fiable para catalogar una fuente terminológica, ya que la diferencia entre una obra dirigida a un público o a otro no se hace realmente perceptible *a posteriori*, es decir, una vez que la obra está ya terminada, a no ser que se especifique en la misma obra, ya sea en el prólogo, instrucciones de uso, introducción, etc.

2.4.1.7. Según el tipo de unidades terminológicas

Las unidades terminológicas que aparecen en una fuente terminológica determinan de qué tipo de fuente se trata. En este sentido, podemos distinguir dos tipos básicos: fuentes sintagmáticas y fuentes paradigmáticas.

⁴⁷⁷ <<http://drw-www.adw.uni-heidelberg.de/drw/>>

Las fuentes sintagmáticas tratan las unidades terminológicas en su aplicación contextual, la cual se ha descuidado normalmente en la mayoría de las fuentes de información terminográfica. Podemos distinguir varios tipos según su contenido:

- Diccionarios de colocaciones, como por ejemplo el Diccionario combinatorio práctico del español contemporáneo, de Bosque (2008); el diccionario *DICE (Diccionario de Colocaciones del Español)*;⁴⁷⁸ el *Kollokationen in der Wortschatzarbeit*, de Reder; el Dornseiff: *Der Deutsche Wortschatz nach Sachgruppen*, de Dornseiff; el *Oxford Collocations Dictionary for Students of English*; el *BBI Dictionary of English Word Combinations*, etc.
- Diccionarios de locuciones o unidades fraseológicas, como es el caso de las obras *Spanische idiomatische Redewendungen*, de Florian y Martínez (1985); la obra *Expresiones idiomáticas. Lexikon der spanischen Redewendungen*, de Moreno Torres (1990); el *Cambridge Idioms Dictionary* (2006); el *Oxford Dictionary of Idioms 2.01.*, etc.
- Diccionarios de refranes, de citas y frases, como el *Diccionario de refranes*, de Carbonell Basset (1996); el *Oxford Dictionary of Quotations* (1992); el *Diccionario de citas literarias*, de Guerrero del Río (2002); la *Guía práctica español-inglés* (1997), y el diccionario *Vox Spanish Phrasebook and Dictionary* (2002).

En cada una de estas obras, encontraremos información sobre el uso contextual de sus entradas que nos ayudarán a utilizar la unidad léxica en su contexto correcto.

El otro tipo de fuentes, las paradigmáticas, reúnen las unidades léxicas en paradigmas que pueden ser de contenido (como en los casos de las fuentes sistemáticas o ideológicas, los diccionarios de sinónimos y de antónimos y los diccionarios visuales) o de expresión (como en los diccionarios de homónimos y parónimos y diccionarios de rimas). Algunos ejemplos de obras paradigmáticas según su contenido pueden ser los siguientes:

⁴⁷⁸ <<http://www.dicesp.com/paginas>>

- Diccionarios de antónimos y sinónimos. Como ejemplos de estos diccionarios, podemos citar el *Diccionario de sinónimos y antónimos*, de la editorial Santillana (2008); el *Diccionario práctico de sinónimos, antónimos y parónimos*, de la editorial Everest (2005); *Dudden 8: Die Sinn- und Sachverwandten Wörter: Synonymwörterbuch der Deutschen Sprache*, de Müller (2002); el *Oxford Dictionary of English Synonyms and Antonyms* (2007), o el *Scholastic Dictionary of Synonyms, Antonyms & Homonyms* (2001).
- Diccionario de rimas, como son el *Oxford Rhyming Dictionary* (2004) o el *Diccionario de Rimas* de Ediciones Cardedoso (2005).
- Diccionario de gestos e imágenes. Como ejemplos, podemos citar el *Diccionario de gestos españoles* (2007), *The-onverbal Dictionary of Gestures, Signs and Body Language Cues*, de Givens (2008), o el Glosario de Iconos turísticos, de IberiaRural.⁴⁷⁹

La distinción entre estos dos tipos de fuentes pertenece realmente a la lengua común, es decir, se trata de fuentes lexicográficas y no terminológicas; no obstante, consideramos pertinente esta clasificación, ya que las fuentes terminológicas suelen ser fuentes paradigmáticas, pero también podemos encontrarnos con fuentes híbridas, es decir, que contengan información paradigmática y sintagmática. De hecho, sería recomendable que todas las fuentes de información terminográfica (y también lexicográfica) incluyeran información sintagmática, con objeto de mostrar a los usuarios el uso real de sus entradas.

2.4.1.8. Según la extensión de la especialidad

Según la representación que se haga de un campo de especialidad, podemos clasificar las fuentes terminológicas en generalistas o en concretas.

⁴⁷⁹ <<http://www.iberiarural.es/glosario.htm>>

Las fuentes generalistas pretenden abarcar todo el campo de especialidad en el que se está trabajando, es decir, tratan el lenguaje de especialidad como un todo y no distinguen las diferentes especialidades, actividades, etc. que se puedan distinguir dentro del ámbito de especialidad. En el caso del lenguaje del turismo, se trataría de una fuente terminológica que incluyera términos de todo el campo de conocimiento en general. Este es el caso del *Diccionario de términos de turismo y de ocio*, de Alcaraz Varó et ál. (2006), donde se trata el turismo en general, sin centrarse en un ámbito concreto.

Por el contrario, las fuentes terminológicas concretas se centran en una parte concreta de la especialidad, ya sea un segmento turístico, una especialidad de medicina, una actividad económica, etc. De esta forma, se consigue una fuente terminológica aún más especializada que la fuente generalista, ya que delimita aún más su campo de actuación. En el ámbito del turismo, por ejemplo, podríamos hablar de una fuente que se centra en el segmento turístico de turismo de aventura o en la actividad de montañismo. Estas fuentes son muy abundantes en la actualidad, especialmente en formato electrónico, y pueden tratar de cualquier aspecto de un lenguaje de especialidad. Una fuente de estas características sería, por ejemplo, el *Glosario de montaña y geografía*, de Montipedia⁴⁸⁰ o el *Lexikon von Fachbegriffen rund um's Klettern*.⁴⁸¹

2.4.1.9. Según la fuente de información utilizada

Este criterio es muy importante a la hora de evaluar las fuentes terminológicas, ya que nos indica la procedencia de la información que se ha seleccionado. Según este criterio, la elaboración de las fuentes terminológicas puede estar basada fundamentalmente en dos tipos diferentes de fuentes: corpus, ya sean orales o escritos, y materiales de otras fuentes terminológicas.

⁴⁸⁰ <<http://www.montipedia.com/glosario.htm>>

⁴⁸¹ <<http://www.on-sight.de/diesdas/index.html?glossar.html>>

La utilización de corpus para la creación de fuentes terminológicas está muy extendida en la actualidad, aunque su incorporación a la terminografía ha sido relativamente reciente. El uso creciente de esta técnica se debe principalmente a la utilización de corpus electrónicos, gracias a la incorporación de las nuevas tecnologías de la información y comunicación (TIC), tanto en el uso de ordenadores e Internet como en la aplicación de herramientas de almacenamiento y de extracción de términos. Tradicionalmente, se utilizaban corpus en papel, es decir, toda la información que se utilizaba para elaborar una fuente terminológica tenía que ser recopilada en papel. Como se puede deducir, el manejo y el almacenamiento de la información eran bastante laboriosos, especialmente si los comparamos con los actuales corpus electrónicos.

Los corpus utilizados para la elaboración de fuentes terminológicas pueden ser orales, formados por la recopilación de actos comunicativos de la lengua hablada, o escritos, constituidos por la recopilación de textos escritos ajustados a unas características concretas. A veces, también se pueden encontrar corpus mixtos, es decir, que se incluyen enunciados orales y textos escritos en un mismo corpus. Esta técnica está muy extendida en la actualidad, ya que permite observar el lenguaje en su contexto y almacenar gran cantidad de datos en un espacio muy reducido, lo que garantiza la selección de material real y utilizado por los usuarios reales.

La elaboración de fuentes terminológicas basadas en materiales de otras fuentes puede ser también de calidad, aunque carece de originalidad y novedad por lo general. Con esta técnica, los autores de la nueva fuente terminológica no realizan investigaciones a fondo sobre el uso de esos términos incluidos en las obras de consulta, por lo que requiere gran confianza en las obras consultadas. De todas maneras, los investigadores deben comprobar si las unidades terminológicas han caído en desuso, si han cambiado de registro, si tiene nuevas acepciones y si la finalidad de la fuente consultada se ajusta a los objetivos de la nueva fuente. Esta técnica abunda especialmente en la elaboración de fuentes lexicográficas monolingües, aunque también existen dentro de las fuentes terminológicas, principalmente en proyectos donde existe falta de financiación.

2.4.1.10. Según la finalidad

Con este criterio, hemos querido diferenciar dos tipos de aplicaciones que se le pueden dar a una fuente terminológica: la codificación y la descodificación de la información.

La codificación de información consiste en la escritura del mensaje codificado con un lenguaje de especialidad propio de un campo de conocimiento, en nuestro caso el lenguaje turístico. Las obras que tienen esta aplicación tienen como objetivo mostrar y facilitar la forma de expresión de un lenguaje determinado, mediante una indicación adecuada acerca de sus características sintácticas, fraseológicas y terminológicas. Estas obras ofrecen una microestructura muy útil y completa, ya que ofrecen información gramatical, de contexto, indicaciones de frecuencia y de uso, notas aclaratorias, sinónimos, etc., es decir, toda la información que se considere pertinente para la expresión escrita de un texto. Como ejemplo de este tipo de fuentes, podemos nombrar las obras monolingües, la mayoría de las bases de datos terminológicas, etc.

Por el contrario, la descodificación tiene como objetivo ayudar a la comprensión de un texto, es decir, a entender perfectamente qué dice el texto, ya sea en el mismo idioma o en otro idioma. Su labor consiste, por tanto, en proporcionar una definición del término que se haya consultado, para aclarar su significado y permitir su comprensión, o proporcionar un equivalente, ya sea en forma de sinónimo (en la misma lengua) o de traducción (en otra lengua). De esta forma, la fuente permite comprender el término consultado proporcionando una aclaración acerca de dicho término. En estas obras, no se suelen incluir ejemplos de uso, indicaciones de frecuencia, campo de aplicación, etc., ya que su objetivo es simplemente facilitar la comprensión de un término. Las obras que cumplen con las características para la descodificación son, principalmente, las fuentes bilingües o plurilingües.

Una vez que hemos establecido la clasificación general de las fuentes terminológicas que se pueden encontrar en cualquier ámbito de especialidad, continuaremos con una propuesta de evaluación de dichas fuentes, tan necesaria a veces para realizar una criba de todo el material disponible en la actualidad, especialmente en

el caso de los traductores, para quienes la competencia documental es una parte esencial en su formación y en su profesión, como veíamos anteriormente.

2.4.2. EVALUACIÓN DE FUENTES DE INFORMACIÓN TERMINOGRÁFICA

Je weniger ein Wörterbuch nur 'Wörter'buch ist, desto hilfreicher ist es für den Translator (Vermeer, 1989: 172).

El traductor requiere unas fuentes de información terminográficas que le ofrezcan las unidades terminológicas más adecuadas para la traducción que está realizando y, para ello, debe conocer las fuentes de información existentes y disponer de criterios para evaluarlas convenientemente. Es decir, es necesario tener «una actitud crítica», como nos decía Haensch (1997: 237) ante la gran cantidad de fuentes terminológicas disponibles.

A pesar de la necesidad de esta actitud crítica, la investigación que se ha realizado sobre la evaluación de fuentes de consulta es todavía muy parca y carece de unos criterios claros y generalmente aceptados.

Hasta el momento, se han realizado algunos intentos de unificar estos criterios de evaluación, especialmente para la realización de reseñas de diccionarios (Steiner, 1984; Nakamoto, 1994), pero su aplicación no se ha extendido ni se ha recibido la aceptación por parte de los especialistas en este campo. Nakamoto, en su tesis doctoral (1994), establece unos criterios para la evaluación de diccionarios dirigidos a estudiantes de lengua extranjera en forma de listas de control (*checklists*), con objeto de homogeneizar y de completar la realización de reseñas de diccionarios. Estas listas contienen unas casillas de verificación sobre la macroestructura y la microestructura de diccionarios, así como sobre el objetivo y el autor de la reseña. Con respecto a la macroestructura, se pregunta acerca de los datos de la publicación y del autor, los contenidos, la organización, etc.; y, con respecto a la microestructura, se cuestiona la estructura de las entradas, las categorías gramaticales, ilustraciones y referencias cruzadas.

Estas listas ofrecen la posibilidad de valorar una fuente de información de forma exhaustiva, ya que tratan todos los aspectos de dichas fuentes (macroestructura, autoría, microestructura). Sin embargo, consideramos que es excesivamente extensa y, por tanto, poco práctica para realizar evaluaciones pertinentes sobre las fuentes que se estén utilizando, o por lo menos para que estas evaluaciones sean realizadas por los usuarios de las fuentes en cuestión. Además, incluyen preguntas sobre el autor de la reseña y el lugar de su publicación, cuyos datos son irrelevantes para la actividad traductora.

Por lo general, las listas de evaluación de recursos están dirigidas a expertos que realizan una evaluación sobre una fuente de información determinada o, a lo sumo, una comparativa entre varias fuentes en forma de reseña, y así se determina en el mismo título del trabajo *Establishing Criteria for Dictionary Criticism: A Checklist for Reviewers of Monolingual English Learners' Dictionaries* (Nakamoto, 1994). Sin embargo, el traductor profesional debe valorar multitud de fuentes de información y, por tanto, necesita una plantilla mucho más manejable y concisa, que contenga la información pertinente para evaluar si una fuente de información es fiable o no. Para conseguir esta plantilla, hemos tenido que seguir investigando sobre los diferentes estudios que se han realizado en torno a este tema, es decir, sobre los elementos básicos que debe tener una obra de referencia para que tenga utilidad para un traductor, además de tener en cuenta las necesidades de estos a la hora de consultar los recursos terminográficos.

Muchos autores han trabajado también sobre el tema de la evaluación de fuentes de información y han realizado sus investigaciones basadas fundamentalmente en estudios de campo, ya fueran mediante encuestas o pruebas empíricas de otra naturaleza. En general, la mayoría de estas investigaciones ha estado dirigida fundamentalmente al estudio del uso de diccionarios por parte de los estudiantes de segundas lenguas (Béjoint, 1981; Ard, 1982; Hartmann, 1982, 1983, 1995; Bogaards, 1988; Cumming et ál., 1994; Kernerman, 1996; Atkins y Varantola, 1998; East, 2008), aunque también podemos encontrar algunos trabajos sobre el uso de diccionarios por parte de estudiantes de traducción (Mackintosh, 1998; Varantola, 1998; Corpas Pastor et ál., 2001; Sánchez Ramos, 2005b; Pujol et ál., 2006), un grupo de usuarios que

presentan unas características y necesidades diferentes a los estudiantes de segundas lenguas.⁴⁸² Estas investigaciones han tenido como objetivo estudiar las necesidades, los hábitos y las preferencias de los usuarios, determinar el éxito o el fracaso de las búsquedas en diccionarios y, en general, estudiar y controlar el uso de los diccionarios por parte de los usuarios.⁴⁸³ En otras palabras, podemos decir que pretenden detectar las fortalezas y las debilidades de los diccionarios así como estudiar la forma en que los alumnos de traducción los utilizan.

A partir de estos estudios, podemos extraer conclusiones sobre los puntos positivos y negativos que presentan las fuentes estudiadas y sobre la información pertinente que debe incluir un diccionario o cualquier otro recurso con objeto de facilitar su labor al traductor. Así, podemos ir estableciendo criterios necesarios para crear nuestra plantilla de evaluación.

Una de las conclusiones más importantes que se obtiene de estos estudios es el hecho de no poder determinar en términos absolutos si una fuente es buena o no, ya que siempre dependerá de factores como la tarea en cuestión, el tipo de unidad léxica, el contexto de la unidad, el conocimiento del usuario, la función de la tarea y las habilidades de búsqueda (Hulstijn y Atkins, 1998: 11). En este sentido, cada traductor deberá decidir qué fuente utilizar para qué y en qué momento a la hora de realizar una traducción de textos especializados de cualquier campo de conocimiento. No obstante, esta afirmación no hace referencia a la calidad en sí de una fuente de información, que será lo que podamos valorar con nuestra plantilla de evaluación.

De la anterior conclusión se desprende otra muy importante con respecto a los usuarios, en nuestro caso los traductores. Esta idea hace referencia a la falta de conocimientos que presentan a menudo al utilizar un diccionario u otra fuente de consulta. Este desconocimiento puede deberse principalmente a dos razones: una falta

⁴⁸² Algunas de estas investigaciones se tratarán con mayor profundidad en el apartado 2.5.1., cuando se haga un repaso sobre los diferentes estudios que se han llevado a cabo sobre el uso de los diccionarios por parte de los traductores, estudiantes de traducción o de segundas lenguas.

⁴⁸³ Para profundizar en este tema, véase el trabajo de Atkins y Hulstijn «Empirical research on dictionary use in foreign-language learning: survey and discussion» (1998). En este artículo, se realiza un estudio sobre cincuenta trabajos que tratan de establecer el uso de los diccionarios por parte de los estudiantes de lengua extranjera.

de habilidad a la hora de realizar búsquedas, que sería debida al usuario, o a una falta de información sobre el uso de la obra, que sería debida a la obra de consulta en sí. Con respecto a la primera causa, Roberts (1997: 50) nos indica lo siguiente:

[...] translators have not learned to use their dictionaries efficiently. Dictionary use is not an integral part of school and university curricula. And even schools of translation and interpretation rarely have courses specifically designed to teach dictionary use.

En este sentido, el traductor no habría aprendido realmente a utilizar los recursos que tiene a su alcance y, por tanto, no haría un buen uso de ellos. Consecuentemente, sus resultados no serían tan positivos ni su labor tan eficiente. No obstante, esta carencia se podría subsanar aprendiendo las técnicas de búsqueda, conociendo los diferentes tipos de fuentes disponibles y seleccionando adecuadamente los materiales utilizados.

La segunda causa, en cambio, es debida a la inadecuación de la información de las fuentes de consulta, es decir, a la ausencia de indicaciones pertinentes para llevar a cabo búsquedas eficientes en la fuente de consulta. En toda obra de consulta, es recomendable incluir unas indicaciones acerca de su funcionamiento, de su organización, de las unidades incluidas, etc. para que el usuario pueda sacar el máximo partido de dicha fuente.

La primera causa no puede influir en la evaluación de una fuente de información, ya que es un problema extrínseco a la fuente; sin embargo, la segunda causa sí es un problema de la obra de consulta y, como tal, debe ser tratado como un criterio de evaluación para determinar la calidad de una obra.

Otra conclusión importante se desprende del estudio realizado por Mackintosh (1998), que expone la necesidad de incluir definiciones y equivalentes en las fuentes de consulta para los traductores. En su trabajo, la autora nos presenta una encuesta realizada a estudiantes de traducción de L2 a L1, en la que pregunta las razones más habituales por las que consultan los diccionarios (*Most common reasons for consulting dictionaries*). De sus respuestas, se deriva el tipo de búsquedas que realizan los alumnos así como la información que pretenden encontrar. Con los resultados, se observa la necesidad que tienen los alumnos de consultar la definición de un término para

comprender su significado, así como la necesidad de buscar equivalentes a la hora de traducir en traducción directa. De esta forma, se hace patente la necesidad de combinar definiciones y equivalentes para facilitar la labor traductora, además de para garantizar un buena comprensión del texto original por parte del traductor.

Cabré Castellví (1999: 37) también nos presenta sus conclusiones con respecto a la evaluación de la utilidad y fiabilidad de las fuentes terminológicas. Según la autora, es importante que estas fuentes presenten una serie de características y de datos para que se puedan considerar fiables y útiles. A continuación, presentamos las indicaciones que consideramos más pertinentes a la hora de realizar la evaluación. Las fuentes terminológicas deberían:

- Explicitar el método de recopilación de sus datos
- Estar avaladas por especialistas o instituciones representativas;
- Aportar información sobre el valor conceptual de las unidades terminológicas dentro de su disciplina;
- Proporcionar información detallada sobre los conceptos, su uso y su aspecto gramatical;
- Ofrecer contexto para las unidades de equivalencia;
- Incluir indicaciones sobre la fuente de los términos y sobre su nivel de fiabilidad.
- Representar remisiones de sinonimia jerarquizadas correctamente de acuerdo con su valor de uso.
- Ofrecer una consulta fácil.

Por último, nos parece importante incluir en este apartado una mención a la fraseología en las fuentes terminológicas. Los autores que hemos consultado para analizar la forma de evaluar este tipo de fuentes no hacen referencia a la fraseología dentro de estos recursos, aunque sí se han publicado abundantes trabajos muy interesantes y completos sobre el tratamiento y la inclusión de fraseología en los diccionarios (Corpas Pastor, 2003; Wotjak: 2005; Sosinski, 2006, entre otros). En nuestra opinión, una fuente de

información, tanto lexicográfica como terminológica, debería incluir siempre la fraseología de su ámbito de aplicación,⁴⁸⁴ ya que es la mejor forma de garantizar la calidad de las traducciones. A pesar de esto, las fuentes lexicográficas y las terminológicas presentan, por lo general, fuertes deficiencias en lo que se refiere a la inclusión de unidades fraseológicas (Corpas Pastor, 2003: 155). Con lo que respecta a nuestra investigación, consideramos imprescindible que las fuentes terminológicas del sector turístico incluyan fraseología y, por este motivo, la presencia de fraseología en las fuentes terminológicas constituirá uno de los criterios de nuestra plantilla de evaluación.

Este apartado lo terminaremos con una cita de Williams (1996: 297), donde nos resume los elementos más relevantes que deben estar presentes en una obra de consulta:

To produce a ‘smooth’ translation, the translator requires more than lexical equivalents for slot-filling. Clear labelling, more multiple compounds, specific collocations, synonyms at different levels of formality, and grammatical information (prepositions) could all usefully be incorporated into the bilingual dictionary to make it user-friendly for translators.

2.4.2.1. Plantilla de evaluación. Parámetros e indicadores

Los motivos que llevan a la necesidad de evaluación de las fuentes de información ya han quedado manifiestos en los apartados anteriores de este capítulo, así como la necesidad de elaborar una plantilla de evaluación útil y práctica que permita realizar valoraciones objetivas sobre las diferentes fuentes disponibles. Como vimos en el apartado anterior, las evaluaciones de obras de consulta, ya sean diccionarios, glosarios, bases de datos, etc. vienen realizadas normalmente por expertos lexicógrafos y su objetivo es fundamentalmente la realización de reseñas para su publicación en revistas especializadas. De esta forma, los criterios y las plantillas de evaluación existentes hasta el momento están destinados a estos expertos en lexicografía y a la

⁴⁸⁴ El DRAE define «Fraseología» como el «Conjunto de frases hechas, locuciones figuradas, metáforas y comparaciones fijadas, modismos y refranes, existentes en una lengua, en el uso individual o en el de algún grupo».

consecución de este fin, con objeto de mostrar una valoración pericial a los usuarios finales de estas obras de referencia.

En nuestra opinión, este cometido de evaluación es imprescindible a la hora de establecer unas fuentes como prioritarias y recomendadas, así como para realizar comparativas entre varias fuentes. Sin embargo, estos criterios de evaluación quedan alejados de los usuarios reales, que son realmente los que utilizan las fuentes de información y lo que deberían realizar las evaluaciones y, por tanto, no se acogen a la realidad y a la práctica de la tarea diaria. Por este motivo, consideramos necesario la elaboración de una herramienta de evaluación que les permita a estos usuarios distinguir las fuentes de información fiables de las que no ofrecen la fiabilidad suficiente.

Después de haber realizado una investigación exhaustiva de los diferentes tipos de plantillas de evaluación disponibles en la actualidad, y que se encuentran a nuestro alcance, hemos considerado adecuado dividir estas plantillas de evaluación en dos tipos principalmente: por un lado, las plantillas de evaluación que se emplean para elaborar reseñas lexicográficas, que comentábamos en el apartado anterior y que están destinadas a la valoración exhaustiva de una obra para su publicación en revistas o en otros medios por parte de expertos lexicógrafos; y por otro, las plantillas de evaluación de recursos electrónicos, cuyo objetivo es la valoración cuantitativa o cualitativa de los diferentes recursos disponibles en la Red.

Para este último caso, hemos encontrado varios modelos (Kirk, 1996; Codina, 2000; Larsen, 2001; Palomares Perraut et ál., 2004; Gonzalo García et ál., 2004; Seghiri Domínguez, 2006, entre otros),⁴⁸⁵ que, a pesar de presentar diferentes criterios de evaluación y sistemas de valoración, comparten el objetivo de evaluar los recursos electrónicos para seleccionar los más fiables que permitan al usuario garantizar un uso adecuado de la información que contienen. De la misma manera, permiten que el traductor pueda seleccionar y filtrar los recursos electrónicos más fiables para realizar una traducción de calidad y acorde con el ámbito de especialidad en el que se encuadre el texto para traducir. Estas plantillas evalúan de forma general todas las fuentes documentales disponibles en Internet y los resultados obtenidos son bastantes positivos,

⁴⁸⁵ Para una mayor profundización sobre este tema, se puede consultar Seghiri Domínguez (2006: 75-86).

ya que permiten distinguir fácilmente los recursos electrónicos que ofrecen una mayor fiabilidad. Sin embargo, y a pesar de estos resultados, consideramos que no son suficientes para realizar una evaluación de fuentes terminológicas, las cuales son objeto de nuestro estudio en la presente tesis doctoral. Asimismo, las plantillas de evaluación existentes solo evalúan recursos electrónicos, por lo que los recursos en papel no se incluirían. De esta forma, para la evaluación de fuentes terminológicas, tanto electrónicas como en papel, nos vemos obligados a crear una plantilla de evaluación concreta que incluya criterios adecuados con los que se puedan evaluar los aspectos característicos de estas fuentes de referencia de forma objetiva y homogénea.

Nuestra plantilla está principalmente dirigida a traductores, profesionales o estudiantes de traducción como usuarios de fuentes de información lexicográfica y terminológica, aunque también se podría utilizar por otros usuarios sin esta formación específica, puesto que incluye unos criterios fáciles de manejar y de evaluar. Asimismo, es una plantilla abierta que permite la evaluación de fuentes tanto en papel como en formato electrónico; fuentes monolingües o bilingües; fuentes onomasiológicas o semasiológicas, etc. Una vez establecida nuestra plantilla de evaluación, nuestro objetivo será la evaluación de las fuentes terminológicas disponibles tanto en papel como en formato electrónico del sector turístico, especialmente del turismo de aventura, a modo de ejemplo de aplicación. Para ello, en el siguiente capítulo, aplicaremos esta plantilla para realizar una evaluación exhaustiva de las fuentes terminológicas que podemos encontrar actualmente en este sector.

Por lo general, las plantillas de evaluación parten de una serie de criterios que se aplican a las fuentes de información en cuestión. Estos criterios se ponen en práctica mediante el uso de «parámetros» e «indicadores» de evaluación, entendiendo por «parámetros» los aspectos generales que son objeto de evaluación, y por «indicadores» los elementos que desarrollan cada uno de los parámetros establecidos para el análisis de la información (Merlo Vega, 2003). Es decir, los parámetros constituirán los grandes bloques sobre los que se realiza el análisis y los indicadores harán referencia a las cuestiones concretas que se evaluarán dentro de esos grandes bloques.

Para nuestro trabajo, consideramos que existen dos indicadores establecidos imprescindibles a la hora de realizar una evaluación objetiva de cualquier recurso, a saber: la autoría de la obra y su fecha de publicación o de actualización, dependiendo de si el recurso se encuentra en formato papel o electrónico.

Con respecto a la *Autoría*, es esencial que la fuente de información en cuestión recoja al menos el nombre de la persona, de la institución o de la empresa que la creó. Asimismo, también sería recomendable que aparecieran sus credenciales y sus datos de contacto, para aportar una mayor fiabilidad al recurso. En otras palabras, cuanto más información se aporte sobre el autor o la institución que haya realizado el recurso, más fiabilidad aportará este al usuario. Además, en el caso de las fuentes electrónicas, se toma también como elemento de referencia el dominio, que indica la procedencia de las páginas web o su autoría, tanto con respecto al país (*.es* de España, *.de* de Alemania, *.uk* de Reino Unido, *.it* de Italia, etc.) como a la institución (*.edu*, *.gov* o *.eu* para instituciones de educación, organizaciones gubernamentales, Unión Europea, etc.), lo que supone un alto grado de fiabilidad y calidad.

La *Fecha de publicación* de la fuente o su *Actualización* son datos primordiales para comprobar si la fuente en cuestión es actual o no. Como decíamos, la distinción entre estos dos parámetros es debido al hecho de que nuestra plantilla tiene como objetivo la evaluación de fuentes en formato electrónico así como en papel. De esta manera, para las fuentes en papel será relevante la fecha de publicación, pero para las fuentes electrónicas nos tendremos que basar en su fecha de actualización, ya que es poco relevante su fecha de creación si, posteriormente, han sufrido actualizaciones.

En lo que respecta a las fuentes electrónicas, puede ocurrir que no aparezcan estos datos, por lo que en este caso el usuario debe ser cauto en cuanto a la actualidad del contenido. No obstante, la comprobación del correcto funcionamiento de los distintos enlaces que contenga la página se puede considerar como un síntoma de fiabilidad (Seghiri Domínguez, 2006: 82).

Junto a los indicadores anteriores, también incluiremos un tercer indicador evaluable: *Especialización en* En este apartado, se deberá especificar si la obra terminológica trata sobre el dominio de interés en cuestión, en este caso el turismo y el

turismo de aventura, o si, por el contrario, trata de otro dominio de especialidad pero incluye términos pertenecientes al dominio de interés. Así pues, en el espacio disponible tras el criterio *Especialización en* se especificará el campo de interés en cuestión y se evaluará según la proximidad y los términos encontrados en la obra que se desea evaluar.

Estos tres indicadores evaluables a saber: *Autoría*, *Fecha de publicación* o de *Actualización* y *Especialización en*, constituyen el primer bloque de la plantilla de evaluación que hemos denominado *Descripción* y están acompañados por otros tres indicadores que ayudan a la descripción de la obra: el *Título*, las *Lenguas* utilizadas y el *Soporte*. Este bloque permite describir o situar la fuente de información en cuestión dentro de la clasificación que elaboramos en el apartado 2.4.1.⁴⁸⁶ A excepción de los indicadores de *Autoría* y de *Fecha de publicación* o de *actualización*, los demás indicadores de este bloque no nos servirán para evaluar la fuente de información, sino más bien para describir su naturaleza y funcionalidad, es decir, para contextualizarla.

De esta forma, el primer bloque que vamos a considerar para nuestra evaluación de fuentes de información lexicográfica y terminológica quedaría como se indica en la siguiente tabla:

| DESCRIPCIÓN | |
|------------------------------------|--|
| Título | |
| Autoría URL (si procede) | |
| Fecha de publicación/actualización | |
| Lenguas | <input type="checkbox"/> Monolingüe <input type="checkbox"/> Bilingüe <input type="checkbox"/> Multilingüe |
| Soporte | <input type="checkbox"/> Papel <input type="checkbox"/> Electrónico |
| Especialización en | <input type="checkbox"/> Sí <input type="checkbox"/> No |

Tabla 7. Indicadores del primer bloque de la plantilla de evaluación.

⁴⁸⁶ La clasificación de las fuentes de información terminológica del apartado 2.4.1. estaba basada en diez criterios, a saber: la presentación, el número de lenguas, el tipo de soporte, la ordenación de sus contenidos, la función, el grupo de destinatarios, el tipo de unidades terminológicas, la extensión en la especialidad, la fuente de información utilizada y la aplicación de la fuente.

Además de este primer bloque, nuestra plantilla está dividida en otros dos bloques, a saber: macroestructura y microestructura, que coinciden con las dos partes principales en que se divide la estructura de una fuente lexicográfica. Por macroestructura, entendemos el diseño general que ofrece la fuente en cuestión, y por microestructura, el conjunto de informaciones que se incluyen en cada entrada de una fuente.

Dentro del parámetro de la macroestructura, nos encontramos con varios indicadores, que nos permitirán evaluar la fiabilidad y la calidad de nuestras fuentes.⁴⁸⁷

En primer lugar, encontramos el indicador de *Indicaciones*. Este indicador hace referencia a la información que la fuente ofrece sobre sí misma y sobre su uso, es decir, sobre el método utilizado para la compilación de la obra y la selección de las entradas, sobre la fuente original utilizada,⁴⁸⁸ las abreviaturas utilizadas, su organización, etc. En las fuentes en papel, estos datos se encuentran normalmente en la Introducción o en el Prólogo; sin embargo, en las fuentes electrónicas, estos datos no suelen aparecer, lo que dificulta la consulta de dichos recursos. En este indicador, además de evaluar si se incluyen o no estas indicaciones, podemos evaluar también la calidad y la complejidad de estas indicaciones, es decir, cuanta más información nos aporten estas indicaciones, mejor será la valoración.

Otro indicador que consideramos imprescindible para la evaluación de las obras de consulta es el *Número de entradas* que contiene la obra en cuestión. En las obras en papel, este indicador no presenta un problema, ya que suelen tener un número de entradas muy elevado debido al coste que supone la creación y la publicación de una obra terminológica. Sin embargo, sí supone un problema al evaluar las obras en formato electrónico, ya que el número puede variar enormemente, como veremos cuando realicemos la evaluación de las obras del sector turístico en el siguiente capítulo. En este sentido, los traductores necesitan que las obras terminológicas que consultan incluyan el mayor número de entradas posibles para su especialización, a fin de contar con una

⁴⁸⁷ Para la elección de estos parámetros, hemos partido de la evaluación formal que realiza Haensch (1997:239) para la elaboración de reseñas de fuentes lexicográficas.

⁴⁸⁸ Como veíamos en la clasificación de fuentes del apartado 2.4.1., la fuente original de una obra de referencia puede ser un corpus textual u otras obras de referencia.

mayor probabilidad de encontrar el término que buscan. Así, cuantas más entradas tenga una obra terminológica, mayor será su puntuación para este indicador.⁴⁸⁹

En este bloque, también podemos incluir la evaluación sobre la ausencia o *Presencia de errores* en la obra de referencia, tanto en las entradas como en las explicaciones y en las indicaciones nombradas anteriormente. Una fuente con errores, ya sean ortográficos, ortotipográficos, de sentido, etc. deberá considerarse como una fuente descuidada y, por tanto, poco fiable. De esta manera, la presencia de errores significará una puntuación baja en este apartado.

La *Ergonomía*⁴⁹⁰ de la fuente es esencial a la hora de utilizarla, ya que permite que el usuario pueda manejarla de forma fácil, cómoda e, incluso, de forma intuitiva. Asimismo, también hace referencia a la posibilidad de realizar diferentes tipos de búsquedas, con la combinación de varias ordenaciones de los lemas (alfabético y semántico, en las fuentes en formato papel) o de varios tipos de consulta (búsqueda en entrada, búsquedas en lista de entradas, búsqueda asistida, búsqueda múltiple, si se trata de fuentes en formato electrónico). Por este motivo, la consideramos como un indicador más de la calidad de la fuente en cuestión.

De esta forma, el bloque de macroestructura quedaría de la siguiente manera:

| MACROESTRUCTURA | |
|---------------------|--|
| Indicaciones | <input type="checkbox"/> Sí <input type="checkbox"/> No |
| Número de entradas | |
| Ausencia de errores | <input type="checkbox"/> Sí <input type="checkbox"/> No |
| Ergonomía | <input type="checkbox"/> Sí <input type="checkbox"/> No |

Tabla 8. Indicadores del segundo parámetro de la plantilla de evaluación.

⁴⁸⁹ En este apartado se tendrá en cuenta la especialización de la obra, ya que cuanto más especializada sea y más acotado tenga el campo de actuación, menor número de entradas podrá ofrecer. En este sentido, se otorgará la máxima puntuación en el criterio de *Especialización en*, siempre y cuando sea una obra del sector del turismo de aventura, y se puntuará este criterio según el número de entradas ofrecidas.

⁴⁹⁰ Codina (2000b) considera este indicador como elemento evaluador de los recursos digitales. En nuestro caso, no hemos incluido el resto de indicadores que considera este autor por tratarse de indicadores exclusivamente para los recursos electrónicos. Sin embargo, somos de la opinión de que este indicador es una muestra de calidad y fiabilidad a la hora de utilizar una fuente lexicográfica o terminológica.

A continuación, nos encontramos con el parámetro de la microestructura. Como hemos indicado anteriormente, la microestructura hace referencia a la información contenida en las diferentes entradas de una fuente de referencia. En este sentido, distinguiremos varios indicadores que nos permitirán evaluar la fuente en cuestión.

En primer lugar, tenemos la *Definición*. Como vimos en el apartado 2.4.1.1, existen diferentes tipos de fuentes según su presentación y, por ello, unas fuentes incluyen definición, como los diccionarios, y otras no, como los glosarios. Derivado de la información extraída con la consulta de investigaciones anteriores, así como de los datos de la encuesta realizada en este trabajo de investigación (cf. 2.5.2.), podemos concluir que un traductor requiere fuentes terminológicas con definiciones, ya que es la manera que tiene de poder conocer con exactitud el significado de la unidad que está buscando. En nuestra plantilla, valoraremos que la fuente tenga definición y, a la vez, que la definición sea adecuada, es decir, que proporcione una explicación concisa, precisa y directa, que no tenga remisiones circulares, que no esté incompleta u ofrezca información errónea o negativa (ISO 704, 2000: 17-22).⁴⁹¹

También consideramos importante la inclusión de *Información gramatical*, es decir, si la entrada se trata de un verbo transitivo, un artículo, un sustantivo masculino, etc. Esta información se incluye por lo general en las fuentes en formato papel, pero las fuentes electrónicas carecen normalmente de ella, lo que disminuye la calidad de una obra de consulta. En este apartado, también se puede evaluar la inclusión de información gramatical adicional, como sería la indicación de las formas conjugadas de verbos, participio, etc.

La información sobre el *Uso de las unidades* léxicas o terminológicas es imprescindible para cualquier usuario y, especialmente, para los traductores. Las fuentes

⁴⁹¹ Además de estos principios recomendados por la norma ISO, hemos tenido en cuenta también las recomendaciones para la elaboración de definiciones propuestos por el IULA (2010), que establece los siguientes puntos: a) evitar expresiones y formulaciones de tipo metalingüístico (por ejemplo, «Dícese de»); b) evitar palabras más complicadas que la palabra definida; c) no utilizar en la definición el término definido; d) partir de un descriptor de la misma categoría gramatical que el término definido; e) utilizar el mismo descriptor para todos los términos que pertenecen a una misma clase o subclase conceptual; f) contener una sola oración —sin puntos—; g) evitar paráfrasis conceptualmente vacías, y h) adaptar la formulación de la definición a las necesidades específicas de los usuarios. Al no ser posible comprobar todas las definiciones que incluyen las diferentes fuentes utilizadas, hemos tomado al azar cinco entradas y hemos valorado sus definiciones.

de consulta nos deben indicar a qué disciplinas o campos pertenece la unidad léxica consultada, a modo de adquirir el conocimiento especializado del que puede carecer un traductor profesional por no ser especialista en el dominio en cuestión. Este es un indicador básico, incluso para una fuente terminológica, ya que en muchas ocasiones se producen trasvases de unidades entre disciplinas o especialidades. Con respecto a este criterio, no solamente es necesario que los recursos indiquen la disciplina o el campo de especialidad al que pertenece un término, sino que sería muy adecuado la inclusión de información pragmática acerca del contexto en el que se utiliza la unidad, algún rasgo relevante en torno al uso cultural, institucional, etc. que pudiera tener, una posible variación ortográfica o geográfica, la frecuencia de uso, entre otras razones.

También es importante que las fuentes de consulta proporcionen *Información sintagmática*, es decir, locuciones, colocaciones y enunciados fraseológicos (Corpas Pastor, 1997), ya que son unos elementos básicos para que el discurso suene natural, incluso en los discursos especializados. Estos elementos son los más complicados de consultar y los que menos aparecen en las fuentes lexicográficas y terminológicas, por lo que dificulta a menudo la labor del traductor, a pesar de ser básicos para la producción de un texto de especialidad de calidad.

Además de la información sintagmática, las fuentes de consulta deben incluir *Información paradigmática*, a saber: sinónimos, palabras afines, antónimos, etc., para ofrecer una visión más completa del lema consultado, así como de la organización conceptual del dominio de especialidad.

Consideramos esencial la inclusión de *Ejemplos* en las entradas de los lemas, ya que, de esta forma, el usuario puede observar la unidad en su contexto y en pleno funcionamiento. Lamentablemente, este elemento no se suele incluir en las fuentes de información, y menos aún, en las fuentes terminológicas.

Otro indicador de calidad y fiabilidad de una fuente de información son las *Remisiones o hiperenlaces* (en el caso de las fuentes electrónicas) a otras entradas. Consideramos que es importante crear una relación conceptual entre las diferentes entradas, especialmente cuando la organización de la fuente es por orden alfabético, por dos razones: en primer lugar, porque facilita la consulta de la información contenida en

el recurso, así como el acceso a unidades relacionadas; y en segundo lugar, porque le permite al usuario adquirir un conocimiento mayor sobre la organización conceptual del dominio.

Por último, consideramos un indicador de calidad la inclusión de elementos distintos de sustantivos (unidades prácticamente únicas en los recursos terminográficos), es decir, consideramos importante la inclusión de verbos, adjetivos, adverbios, locuciones, abreviaturas y acrónimos, ya que estos elementos también forman parte del discurso de especialidad y deberían recogerse en las fuentes que pretenden representar dicho discurso. Asimismo, la inclusión de estas unidades facilita la consulta y obtención de información por parte del traductor.

Así, la tabla relativa al parámetro de microestructura quedaría como sigue:

| MICROESTRUCTURA | |
|---|--|
| Definición | <input type="checkbox"/> Sí <input type="checkbox"/> No |
| Información gramatical | <input type="checkbox"/> Sí <input type="checkbox"/> No |
| Uso de las unidades | <input type="checkbox"/> Sí <input type="checkbox"/> No |
| Información sintagmática (locuciones, colocaciones y enunciados fraseológicos) | <input type="checkbox"/> Sí <input type="checkbox"/> No |
| Información paradigmática (sinónimos, antónimos) | <input type="checkbox"/> Sí <input type="checkbox"/> No |
| Ejemplos | <input type="checkbox"/> Sí <input type="checkbox"/> No |
| Remisiones/hiperenlaces | <input type="checkbox"/> Sí <input type="checkbox"/> No |
| Tipos diferentes de entradas (sustantivo, adjetivo, verbo, adverbio, acrónimos, abreviaturas) | <input type="checkbox"/> Sí <input type="checkbox"/> No |

Tabla 9. Indicadores del tercer parámetro de la plantilla de evaluación.

Una vez expuestos los parámetros e indicadores de nuestra plantilla, explicaremos el sistema de puntuación que hemos utilizado para establecer una evaluación objetiva. Nuestra propuesta se basa en un sistema de puntos con una escala del 0 al 2, siendo 0 el valor mínimo y 2 el valor máximo. Actualmente, hay gran variedad de sistemas de evaluación de recursos, ya sean mediante la asignación de una puntuación (Seghiri Domínguez, 2006; Schroeder, 2006) o mediante *checklists* de una serie de parámetros (Beckwith, 2005; Videon, 2006). En nuestro caso, nos hemos

decantado por un sistema de puntos, puesto que es el sistema que nos permite cuantificar la calidad de un recurso terminológico de forma rápida, sencilla y objetiva.

Los sistemas de puntos desarrollados hasta el momento pretenden evaluar solo recursos electrónicos con sus plantillas de evaluación, por lo que sus plantillas no se podrían aplicar a recursos en papel; no obstante, nos han servido de modelos para crear nuestro propio sistema, ya que compartimos la misma aplicación y los mismo objetivos, aunque el contenido sea diferente.

Con este sistema de puntuación del 0 al 2, el evaluador de una determinada fuente de información terminológica realiza una reflexión crítica y objetiva de cada indicador para otorgarle una puntuación individual según el grado de fiabilidad. La escala de puntos tiene solo tres valores (0, 1, 2), puesto que después de comprobar las diferentes propuestas de valores para la evaluación de recursos electrónicos, consideramos que es la escala que presenta una mayor objetividad y sencillez a la hora de realizar una evaluación.

Algunos autores establecen esta puntuación individual por parámetros (Seghiri Domínguez, 2006), pero en nuestra plantilla todos los indicadores tienen un valor en sí mismo y, por tanto, representan parte de la fiabilidad de la obra. En este sentido, hemos incluido una valoración del 0 al 2 en cada indicador evaluable de la plantilla.⁴⁹²

Esta plantilla contaría, por tanto, con 15 indicadores determinantes de la fiabilidad de una fuente de información evaluada y su puntuación máxima sería de 30 puntos en total. Además, se ha añadido un último valor evaluable, para que el evaluador pueda evaluar cualquier aspecto positivo que considere que no se ha evaluado en ninguno de los indicadores anteriores. Por ejemplo, si evaluamos una fuente terminológica de iconos, como el *Glosario de Iconos*, de Iberia Rural SL., habrá indicadores que no se ajusten a la evaluación; en este caso, se podría puntuar con dos puntos más para equilibrar la evaluación con otros recursos terminográficos.

Consideramos que la puntuación mínima que debe alcanzar una fuente de información para ser considerada como fiable, después de haber sido valorada con todos

⁴⁹² La plantilla de evaluación completa se puede consultar en el Anexo II.

los indicadores, debe ser de 8 puntos (escala de 0 a 8 puntos), ya que por debajo de esta cantidad no contaría con la calidad mínima recomendada. A partir de esa puntuación, nos encontramos con otros tres grupos: el grupo de 9 a 16 puntos, que se trataría de un recurso con una calidad medio-baja; el siguiente grupo con una puntuación comprendida entre 17 a 23 puntos, que presentaría una calidad medio-alta, y el último grupo de 24 a 30 puntos, que contaría con una fiabilidad alta.⁴⁹³ De esta forma, cuanto mayor sea la puntuación obtenida, mayor será su fiabilidad.⁴⁹⁴

Para terminar, nos gustaría enfatizar una vez más la necesidad que existe de evaluar las fuentes terminológicas disponibles en la actualidad, tanto en papel como en formato electrónico, ya que hay un gran volumen de ellas y su calidad es a menudo cuestionable. Cualquier usuario de estas fuentes, en especial los traductores profesionales, deberían conocer y utilizar herramientas de evaluación para poder discernir entre la información fiable de aquella que no lo es. Como muestra de este funcionamiento, aplicaremos la plantilla de evaluación desarrollada en este apartado a las fuentes terminológicas del ámbito del turismo en el siguiente capítulo, con objeto de estudiar sus deficiencias y puntos fuertes y débiles, así como de mostrar la aplicación real de la plantilla de evaluación desarrollada en este apartado.

2.5. EXPECTATIVAS Y NECESIDADES DE CONSULTA DE TRADUCTORES PROFESIONALES

Como hemos indicado anteriormente, en el ámbito terminográfico se ha prestado escasa atención a los traductores profesionales como grupo potencial de destinatarios en

⁴⁹³ La esquema de puntos se resume de la siguiente manera:

0- 8 puntos: Fiabilidad baja

9-16 puntos: Fiabilidad media-baja

17-23 puntos: Fiabilidad media-alta

24-30 puntos: Fiabilidad alta

⁴⁹⁴ El umbral de puntuación de esta evaluación se ha calculado a partir de los resultados obtenidos en una evaluación de prueba de 15 recursos terminográficos y teniendo en cuenta, además, el procedimiento seguido en un estudio piloto realizado por Seghiri Domínguez (2006: 94) en su tesis doctoral.

la elaboración de los recursos terminográficos y, por ello, la mayoría de estos recursos presentan serias carencias con relación a las expectativas y necesidades de estos profesionales durante su consulta, tanto en lo que respecta a la calidad del contenido como a la cantidad de información que ofrecen. Además de esta falta de atención por parte de los terminógrafos, la baja calidad de estos recursos se debe también a veces a una falta de trabajo sistemático en su elaboración. En este contexto, los traductores se ven a menudo obligados a crear sus propios recursos terminográficos, ya sea a partir de corpus *ad hoc*, de memorias de traducción o directamente sobre el texto de origen de la traducción, a pesar de los plazos restrictivos con los que trabajan normalmente.

Bergenholtz y Tarp (1995: 77) tratan el problema de la elaboración de diccionarios de forma general y defienden la necesidad de utilizar una comunicación previa con el usuario potencial de la fuente lexicográfica en cuestión para incorporar a la obra la información deseada por este usuario.

Lexicographical work often proceeds without any prior knowledge of the potential user group, and the dictionary may therefore be said to be the result of the lexicographer's own conjectures concerning user needs for lemmata, collocations, sentence examples, encyclopaedic and linguistic information, etc. To acquire more precise knowledge, the lexicographer may make a user survey before starting actual work on the dictionary, with the aim of uncovering the needs of potential users in relation to the information categories to be incorporated in the dictionary as well as the representation of this information (Bergenholtz y Tarp, 1995: 77).

En la misma línea, se encuentra Lew (2011: 3):

Expert opinion, intuition, or purely deductive reasoning cannot replace solid empirical evidence from user studies: dictionary use is just too complex an affair to be that predictable.

En nuestro caso, estamos completamente de acuerdo con los autores mencionados y, por tanto, con los trabajos terminográficos que emplean un estudio acerca de las necesidades de los usuarios potenciales en la fase previa a la elaboración de los recursos, al igual que la mayoría de los terminógrafos profesionales en la actualidad.

En la misma obra, Bergenholtz y Tarp (ibíd.) realizan una distinción de los diferentes tipos de estudios previos que se pueden llevar a cabo para la elaboración de

un recurso de este tipo: en primer lugar, consideran los cuestionarios, donde se emplea la misma metodología que en los estudios de mercado, es decir, se realizan un número de preguntas estándar a los usuarios potenciales en relación con un producto; en segundo lugar, presentan las entrevistas, que se utilizan normalmente con una muestra muy concreta como medio para detectar los problemas y mejorar las preguntas de los cuestionarios; en tercer lugar, los protocolos, que consisten en describir el proceso de trabajo de una tarea concreta que necesita el uso de recursos lexicográficos; y por último, las pruebas de casos, que consiste en asignar una tarea que debe resolverse mediante el uso de los recursos lexicográficos.

Cabré Castellví (2004c: 125), por su parte, proclama la necesidad de que los traductores participen y desarrollen más la terminología sistemática (terminología sectorial, según la autora), a fin de alcanzar unos mejores resultados en los recursos terminográficos dirigidos a los traductores.

[...] terminología sectorial, campo en el que los traductores deberían desarrollar más actividad de la que ejercen. Los glosarios elaborados por terminógrafos no suelen satisfacer las necesidades de los traductores, porque no parten del análisis de las necesidades reales que plantea la traducción. Por ello, clamamos para que el colectivo de la profesión traductora entre con mayor profusión en la elaboración de recursos terminográficos que respondan más coherente y adecuadamente a las necesidades reales de la traducción.

En nuestro trabajo de investigación, llevaremos a cabo las dos actividades propuestas por estos expertos, a saber: por un lado, realizaremos un estudio previo sobre las necesidades y expectativas de los usuarios potenciales, en nuestro caso los traductores profesionales, a fin de elaborar el recurso terminológico según sus criterios; y por otro lado, como profesionales de la traducción, realizaremos un trabajo sistemático sobre el segmento del turismo de aventura basado en una metodología propuesta en esta investigación, con objeto de incorporar las necesidades expuestas en la encuesta y los criterios según nuestra experiencia.

A continuación, presentaremos el proceso de elaboración y desarrollo de nuestra encuesta, así como el análisis de los resultados obtenidos, precedidos de una

recopilación de estudios previos similares realizados en el campo de la traducción y de la enseñanza de lenguas.

2.5.1. ESTUDIOS PREVIOS REALIZADOS

Los estudios y las encuestas realizadas hasta ahora en relación con la terminología y los recursos lexicográficos se han centrado principalmente en el estudio del uso que los estudiantes de traducción o de segunda lenguas dan a los diccionarios (Welker, 2010, 2008; Tarp, 2009; East, 2008; Měchura, 2008; Pujol et ál., 2006; Sánchez Ramos, 2005; Bogaards, 2005, Corpas Pastor et ál., 2001; Mackintosh, 1997; Dancette, 1995; Varantola, 1994; Duvå y Laursen, 1994; Bejoint, 1981; Benbow, 1980). En nuestro caso, el grupo objeto de estudio son los profesionales de la traducción, que consideramos como los verdaderos protagonistas de la traducción⁴⁹⁵ y los que, con su experiencia, han aprendido a conocer sus necesidades con respecto a los recursos terminográficos. Para este grupo receptor específico, no hemos encontrado hasta ahora estudios donde se les requiera información previa sobre sus necesidades y sus expectativas a la hora de consultar una obra de referencia, a excepción de un estudio realizado por Agirre et ál. (2000) y por el proyecto POINTER (1996), que comentaremos más adelante.

Los estudios previos que han realizado análisis similares al que pretendemos llevar a cabo en este trabajo se pueden dividir en dos grupos principalmente: por un lado, los estudios que analizan las destrezas y habilidades del uso del diccionario; y por otro, aquellos que pretenden estudiar las necesidades y expectativas de los traductores (Durán Muñoz, 2010a).

El primer grupo hace referencia a estudios que proclaman la necesidad de enseñar a los estudiantes, ya sean de segunda lengua o de traducción, a utilizar los recursos lexicográficos en general. Para ellos, el diccionario es un «libro especial» y, a fin de que se utilice de forma efectiva y eficiente, los usuarios necesitan adquirir una

⁴⁹⁵ En este caso, hablamos de «traducción» en general y, con ello, pretendemos englobar todas las actividades que consisten en la traslación del mensaje de una lengua a otra, ya sea de forma oral o escrita.

serie de habilidades. De hecho, estos estudios defienden la necesidad imperante de este tipo de enseñanza, ya que, de lo contrario, no se sacaría partido de las ventajas que ofrecen este tipo de recursos. Los trabajos relacionados con este objetivo han tenido principalmente como grupo de destinatarios a los estudiantes de segundas lenguas (Bejoint 1981, Hartmann, 1999; Bogaards, 2005; East, 2008), aunque también podemos encontrar algunos estudios acerca de estudiantes de traducción (Roberts, 1992; Atkins y Varantola, 1998; Mackintosh, 1998; Varantola, 1998; Sánchez Ramos, 2005). En este contexto, este primer grupo de estudios dirigidos a analizar las destrezas y habilidades del uso del diccionario se puede dividir a su vez en dos subgrupos: por un lado, estudios centrados en los estudiantes de segundas lenguas; y por otro, en estudiantes de traducción.

Con respecto al grupo de estudios basados en estudiantes de traducción, varios autores (Roberts, 1992, entre otros) defienden que los traductores necesitan aprender a consultar los diccionarios y demás recursos de forma eficiente, con objeto de alcanzar una buena calidad en los resultados de sus traducciones. En este sentido, es esencial para ellos estudiar en mayor profundidad la relación entre los estudiantes de traducción y los diccionarios (Durán Muñoz, 2010a: 60) y, por tanto, llevar a cabo estudios empíricos sobre los hábitos de uso, las habilidades y los diferentes problemas a los que se enfrentan estos estudiantes cuando utilizan los diccionarios y otros recursos lexicográficos.

Dentro de este contexto de análisis de habilidades y protocolo de trabajo con recursos lexicográficos, hemos encontrado un único estudio basado en traductores profesionales realizado como parte de la tesis doctoral de Nord (2002). Se trata de un estudio empírico que consistió en una técnica de observación de 30 traductores profesionales con diferentes lenguas maternas durante sus horas de trabajo cuyo objetivo consistía en conocer el comportamiento real de dichos sujetos en situaciones reales. Durante la observación se le fueron realizando preguntas acerca de las herramientas utilizadas, de las razones para la utilización de una u otra y de los déficits lingüísticos y cognitivos que provocaban la consulta. Las principales conclusiones extraídas de este estudio empírico fueron las siguientes:

- El tipo de fuente consultada: más de dos tercios de las búsquedas se realizaron en diccionarios, principalmente bilingües especializados, bilingües generales, monolingües generales, en este orden. También se realizaron consultas en otros tipos de diccionarios (de sinónimos, por ejemplo) y en textos paralelos.
- Frecuencia de consulta: durante la traducción de textos generales, la media de consulta era de tres minutos, mientras que en la traducción de textos especializados era de dos minutos.
- Principales búsquedas: las búsquedas más comunes eran para detectar equivalentes, comprensión de significados, búsqueda de uso y selección del equivalente más adecuado, en ese orden.
- Principales unidades buscadas: sustantivos, grupos de palabras y abreviaturas, en ese orden.

Como decíamos, este estudio es el único realizado sobre traductores profesionales al que hemos tenido acceso, lo que demuestra el déficit de este tipo de estudios acerca de este grupo.

Por su parte, el segundo grupo, dedicado al estudio de las necesidades y expectativas de los traductores, se aproxima más a nuestra investigación, aunque el objetivo común de la mayoría de los estudios recae en conocer las necesidades de los estudiantes de traducción. Entre otros, podemos citar los trabajos de Duvå y Laursen (1994), Dancette y Réthoré (1997) y Corpas Pastor et ál. (2001).⁴⁹⁶ En general, estos estudios pretenden conocer, a través de diferentes técnicas, las necesidades o las dificultades que encuentran sus informantes durante el proceso de traducción, es decir, cuál es la información que consultan (gramática, definición, etc.), cuáles son las fuentes de consulta, cómo realizan la consulta, cuáles son las dificultades que encuentran durante las consultas, entre otras cuestiones de interés.

⁴⁹⁶ Debemos mencionar que las autoras Duvå and Laursen (1994) contaron con un grupo de informantes que estaba formado parcialmente por graduados (sin especificar profesión) y por traductores profesionales, que representaba un 38% del total de la muestra.

No obstante, como decíamos al principio de este apartado, hemos podido acceder a un estudio (Agirre et ál., 2000: 302-303) que, aunque no especifica la composición de su muestra, habla en ocasiones de traductores profesionales a lo largo de la presentación de su estudio. Por ser este un estudio que se aproxima al nuestro, vamos a presentar a continuación una breve descripción del proyecto y de los resultados obtenidos.

En el estudio realizado por Agirre et ál. (2000), se utilizaron dos tipos de técnicas diferentes: una, el protocolo de observación directa, con la que se entregaron a los traductores (en este caso de francés y de vasco) varios textos que tenían que ser traducidos junto con diccionarios (monolingües y bilingües con diferentes características) para que registraran sus problemas, las soluciones que adoptaban y las tareas desempeñadas. Según los mismos autores, el objetivo de este estudio consistía en caracterizar la actividad de los traductores mediante la observación de las tareas de traducción de palabras, expresiones, frases contextualizadas e incluso párrafos. Cada vez que los traductores consultaban un diccionario, estos tenían que registrar la unidad objeto de traducción, el diccionario utilizado, la entrada del diccionario consultada y el tipo de consulta. La otra técnica utilizada se basaba en entrevistas personales con traductores profesionales, que permitieron detectar los diferentes usos del diccionario según sus experiencias. A los entrevistados se les preguntó por las características que debían tener los diccionarios para que los utilizaran, por su interés por los diccionarios electrónicos y por las funciones que esperaban, entre otros temas. Los datos que se extrajeron de este estudio dieron lugar a unas conclusiones determinantes y algunas, ya supuestas:

- Los traductores expertos y los ocasionales necesitan diferente información.
- Algunos traductores, especialmente los ocasionales, encontraron los diccionarios bilingües muy útiles.
- Los términos formados por más de una unidad causaban problemas a la hora de buscarlos en los diccionarios y de encontrar un equivalente.
- El contexto es importante cuando se traduce un texto.

- Los diccionarios dirigidos a la traducción deben proporcionar información gramatical y pragmática.
- La proximidad de las lenguas es un aspecto útil, pero es importante que los diccionarios incluyan información sobre los falsos cognados.

Por su parte, el proyecto POINTER (1996),⁴⁹⁷ como decíamos, también incluye a traductores profesionales entre sus informantes, aunque estos no son los únicos que forman la muestra como veremos a continuación.

Este proyecto se llevó a cabo con la finalidad de estudiar detalladamente los recursos terminográficos existentes en Europa en su momento (1995-1996) y recomendar en consecuencia una serie de propuestas que asegurasen el acceso, distribución y reutilización de dichos recursos, así como la creación de recursos nuevos que se ajustasen a las demandas del creciente mercado europeo. Para ello, uno de sus objetivos consistió en la especificación de las necesidades de los usuarios con relación a los recursos terminográficos disponibles. Para llevar a cabo este objetivo, realizaron un estudio basado en entrevistas personales a un grupo representativo de usuarios relacionados con el trabajo terminológico. Estos usuarios se clasificaban en tres grupos diferentes: un grupo estaba formado por expertos (organismos de estandarización, redactores técnicos, docentes); otro, por ingenieros informáticos (inteligencia artificial, lingüistas computacionales), y por último, otro grupo formado por profesionales de la lengua (traductores, documentalistas, periodistas, editores).⁴⁹⁸ Estos entrevistados debían responder a dos cuestiones principalmente: por un lado, el tipo de información que debería contener un recurso terminológico; y por otro, la estructura que debería presentar un recurso de este tipo para garantizar una consulta accesible y útil. En primer lugar, se les realizaba una serie de preguntas sobre el tipo de información que debería contener un recurso terminológico para ser considerado de calidad y se les pedía que la

⁴⁹⁷ <<http://www.computing.surrey.ac.uk/ai/pointer/>>

⁴⁹⁸ Como se puede observar, el grupo de traductores no es el principal objeto de estudio, pero tanto el estudio como las conclusiones son relevantes para nuestro estudio y, por ello, lo hemos querido incluir en este apartado. Asimismo, este estudio nos ha servido como base en algunas ocasiones para elaborar preguntas de nuestro estudio. No obstante, las conclusiones que se presentan están relacionadas con el grupo de profesionales de la lengua, por lo que hay que ser prudente a la hora de generalizar a los traductores profesionales.

clasificaran según tres tipos: información esencial, deseable e irrelevante. A continuación, se les pedía que indicaran sus preferencias de consulta del formato y la organización de contenido de los recursos terminográficos, es decir, se les preguntaba si preferían recursos electrónicos o en papel y ordenados de forma alfabética o conceptual

Como conclusiones específicas del grupo de usuarios que aquí nos ocupa, los traductores profesionales, indicamos a continuación las siguientes conclusiones extraídas del proyecto POINTER (1996):

- Los equivalentes en los recursos se consideran como información esencial y, por tanto, es imprescindible su inclusión junto con el término de origen.
- Se considera importante también la información gramatical sobre las denominaciones que describen los conceptos.
- Las definiciones se consideran como información esencial, a fin de comprender el concepto, de conocer su finalidad y sus campos de aplicación, valorar la adecuación de los equivalentes de traducción, facilitar la indexación y recuperación de los documentos, etc.
- La mayoría de los profesionales de la lengua consideran la información semántica (relaciones conceptuales entre conceptos) como información esencial.
- La indicación del campo de especialidad se considera como información esencial.
- El contexto, las notas, los ejemplos y las definiciones de equivalentes se consideran como información deseable.
- El formato electrónico es preferible frente al formato en papel.

Además de estas conclusiones derivadas del estudio llevado a cabo por Agirre et ál. (2000) y por el proyecto POINTER, debemos hacer mención también a los resultados obtenidos por los demás estudios basados en las expectativas y necesidades de los traductores, aunque sean estos estudiantes y no profesionales. Según estos estudios, los traductores (sin hacer distinciones) necesitan los siguientes tipos de información: 1.

información lingüística (por ej., definiciones), 2. información semántica (por ej. relaciones semánticas entre conceptos), y 3. información pragmática (por ej. contexto).

En nuestro caso, nuestra investigación pertenece también a este segundo grupo, aunque hay que tener en cuenta que nuestro grupo de destinatarios está formado totalmente por traductores profesionales. De este modo, pretendemos conocer qué recursos terminográficos consultan los traductores profesionales, qué tipo de información esperan encontrar, cómo consultan estos recursos y cuáles son los problemas o dificultades que encuentran a la hora de consultarlos. Por último, realizaremos una comparación entre los resultados obtenidos por estos estudios previos y por los alcanzados mediante nuestra encuesta.

A continuación, realizamos la presentación del estudio realizado en forma de encuesta, así como la fase de elaboración y consecución de esta, a fin de llevar a cabo un estudio previo a la elaboración de nuestra base de datos ontoterminográfica con el objetivo de conocer las necesidades y expectativas del grupo de destinatarios de nuestra obra (los traductores profesionales). Por último, realizaremos un análisis de los resultados obtenidos y una comparación entre nuestras conclusiones y las extraídas en los estudios anteriores.

2.5.2. PRESENTACIÓN DE LA ENCUESTA Y SU ELABORACIÓN

Como decíamos con anterioridad, se pueden utilizar diferentes métodos para llevar a cabo un estudio previo de base empírica sobre las necesidades de los usuarios potenciales de un recurso terminológico. De esta manera, entre los estudios previos que hemos comentado con anterioridad, se han empleado los tres tipos de técnicas: técnica de protocolo (Duvå y Laursen, 1994; Atkins y Varantola, 1998; Agirre et ál., 2000), entrevista personal (Duvå y Laursen, 1994; proyecto POINTER, 1996; Agirre et ál., 2000) o cuestionario/encuesta (Corpas Pastor et ál. 2001).⁴⁹⁹ Cada uno de estos métodos

⁴⁹⁹ Obsérvese que es común que algunas técnicas se utilicen de forma combinada para intentar obtener las ventajas de las técnicas y paliar sus inconvenientes.

presenta sus ventajas y sus inconvenientes, y se deberán seleccionar según los objetivos y características del estudio que se vaya a llevar a cabo.⁵⁰⁰

En nuestro caso, hemos optado por el empleo de la encuesta, puesto que se trata de un método muy eficiente gracias a las ventajas que presenta y, además, se ha adaptado perfectamente a las necesidades de nuestro estudio. Asimismo, hemos podido subsanar en gran medida sus desventajas. La principal ventaja que presenta consiste en la posibilidad de conseguir un número muy elevado de respuestas homogéneas y comparables en un tiempo reducido, lo que no es posible con los otros métodos empíricos. Además, permite su administración desde un punto remoto mediante el uso de correo ordinario, correo electrónico o teléfono, lo que permite un mayor alcance de los encuestados y la eliminación de la barrera geográfica; permite un análisis cuantitativo automático, rápido y preciso, así como el procesamiento de un número de respuestas muy elevado. Otra ventaja de este método consiste en el reducido tiempo que se requiere para contestarlo, en comparación con el resto de los posibles métodos, lo que resulta muy relevante si tenemos en cuenta el tiempo limitado del que disponen los traductores profesionales.

No obstante, también hay que tener presente las desventajas de este método, que radican básicamente en la poca fiabilidad que pueden tener las respuestas obtenidas y en la falta de interés de los receptores. En nuestro caso, hemos intentado solventar estos problemas mediante la elaboración de una encuesta específica dirigida a un grupo determinado de receptores (en este caso, los profesionales de la traducción) y, de esta manera, motivar su participación. Asimismo, hemos distribuido la encuesta a través del correo electrónico, acompañada de una carta de presentación donde se explicaba la naturaleza del estudio, su marco de actuación y sus principales objetivos, y la hemos dirigido a empresas de traducción, instituciones y asociaciones de traductores e intérpretes profesionales de todo el mundo para que las hicieran llegar a sus miembros, a la vez que la hemos publicado en listas de distribución, foros de discusión, etc. de

⁵⁰⁰ En Welker (2010: 11-21) se realiza una descripción de cada técnica y se presentan las ventajas e inconvenientes de cada una.

traductores profesionales, con la intención de alcanzar el mayor número posible de encuestados de calidad.⁵⁰¹

Una vez que se ha seleccionado la metodología empírica que se va a emplear para el estudio, así como el universo al que irá dirigida (en nuestro caso, los traductores profesionales), es necesario pasar a la preparación del estudio. En primer lugar, es imprescindible determinar las hipótesis del estudio y los objetivos que se desean alcanzar con el trabajo.

En este estudio, la hipótesis principal ha consistido en la siguiente: «Los traductores profesionales no están satisfechos con los recursos terminográficos disponibles actualmente». A su vez, existen otras tres hipótesis secundarias: a) «Los recursos actuales no incluyen la información necesaria para los traductores profesionales»; b) «Para la labor traductora, los traductores necesitan información lingüística, pragmática y semántica», y c) «Los traductores profesionales se ven obligados a utilizar muchos recursos terminográficos diferentes durante sus trabajos».

A su vez, los objetivos principales planteados para este estudio se indican a continuación:

1. Conocer los recursos más habituales y preferidos por los traductores profesionales.
2. Conocer la forma de consulta más habitual por parte de los traductores profesionales.
3. Conocer la carga de terminología puntual que realizan y los métodos frente a la terminología sistemática, y sus preferencias.
4. Conocer los métodos de evaluación de los traductores profesionales.
5. Conocer las expectativas y necesidades específicas de los traductores profesionales durante la consulta de un recurso terminológico.

⁵⁰¹ Hemos utilizado las listas de distribución de RedIris, Traducción y Tradumática, la lista de The Linguist, Corpora y el foro de ProZ.com, entre otras, así como algunas asociaciones de traductores profesionales.

6. Conocer la información lingüística, pragmática y semántica que consideran como esencial, deseable e irrelevante.

Después de haber tomado las decisiones iniciales sobre el estudio, es necesario pasar a la fase de preparación. Para comenzar con esta fase, es importante realizar primero un estudio piloto o de prueba, en el que se pueden emplear diferentes métodos (Mucchielli, 1993:13): un método *documentalista*, que consiste en la recogida de material documental (libros, artículos, películas, conferencias, discos, etc.) que se haya publicado relacionado con el tema de la encuesta, igualmente como todas las encuestas anteriores de interés parecido y estudios previos; un método *introspectivo hermenéutico*, de «silencio y despacho cerrado», como dice el autor, evaluando todas las posibles hipótesis que pasen por la cabeza del estudioso; un método de *grupo de discusión*, que consiste en reunir personas no relacionadas con el tema propuesto que pueden tener ideas innovadoras; un método de *panel de expertos*, pidiendo opinión a los especialistas del tema; y un método de *entrevistas exploratorias* con los expertos o las personas que no han tenido hasta ahora ninguna experiencia en el tema.

En nuestro caso, además de haber llevado a cabo un método documentalista, como se ha expuesto en el apartado anterior con la recopilación de los estudios previos realizados, hemos puesto en práctica un método de panel de expertos, con un estudio piloto que se llevó a cabo antes de elaborar la versión definitiva de la encuesta. Este estudio se realizó con una serie de expertos en los campos relacionados con la encuesta, es decir, expertos en traducción y en terminología, con objeto de mejorar la versión inicial de las preguntas. Para ello, se contactó personalmente a través de correo electrónico con la muestra de las personas expertas seleccionadas y se les pidió que rellenaran la encuesta y que realizaran comentarios, recomendaciones, propuestas, etc. para conseguir una encuesta que pudiera cubrir las necesidades y los objetivos propuestos para el estudio. Una vez que se recogieron todos los comentarios de los expertos que participaron, se realizaron los cambios pertinentes y se elaboró la versión definitiva. En este sentido, se consiguió elaborar una encuesta de calidad en la que se recogían las preguntas más pertinentes para obtener los objetivos planteados.

La estructura de esta versión definitiva estaba formada por 20 preguntas en total, todas cerradas, a excepción de una pregunta abierta, con respuestas de una única opción u opción múltiple. Las preguntas se clasificaron en cuatro bloques diferentes, a saber:

- Datos generales acerca de los profesionales encuestados (experiencia, formación, vinculación con alguna asociación profesional).
- Datos generales acerca del entorno de trabajo de los profesionales encuestados (campos de especialidad, lenguas de trabajo, tipo de contrato).
- Datos sobre los recursos terminográficos (tipos de recursos que utilizan, soporte que prefieren, uso de la terminología puntual).
- Evaluación de los recursos terminográficos y sus ideales (tipo de evolución que realizan, tipo de información que requieren, problemas que encuentran).

Una vez que se obtuvo la versión definitiva de la encuesta,⁵⁰² tras recibir los comentarios de los expertos, aplicar las recomendaciones y correcciones de estos y comprobar la formulación adecuada de las preguntas,⁵⁰³ así como después de elaborar las diferentes versiones (español, inglés, alemán e italiano), se distribuyó la versión adecuada (según país y lengua materna) entre asociaciones de traductores, listas de distribución, foros de discusión, etc. acompañada por la carta de presentación indicada con anterioridad donde se especificaban los objetivos del estudio. Asimismo, se elaboró una versión electrónica de la encuesta en inglés, con el software libre de UCCASS (Unit Command Climate Assessment and Survey System),⁵⁰⁴ que se envió junto con las versiones anteriores, a fin de facilitar a los informantes la realización de la encuesta y el análisis posterior de esta. Esta encuesta electrónica se activó en julio de 2008 y se mantuvo activa durante tres meses, periodo durante el cual se recibieron respuestas de numerosos profesionales contactados a través de los medios indicados con anterioridad.

⁵⁰² La encuesta se encontraba disponible en la siguiente URL: <<http://clg.wlv.ac.uk/surveys/survey.php?sid=29>>, en versión inglesa. Para consultar la versión en español, véase Anexo I.

⁵⁰³ Se pretendió acabar con cualquier posible ambigüedad, con objeto de evitar dobles sentidos o malentendidos. De esta forma, se consiguieron preguntas claras, simples, cortas y directas, que no dieran lugar a subjetividad y que fueran fáciles de contestar.

⁵⁰⁴ <<http://www.bigredspark.com/survey.html>>

Tras el cierre de esta encuesta electrónica, se procedió a analizar las respuestas recibidas a través de este medio, así como las respuestas de las personas que la devolvieron a través del correo electrónico. En total, se recibieron 402 respuestas, de las cuales se pueden extraer conclusiones clarificadoras para la elaboración de recursos terminográficos de cualquier campo de especialidad.

Después de haber presentado el proceso de diseño y elaboración de la encuesta, seguiremos con la presentación de los encuestados y de los resultados obtenidos.

2.5.3. RESULTADOS DE LA ENCUESTA Y CONCLUSIONES

Como se ha indicado anteriormente, se ha pretendido contactar con el mayor número de profesionales de la traducción para, así, poder conocer su opinión sobre los recursos terminográficos disponibles en la actualidad y, a su vez, recopilar información sobre el uso que realizan de los recursos terminográficos y las necesidades que presentan en su trabajo con respecto a este tipo de fuentes de consulta. Asimismo, también se ha querido obtener información sobre las diferentes fuentes terminológicas que emplean y sus preferencias con relación al contenido y a la forma.

Los dos primeros bloques (las siete primeras preguntas) de la encuesta están dirigidas a caracterizar a los encuestados, a razón de recibir información sobre su profesión, su experiencia, sus lenguas de trabajo, sus campos de especialidad, etc. A partir de los datos obtenidos, podemos establecer lo siguiente acerca de los informantes:

La mayoría de los encuestados, el 62,55%, son traductores mientras que 13,67% son intérpretes, el 6,67% son gestores de proyectos, el 5,24% son terminólogos, el 3,00% son subtituladores y el 8,80% se dedican a otra profesión que no han podido especificar.

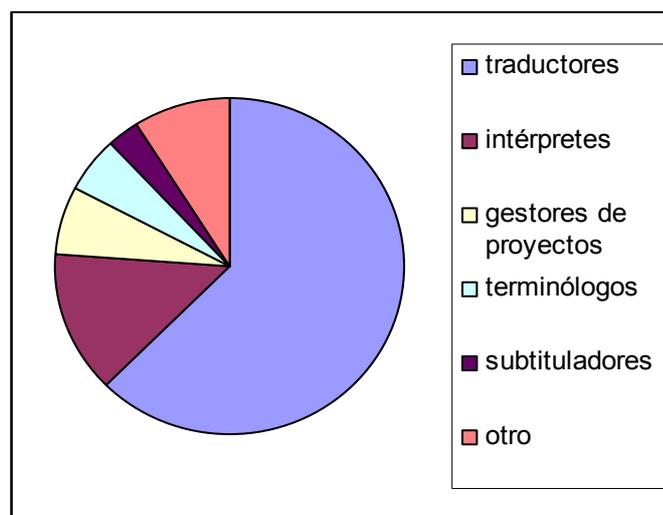


Gráfico 1. Datos profesionales de los encuestados.

Asimismo, la mayor parte de los encuestados son trabajadores autónomos, con un 68,48%, mientras que el 13,93% y el 13,18% son empleados de empresas e instituciones, respectivamente. Casi la mitad de los encuestados (42,49%) llevan más de 10 años en la profesión y el 27,19% lleva cuatro o menos de cuatro años como traductor profesional.

Por su parte, el 61,96% de los encuestados cuenta con una licenciatura en Traducción; el 22,22% pertenece a alguna asociación u organización profesional; el 10,73% cuenta con otro tipo de certificación no especificada y solo el 5,08% de los encuestados no tiene ninguna formación relacionada con la traducción.

Un aspecto importante en la descripción de los encuestados es el referente a los campos de especialidad en el que trabajan.⁵⁰⁵ De los datos obtenidos, se pueden obtener dos conclusiones importantes relacionadas con el campo de la traducción: por un lado, se observa el número medio de campos de especialidad en los que trabajan los encuestados y, por otro, se obtienen los campos de especialidad con un mayor volumen de trabajo. En el primer caso, la media de dominios diferentes a los que se dedican los traductores es de seis, aunque encontramos tanto a profesionales que se dedican a un

⁵⁰⁵ Los campos de especialidad incluidos en la encuesta se han seleccionado de la página web del Institute of Translation and Interpreting (ITI). URL: <http://www.iti.org.uk/indexMain.html>.

único campo de especialidad como a más de diez campos diferentes, llegando incluso a encontrar a traductores que trabajan en quince dominios diversos. Asimismo, algunos profesionales trabajan en campos muy cercanos entre sí, como serían el de negocios, matemáticas/estadística, economía y marketing, y en otros casos en campos alejados, como derecho, artesanía, geografía e informática. En el segundo caso, se puede deducir el peso en la actualidad de ciertos ámbitos especializados para la traducción. En el siguiente gráfico se observan los campos de especialidad en el mercado de traducción que han obtenido un porcentaje alto, más del 20%.

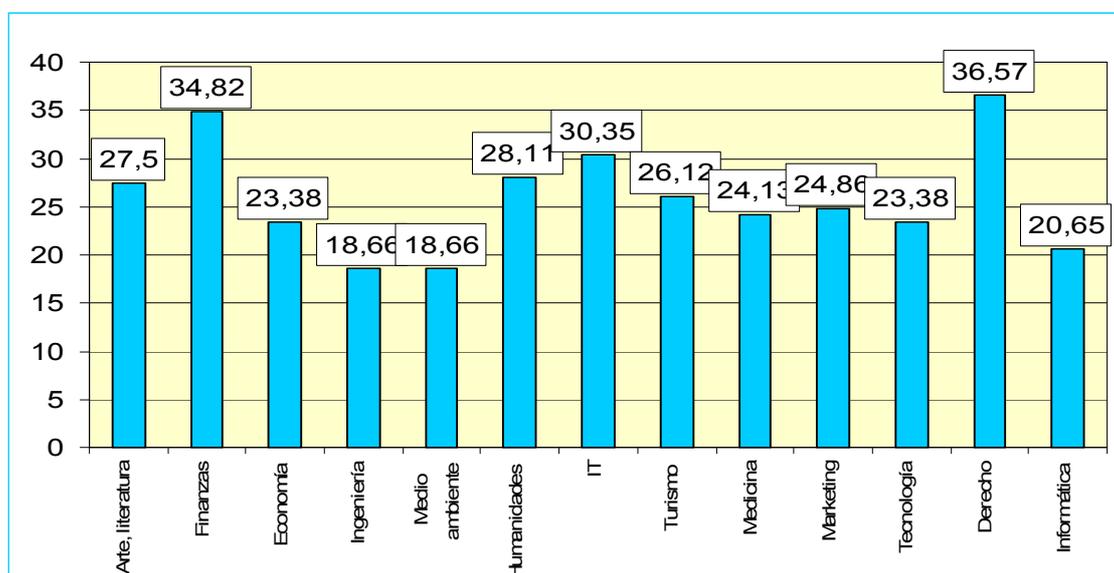


Gráfico 2. Volumen de mercado de traducción en diferentes campos de especialidad.⁵⁰⁶

Como se aprecia en el gráfico, el volumen de traducciones del ámbito jurídico ocupa el primer lugar con respecto a los demás tipos, con un 36,57%, seguido muy de cerca por el campo de negocios (34,82%). A continuación, encontramos las traducciones del campo de la tecnología de la información, con un 30,35%, y las de humanidades, con un 28,11%. El quinto lugar lo ocupan el amplio campo de

⁵⁰⁶ Este gráfico se ha obtenido de los resultados de la encuesta, por lo que no se pueden considerar concluyentes si tenemos en cuenta el número de encuestados, pero sí se pueden tratar como una aproximación al mercado de la traducción.

especialidad de arte, literatura o medios de comunicación, con un 27,6%, y el sexto puesto lo ocupa la traducción turística, con un 26,12%. Después, nos encontramos las traducciones de marketing (24,86%), medicina y farmacia (24,13%), etc.

Como dato significativo e interesante para nuestra tesis doctoral, la traducción turística ocupa el sexto lugar en volumen de trabajo, lo que significa que representa una porción de mercado muy elevada según estos datos. Si tenemos en cuenta el *Estudio de la situación del mercado español de servicios profesionales de traducción*, publicado por la ACT en 2005, donde se indicaba que el mercado de la traducción turística ocupaba el quinto puesto con relación al volumen de traducciones, podemos corroborar que esta traducción de textos especializados se sitúa en un lugar preferente, tanto dentro de España como en el mercado internacional de la traducción, ya que solo varía un puesto al ampliar el número de países analizados.

Asimismo, también es interesante constatar que el idioma inglés sigue siendo el principal idioma de traducción en todo tipo de traducción de textos especializados. En lo referente a la traducción turística, por tanto, también ocurre lo mismo, donde aparece como el idioma de trabajo de la gran mayoría de los encuestados (95%). El segundo idioma de trabajo en este tipo de traducción, según los resultados de esta encuesta, es el español, seguido del francés, del alemán, del italiano y del portugués. En esta traducción de textos especializados también se encuentran el rumano, el ruso, el japonés, entre otros, como idiomas de trabajo, pero con una presencia muy inferior.

Una vez presentadas brevemente las características de los encuestados y del mercado de la traducción turística, pasaremos a describir las preferencias de estos profesionales así como sus necesidades con respecto a los recursos terminográficos, en general, y de la traducción turística, en particular.

En su mayoría, los encuestados prefieren el formato en línea (56,47%) a otro tipo de formato electrónico (24,71%) o en papel (18,82%). En la encuesta, no se les preguntó las razones de esta elección a los profesionales, pero podemos afirmar que esta selección se debe principalmente a las ventajas que presentan este tipo de formato, como veíamos en el apartado anterior, a saber: su fácil acceso y manejo, su gratuidad, su actualización, su inmediatez, su posibilidad de realizar diferentes tipos de búsqueda, etc.

Este hecho nos indica que los traductores profesionales necesitan recursos terminográficos de calidad en línea para realizar sus traducciones, por lo que podemos concluir que es necesario este tipo de recursos para satisfacer las necesidades de estos profesionales de la lengua.

Asimismo, los encuestados prefieren los recursos bilingües (39,45%) frente a los recursos monolingües de la lengua de llegada (25,56%) y de partida (24,12%), y los multilingües (10,88%). Este dato nos desvela también una conclusión fundamental a la hora de realizar recursos dirigidos a estos destinatarios: los recursos deben ser bilingües o, en todo caso, monolingües, pero no multilingües, ya que su elección en las respuestas totales de la encuesta es muy baja.

Por otro lado, los profesionales que han respondido a la encuesta prefieren los recursos disponibles (48,75%) (sin indicar el formato del mismo) antes que los recursos creados por ellos mismos (30,88%) o los recursos que proporcionan los clientes de los encargos de traducción (20,37%).

A la pregunta «¿Qué tipo de recurso terminológicos prefiere cuando traduce?», los encuestados respondieron con una selección de los recursos que utilizan de forma preferente durante su trabajo y, a continuación, se pueden observar los resultados:

| | |
|---|--------|
| Diccionario/glosario monolingüe especializado (L1) | 8,63% |
| Diccionario/glosario monolingüe especializado (L2) | 7,83% |
| Diccionario/glosario bilingüe especializado | 18,94% |
| Diccionario/glosario multilingüe especializado | 2,81% |
| Diccionario visual monolingüe | 0,54% |
| Diccionario visual bilingüe | 1,34% |
| Sección Imágenes de un motor de búsqueda (por ej, <i>Google</i>) | 3,55% |
| Búsquedas en motores de búsqueda (por ej, <i>Google</i>) | 16,13% |
| Corpus paralelo (textos originales y sus traducciones) | 5,09% |
| Corpus comparable (textos originales en ambas lenguas) | 4,62% |
| Base de datos terminológica | 8,84% |
| Enciclopedia | 3,08% |
| Wikipedia | 8,63% |
| Listas de distribución | 1,61% |
| Foros de debate en Internet | 4,15% |
| Tesauro | 3,21% |
| Otros | 1,00% |

Tabla 10. Porcentajes obtenidos de la pregunta «¿Qué tipo de recurso terminológicos prefiere cuando traduce?».

Según estos resultados, los recursos preferidos de los encuestados son los diccionarios bilingües especializados, con un 18,65%, seguido por las búsquedas en los

motores de búsquedas, como por ejemplo *Google*, con un 16,13%. Por su parte, las bases de datos terminológicas ocupan el tercer lugar de preferencia de los traductores, con un 8,84%, seguido por los diccionarios monolingües (L1), con un 8,63%, y la Wikipedia, que ha obtenido un 8,63%. Detrás de estos, se encuentran el resto de recursos que se observan en la Tabla 10, los cuales se ofrecían también como opción en la pregunta.

Entre los resultados obtenidos por las respuestas de los encuestados, observamos un dato curioso referente al recurso Wikipedia, ya que, según estos datos, ocuparía el quinto lugar de entre los recursos preferidos por los traductores profesionales. En este sentido, se convierte en un recurso empleado frecuentemente por profesionales de la traducción y, por tanto, en una posibilidad nueva de búsqueda de información, a pesar de las numerosas críticas que recibe debido a su falta de fiabilidad en algunos de sus contenidos.⁵⁰⁷

Otro dato interesante de esta encuesta es el referente a lo que opinan los encuestados sobre los recursos terminográficos disponibles en la actualidad. El 49,48% de ellos indican que algunos de los recursos son buenos, mientras que el 20,57% indican que es difícil encontrar recursos de calidad. Asimismo, el 9,11% indica que prefiere elaborar sus propios recursos terminográficos, lo que provoca una pérdida de tiempo muy importante, especialmente si se tiene en cuenta que estos profesionales trabajan normalmente con mucha presión y con un tiempo muy limitado.

Con respecto a la evaluación de los recursos terminográficos disponibles, el 38,71% de los encuestados indica que no utiliza un recurso si no es fiable frente a un 33,71% que considera que no siempre es posible utilizar recursos fiables y un 19,11% que afirma que no realiza siempre una evaluación del recurso.

⁵⁰⁷ Actualmente se están llevando a cabo investigaciones sobre la aplicación de Wikis en el campo de la Traducción, tanto en lo que respecta a su aplicación profesional como en la enseñanza. El principal impulsor de estas investigaciones es A. Désilets, un investigador del Institute for Information Technology, de Canadá. En el grupo de investigación *Traducción y Lexicografía*, de la Universidad de Málaga, dirigido por la Dra. Gloria Corpas Pastor, se están realizando también estudios sobre la aplicación de este recurso en la enseñanza, cuyos resultados se publicarán próximamente.

Para conocer los criterios que utilizan los encuestados para determinar la fiabilidad y la calidad de los recursos, se les pidió en la encuesta que ordenaran del 1 al 7 los criterios de evaluación utilizados en la plantilla de evaluación presentada en este estudio. Según los resultados, los encuestados se fijan, en primer lugar, en la autoría; en segundo lugar, en la especialización de la página web; en tercer lugar, en la variedad de información proporcionada; en cuarto lugar, en la actualización del recurso; en quinto lugar, en la facilidad de acceso; en el sexto lugar, en los comentarios externos realizados sobre el recurso y, en último lugar, en las instrucciones de uso.

Para terminar, comentaremos la pregunta clave de esta encuesta: «¿Qué considera que debería incluir un recurso terminológico dirigido a traductores?». Las respuestas han variado en algunas preguntas pero, por lo general, han coincidido en lo que debería considerarse información necesaria, información deseable e información innecesaria, como se puede observar en la siguiente tabla:

| | Información necesaria | Información deseable | Información innecesaria |
|---|-----------------------|----------------------|-------------------------|
| Un mayor número de entradas | 50,94% | 44,77% | 4,29% |
| Una mayor variedad de unidades (nombres, verbos, adverbios, adjetivos, nombres propios) | 33,79% | 55,22% | 10,99% |
| Abreviaturas y acrónimos | 55,26% | 41,78% | 2,96% |
| Definiciones | 74,86% | 24,04% | 1,09% |
| Definiciones más clara | 69,15% | 30,32% | 0,53% |
| Definiciones en ambas lenguas (si es bilingüe) | 45,11% | 45,38% | 9,51% |
| Información etimológica | 6,65% | 38,23% | 55,12% |
| Pronunciación | 5,40% | 33,52% | 61,08% |
| Silabación | 4,29% | 25,43% | 70,29% |
| Ejemplos | 57,33% | 40,53% | 2,13% |
| Una mayor variedad de ejemplos | 29,21% | 60,11% | 10,67% |
| Ejemplos tomados de contextos reales | 56,06% | 40,70% | 3,23% |
| Unidades fraseológicas | 51,51% | 41,92% | 6,58% |
| Derivados y nombres compuestos | 22,35% | 41,92% | 6,58% |
| Especificaciones de dominio | 54,97% | 64,80% | 12,85% |
| Más equivalentes de traducción | 34,60% | 60,49% | 4,90% |
| Una explicación de cada equivalente de traducción | 28,69% | 58,22% | 13,09% |
| Información gramatical (parte de la oración, inflexión, número, género) | 26,94% | 50,83% | 22,22% |
| Instrucciones de uso | 13,76% | 53,65% | 32,58% |
| Información semántica (sinónimos, antónimos, hiperonimia, homonimia) | 21,17% | 63,79% | 15,04% |
| Imágenes | 13,56% | 61,58% | 24,86% |

Tabla 11. Porcentajes obtenidos en la pregunta «¿Qué considera que debería incluir un recurso terminológico dirigido a traductores?» de la encuesta.

Según se observa en la tabla, los traductores consideran como contenido necesario en las fuentes terminológicas un mayor número de entradas, la inclusión de abreviaturas y acrónimos, definiciones, definiciones más claras, la aparición de ejemplos tomados de contextos reales y de información fraseológica.

De la misma manera, consideran como contenido innecesario la información etimológica, la silabación y la pronunciación, por lo que estos datos no deberían ser incluidos en los recursos cuyos destinatarios fueran los propios profesionales de la lengua, principalmente los traductores.

Por último, se encuentran los datos deseables, es decir, los datos que los encuestados no han considerado como esenciales pero que sí les parecería adecuado que se recogieran en los recursos terminográficos. Algunos de estos datos están muy igualados con los datos obtenidos en lo que se ha considerado como información necesaria, como por ejemplo en la opción de definiciones en ambas lenguas (45,11% y 45,38%). Aquí, se ha producido una división de la opinión de los encuestados, por lo que no se puede determinar con exactitud cuál es la opción más adecuada para estos profesionales. No obstante, las dos respuestas han sido positivas, por lo que la inclusión de las definiciones en ambas lenguas sería acogida positivamente por casi la totalidad de los encuestados.

De esta manera, a través de esta pregunta y de sus resultados, se pueden establecer las necesidades específicas que presentan los traductores profesionales, es decir, la información que aparece en la columna de «Información necesaria» y que no debería faltar en ningún recurso dirigido a estos usuarios, así como la información que consideran irrelevante y que, por tanto, no sería necesario incluir en estos recursos. Por su parte, la opción «Información deseable» proporciona información acerca de lo que se podría incluir en un recurso de este tipo, pero que no es considerada como básica para los traductores. Así pues, podríamos dejar esta información a la decisión de los terminógrafos que se encarguen del trabajo sistemático para la creación de recursos terminográficos dirigidos a este grupo de destinatarios.

Para finalizar la encuesta, se incluyó una pregunta abierta, a fin de darles la oportunidad a los informantes de proporcionar algún comentario más que no estuviera

recogido en las preguntas anteriores. La pregunta era la siguiente: «¿Tendría alguna otra sugerencia sobre el contenido que debería incluir una fuente terminológica para traductores?». La mayoría de las respuestas ofrecidas en esta pregunta fueron muy interesantes y ofrecieron luz a algunos aspectos que no se habían incluido en la encuesta, que se habían incluido de forma diferente, etc. Al no poder incluir aquí todas las respuestas obtenidas, haremos mención a las más frecuentes y, en nuestra opinión, más relevantes:

- Exportabilidad a otros formatos, como TXT o TBX.
- Explicaciones y ejemplos sobre el uso de unidades terminológicas, por ejemplo: esta unidad no se debería utilizar en este contexto, falso cognado, uso obsoleto, etc.
- Información de la unidad en otros recursos (referencias, páginas web, textos, etc.).
- Diferencias culturales entre la unidad y su equivalente y variaciones regionales.
- Enlaces a otros recursos.
- Accesibilidad fácil y rápida.

Como principales conclusiones de este estudio, podemos extraer las que presentamos a continuación de forma enumerada:

- Los traductores profesionales deben considerarse como un grupo de usuarios de recursos terminográficos distintos de otros grupos, como pueden ser los especialistas del campo de especialidad, el público en general, los estudiantes, etc.
- Los traductores profesionales necesitan unos recursos específicos que satisfagan sus necesidades de consulta y que les permitan alcanzar los mejores resultados sin la necesidad de consultar numerosas fuentes.
- La mayoría de los traductores trabajan en una gran variedad de campos de especialidad y, por tanto, no pueden ser especialistas reales en todos los

campos en los que traducen. Por ello, su trabajo se basa principalmente en los recursos terminográficos que consultan, los cuales no son apropiados en su gran mayoría para asistir a los traductores profesionales.

- Los recursos específicos de los traductores profesionales deben incluir información lingüística (definiciones, equivalentes, colocaciones, contextos, etc.), así como información pragmática (referencias culturales, consejos de uso, especificaciones de dominio, etc.). El resto de información, como la información semántica, es deseable y, por tanto, puede ser complementaria, pero no es esencial.
- Los traductores profesionales necesitan recursos que ofrezcan una forma fácil y rápida de consulta a través de Internet, así como enlaces externos para facilitar la complejidad o mejora de la información encontrada.
- Las definiciones de los recursos deben ser correctas y concretas, a fin de mejorar la comprensión del traductor de forma rápida.
- Los traductores profesionales emplean diferentes recursos para buscar la información necesaria, siendo la Wikipedia uno de los recursos nuevos más utilizados.
- Los traductores profesionales prefieren la consulta de recursos ya elaborados y disponibles a la elaboración propia de los suyos, debido principalmente al tiempo necesario para llevar a cabo esta terminología puntual.

Asimismo, podemos también extraer otras conclusiones derivadas de la comparación de este estudio con los estudios similares realizados previamente. Destacamos las tres siguientes, que se encuentran relacionadas entre sí:

- Los traductores profesionales presentan necesidades y expectativas diferentes de los traductores no profesionales o estudiantes de traducción (u otros profesionales de la lengua), por lo que es necesario atender estas diferencias a la hora de realizar recursos terminográficos y, por tanto, no generalizar en la muestra utilizada.

- Los traductores profesionales consideran como deseable la información semántica incluida en recursos terminográficos, por lo que se observa la diferencia con los estudios previos, que defendían la incorporación de los tres tipos de información: lingüística, semántica y pragmática.
- Los traductores profesionales consideran la información pragmática como esencial, puesto que es la información que les permite a ellos acceder a la situación comunicativa del término, así como a los posibles contextos.

Para finalizar este apartado, debemos decir que estamos satisfechos con los resultados de esta encuesta, ya que, en nuestra opinión, ha esclarecido las necesidades que presentan los profesionales de la lengua actualmente, se ha recogido su opinión sobre los recursos terminográficos disponibles en la actualidad y se les ha dado la oportunidad de describir su recurso terminológico «ideal».

Una vez que hemos presentado los resultados del estudio llevado a cabo para analizar las necesidades de los traductores profesionales y hemos establecido la clasificación de las fuentes terminológicas que se pueden encontrar en cualquier ámbito de especialidad, realizaremos una revisión de algunos trabajos terminográficos (diferentes tipos de recursos terminográficos, proyectos y herramientas) que se han elaborado para los traductores como el grupo potencial de destinatarios y con el objetivo de satisfacer sus necesidades y mejorar su trabajo diario.

2.5.4. Algunos trabajos terminográficos dirigidos a traductores

The dictionary should encourage and support the creativeness and the autonomy of the translator (Temmerman, 2003: 132).

Al comenzar cualquier trabajo terminográfico es universalmente aceptado que el primer paso consiste en establecer el grupo potencial de destinatarios del recurso que se va a elaborar, es decir, es imprescindible determinar si el recurso va a servir para el público en general, para especialistas, para estudiantes, etc. Por ello, hemos realizado

nuestro estudio de las necesidades de los destinatarios de nuestro recurso terminológico. Lamentablemente, se ha obviado frecuentemente uno de los mayores grupos de usuarios de los recursos terminográficos: los traductores profesionales. De esta manera, los traductores, el grupo de destinatarios que utiliza a diario recursos terminográficos de diferentes especialidades y naturalezas y que, como hemos visto con los resultados de la encuesta, presentan unas necesidades específicas, se ven obligados frecuentemente a utilizar recursos que no han sido elaborados para ellos y que, por tanto, no cumplen sus expectativas. Asimismo, se ven obligados a ejercer una labor de terminología puntual para encontrar respuestas y solucionar dudas que surgen durante el proceso de traducción.

Observando estas necesidades y carencias descritas anteriormente, se ha empezado a trabajar y a elaborar herramientas de diferente naturaleza con el objetivo de satisfacer a los usuarios, entre ellos los traductores, en mayor medida. Según una investigación realizada en este sentido, consideramos que estas herramientas se pueden dividir en los siguientes grupos: 1. Recursos terminográficos (diccionarios, glosarios, bases de datos, etc.) y proyectos que o bien van dirigidos directamente a traductores o contemplan a los traductores como uno de sus posibles usuarios; 2. Portales y recopilatorios de recursos terminográficos que, a pesar de que la mayoría no están dirigidos directamente a los traductores, se han venido utilizando como herramientas útiles por ese grupo de usuarios e incluso se han denominado a sí mismos como *diccionarios de traductores*; 3. Metabuscadores dirigidos a traductores, y 4. Aplicaciones informáticas destinadas a facilitar la labor de consulta de los traductores.

Dentro del primer grupo especificado, podemos distinguir a su vez dos tipos de herramientas de consulta: por un lado, los diccionarios propiamente dichos, según la clasificación realizada en el apartado 2.4.1.1.; y por otro lado, las bases de conocimiento, es decir, los trabajos terminográficos basados en la representación del conocimiento de un campo de especialidad determinado (cf. 1.5.).

Como ejemplo del primer subgrupo, podemos citar el *Dictionnaire analytique de la distribution. Analytical dictionary of retailing*, de J. Dancette y C. Réthoré (2000),⁵⁰⁸ que tienen como grupo potencial de usuarios a los profesionales del comercio, a los estudiantes y profesores de esta especialidad, a los traductores y redactores de inglés a francés, y el *Diccionario Analítico de la Globalización del Trabajo* (2009), de J. Dancette y R. Castrillón, cuyo grupo potencial son los profesores y estudiantes de relaciones laborales, traductores/terminólogos y actores económicos y sociales que trabajen en las lenguas inglés, francés y español. Ambos diccionarios siguen la misma metodología, que consiste en la elaboración de un diccionario con rasgos de enciclopedia y de ontología que pretende responder a la necesidad de información y de clarificación de los conceptos básicos del campo de especialidad en cuestión (Dancette y Castrillón, 2009: 65). De esta manera, ambos diccionarios parten de una metodología semasiológica y una ordenación alfabética, pero, a la vez, se incorporan algunas ventajas que presenta el trabajo onomasiológico (ordenación conceptual de los términos), con el fin de facilitar la comprensión del universo conceptual de un término (Dancette y L'Homme, 2001: 3) y de estimular, así, la autonomía del traductor a la hora de utilizar el equivalente de traducción.

A continuación, comentaremos un poco más en profundidad el segundo diccionario citado por su reciente elaboración. En este diccionario, los propios autores describen que su formato está basado en una base de conocimientos terminológica donde la estructura que se genera es una gran red de términos relacionados. Una de las novedades de esta base es el hecho de que resalta los enlaces semánticos entre los términos y despliega formalmente estas relaciones en tablas en el cuerpo de las entradas. El *Diccionario* cuenta unas 4.000 entradas en inglés, francés y español, lo que hace un total aproximado de 10.000 términos si contamos con los sinónimos de algunas unidades. La originalidad de este *Diccionario* es el gran número de tipos de relaciones semánticas que se pueden consultar, puesto que cuenta aproximadamente con 1.000 relaciones de equivalencia, 1.330 relaciones jerárquicas, cerca de 11.000 relaciones

⁵⁰⁸ Actualmente este diccionario está disponible tanto en papel como en formato electrónico a través de la siguiente URL: <<http://olst.ling.umontreal.ca/dad/>>.

asociativas de todas las clases, incluyendo treinta relaciones sintagmáticas (verbo y adjetivo) (ibíd.: 75). Esta herramienta se presenta como un recurso útil para el público interesado en dicho campo, incluidos los traductores, puesto que ofrece información enciclopédica apoyada con una representación conceptual, a fin de facilitar la comprensión de las unidades terminológicas y de que los menos especialistas puedan comprender fácilmente de qué se habla en el texto de origen.

Como ejemplos del segundo grupo clasificado, las bases de conocimiento, podemos citar los trabajos que presentamos en el capítulo anterior (cf. 1.5.). Así pues, tenemos en este grupo, el proyecto ONCOTERM, de la Universidad de Granada; el proyecto Genoma-KB, de la Universidad Pompeu Fabra; entre otros. Todos estos proyectos, al igual que el grupo anterior, tienen como objetivo representar el dominio de especialidad mediante una estructura conceptual y con definiciones enciclopédicas. La diferencia entre un grupo y otro es la presentación de cada uno (según la clasificación presentada en el apartado 2.4.1.1): por un lado, tenemos las características propias de un diccionario, y por otro, las de una base de conocimientos.

Este grupo de herramientas son útiles con respecto a la información pragmática que ofrecen y que es tan necesaria para los traductores. No obstante, no todas cumplen las necesidades de este grupo de usuarios o siguen sus expectativas, ya que la mayoría de los traductores profesionales no consideran la información semántica contenido esencial de los recursos terminográficos. En este sentido, las bases de conocimiento o los diccionarios basados en ontologías puede incluir en algunos casos excesiva información para estos usuarios, teniendo en cuenta que estos requieren definiciones claras y concisas.⁵⁰⁹

A continuación, comentaremos el segundo grupo en el que hemos dividido estos recursos para traductores: los portales y recopilatorios de recursos terminográficos. Estos recursos son los que más abundan en Internet en la actualidad y consisten en una recopilación de diccionarios, glosarios, bases de datos, etc. de diferentes idiomas, combinaciones y especialidades. Aunque la mayoría no están dirigidos directamente a

⁵⁰⁹ Las necesidades y expectativas de los traductores profesionales se trata en profundidad en el siguiente capítulo a través de la presentación de un estudio llevado a cabo como parte de esta investigación.

los traductores, sí que facilitan la búsqueda terminológica mediante la oferta de multitud de fuentes terminológicas de diferentes especialidades, idiomas, naturaleza, etc., lo que satisface una de las necesidades imperantes de los traductores profesionales: la consulta rápida y fácil de recursos. Asimismo, también tiene en cuenta uno de los principales problemas de los traductores profesionales: la falta de recursos completos y adecuados para su consulta y, por tanto, la obligación de consultar numerosos recursos para garantizar una traducción de calidad.

Como decíamos, existen numerosos recursos de este tipo, por lo que aquí solo citaremos algunos de ellos como ejemplos:

1. Directorios de diccionarios generales y especializados:

- *Yourdictionary*,⁵¹⁰ un directorio con unos 2.500 diccionarios y gramáticas en 300 lenguas.
- *Onelook.com*,⁵¹¹ un directorio con cerca de 900 diccionarios.
- *Diccionarios.com*,⁵¹² un directorio de Larousse que ofrece diferentes tipos de búsquedas en diccionarios.
- *Lexicool*,⁵¹³ un directorio con más de 7.000 enlaces de diccionarios y glosarios bilingües y multilingües.

2. Portales lingüísticos y terminológicos:

- *Foreignword.com*,⁵¹⁴ un portal que recoge información sobre bancos de datos terminológicos, glosarios, diccionarios, información sobre traductores, traductores automáticos, etc.

⁵¹⁰ <<http://www.yourdictionary.com>>

⁵¹¹ <<http://www.onelook.com/browse.shtml>>

⁵¹² <<http://www.diccionarios.com>>

⁵¹³ <<http://www.lexicool.com>>. Se trata de un directorio de «todos» los diccionarios y glosarios en línea bilingües y multilingües que se pueden consultar gratuitamente por Internet. Este recurso va dirigido a los traductores, a los lingüistas, a los estudiantes de lenguas y a todos los que se interesan por los idiomas.

⁵¹⁴ <<http://www.foreignword.com/>>

- *TermSciences*,⁵¹⁵ un portal terminológico multidisciplinario desarrollado por INIST (Institut de l'Information Scientifique et Technique francés), cuyo objetivo consiste en recopilar y proporcionar recursos terminográficos del sector público para crear un recurso terminológico común de referencia. Permite la consulta de términos en francés, inglés, español y alemán.
 - *Terminology Collection*,⁵¹⁶ un portal de la universidad de Vaasa (Finlandia), donde se ofrece una recopilación de diccionarios especializados y generales monolingües y de diferentes combinaciones lingüísticas.
 - Diccionarios de TermCat,⁵¹⁷ un centro de terminología catalán que se dedica a la elaboración de recursos terminográficos para ponerlos a disposición de los usuarios.
 - *Vademécum del traductor externo*,⁵¹⁸ un portal del Parlamento Europeo donde se ofrecen recursos varios (diccionarios especializados, boletines terminológicos y normativos, artículos, etc.) para facilitar la labor del traductor externo de la institución.
3. Otros recursos:
- *Visual dictionary*,⁵¹⁹ un portal que permite la búsqueda y consulta de diccionarios visuales de diferentes especialidades.
 - *Glossarist*,⁵²⁰ un buscador de glosarios de diferentes idiomas y especialidades en Internet.

En general, estos directorios de recursos, portales, buscadores, etc. permiten la búsqueda y consulta de gran número de recursos terminográficos, así como el acceso a otro tipo de información y aplicaciones (traducción automática, diccionarios de sinónimos, etc.).

⁵¹⁵ <<http://www.termsciences.fr/>>

⁵¹⁶ <<http://www.uwasa.fi/viestintatieteeet/terminology/>>

⁵¹⁷ <<http://www.termcat.cat/productes/>>

⁵¹⁸ <http://www.europarl.europa.eu/transl_es/plataforma/pagina/celter/celula.htm>

⁵¹⁹ <<http://www.infovisual.info/>>

⁵²⁰ <<http://www.glossarist.com/>>

Sin embargo, la gran mayoría de los glosarios o diccionarios que está indexado en un directorio no cumplen las necesidades y expectativas de los traductores profesionales que veíamos como resultado de nuestra encuesta, por lo que la recopilación de estos recursos no cumple totalmente su función de servir como recurso terminológico pleno y adecuado. A pesar de este inconveniente, es necesario reconocer las ventajas de consulta que suponen estos directorios o portales lingüísticos y terminológicos, frente a la búsqueda individual y ardua de los traductores a través de Internet, puesto que agiliza el acceso a la información desde un único punto de acceso.

El tercer grupo de aplicaciones que vamos a comentar hace referencia a los metabuscadores, es decir, a buscadores especializados en un determinado aspecto, ya sea la temática, combinación de lenguas, etc. En este caso, vamos a dar como ejemplo dos metabuscadores diferentes, que presentamos a continuación:

Por un lado, el metabuscador *Tradudidática*, un buscador especializado en recursos de apoyo a la traducción de textos con la combinación alemán-español. Este recurso ha sido desarrollado por el profesor E. Martínez Belchí en el marco de sus asignaturas de traducción español-alemán y cuyo objetivo consiste en ayudar a sus alumnos durante las traducciones de clase. Se trata de un recurso muy simple, puesto que consiste en ofrecer acceso a varios recursos terminológicos para dicha combinación, pero a la vez muy útil y práctico, ya que facilita el acceso a los recursos más frecuentes utilizados para este tipo de traducción.



Figura 17. Metabuscador *Tradudidáctica*.

Por otro lado, encontramos el buscador bilingüe *2lingual*,⁵²¹ que no es una herramienta dirigida a los traductores, aunque puede ser de gran utilidad para facilitar la búsqueda documental de estos. Esta aplicación es una sencilla herramienta web que realiza una búsqueda bilingüe de forma simultánea en sitios web y portales bilingües de Google. Esto permite que las búsquedas sean más rápidas y que, además, estén bien agrupadas en una única ventana. Esta aplicación funciona de forma similar a un buscador normal, a saber: se introduce un término en el cuadro de búsqueda, se selecciona la combinación de idiomas que se desea y, para terminar, se pulsa el botón «2lingual search» (búsqueda bilingüe). A continuación, aparecerá una ventana dividida en dos partes: a la izquierda, los enlaces que corresponden con el término buscado, y a la derecha los enlaces que contienen el equivalente del término buscado.

⁵²¹ <<http://www.2lingual.com/>>

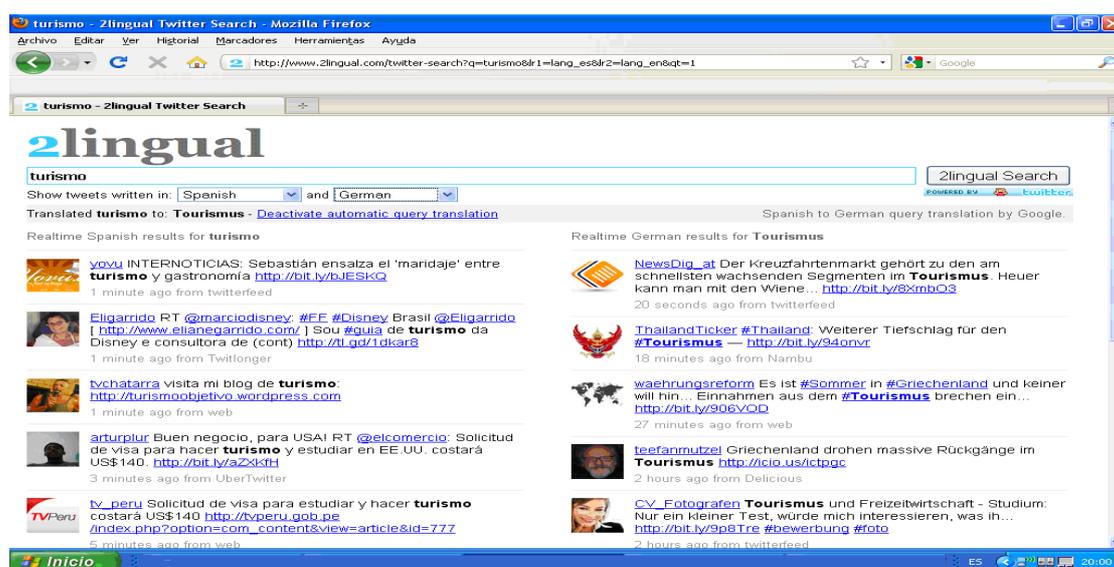


Figura 18. Resultado de búsqueda en el motor de búsqueda 2lingual para el término «Turismo» en español y alemán.

Actualmente cuenta con alrededor de 550 combinaciones diferentes de lenguas disponibles para realizar búsquedas en este buscador. Aunque es una herramienta que facilita la labor de la documentación bilingüe, se trata de una herramienta muy reciente y que, por tanto, necesita mejorar sus resultados de búsqueda para ofrecer resultados más precisos y adecuados.

Por último, llegamos al último grupo: aplicaciones informáticas destinadas a facilitar la labor de consulta de los traductores. En este contexto, debemos decir que cada vez se están elaborando más aplicaciones de este tipo, gracias a la rápida evolución que hay en el campo de la informática y de las tecnologías. No obstante, con relación a la asistencia terminológica para traductores, estas herramientas siguen siendo bastante escasas. En este caso, vamos a comentar tres: *IntelliWebSearch*, una herramienta diseñada especialmente para los traductores y para satisfacer sus necesidades de consulta; y *Linguee*, un “autodenominado” diccionario inteligente que combina un buscador y un diccionario bilingüe.

En primer lugar, trataremos *IntelliWebSearch*,⁵²² una herramienta gratuita para Windows cuyo objetivo consiste en ahorrar tiempo de búsquedas terminológicas a través de Internet. Es una aplicación creada por M. Farrel, un traductor técnico de italiano a inglés, que pone a disposición de todos los traductores esta herramienta, así como una sección de ayuda, un foro de discusión para usuarios y diferentes cursos de formación en línea para aprender a utilizar adecuadamente la aplicación. La idea de esta herramienta consiste en reducir al máximo los pasos necesarios para efectuar búsquedas. Autodenominado «el paraíso del buscador de términos o el mejor buscador de terminología», *IntelliWebSearch* puede configurarse según las necesidades del usuario en cuanto a sitios para la búsqueda (diccionarios, glosarios, vocabularios, bases de datos, portales, etc.), idioma de la interfaz y también asignar combinaciones de teclas diferentes a las predeterminadas tanto para la búsqueda típica o como para la completa (denominada en el programa *Plurisearch*) entre otras. Actualmente, los idiomas disponibles son el inglés británico, el alemán, el danés, el francés, el italiano, el español y el catalán; y sus principales tareas son las siguientes:

- Realizar búsqueda simple (en un único recurso);
- Realizar varias búsquedas simultáneas (en varios recursos predefinidos);
- Copiar texto para la búsqueda; y
- Buscar el texto seleccionado.

El uso y la configuración de esta herramienta son muy accesibles, por lo que resulta de gran utilidad a los traductores (y a cualquier otro usuario) a la hora de realizar búsquedas terminológicas, puesto que es posible consultar simultáneamente diferentes recursos terminográficos previamente determinados por el usuario. Además de la facilidad de uso, esta herramienta presenta otras ventajas: la instalación es muy sencilla y rápida, la interfaz dispone de varios idiomas, el foro de discusión y la ayuda permiten resolver cualquier duda que pueda surgir, la selección de los recursos la realiza el propio usuario, entre otras.

⁵²² <<http://www.intelliwebsearch.com/>>

En segundo lugar, trataremos brevemente la aplicación alemana *Linguee*,⁵²³ creada por G. Frahling y L. Fink y que se encuentra disponible en la Red desde la primavera de 2008. Según los propios autores, esta herramienta se trata de un diccionario inteligente que pretende sustituir a los diccionarios tradicionales electrónicos con sus posibilidades de búsqueda al proporcionarles una asistencia más completa a la traducción y redacción de textos en lengua extranjera. Se trata de una aplicación similar al buscador *2bilingual*, ya que utiliza los sitios web y portales bilingües para extraer información y, además, permite la búsqueda de palabras y frases en gran cantidad de textos bilingües en el par de lenguas seleccionado.⁵²⁴ Sin embargo, es una herramienta bastante más elaborada y que, además, filtra la información obtenida mediante un algoritmo de aprendizaje automático que determina la calidad de los textos para crear su propio corpus interno.⁵²⁵

Al igual que *2lingual*, la búsqueda inicial con *Linguee* se realiza con la introducción de la palabra o la expresión que se desea buscar. Una vez introducida la unidad, el programa realiza las búsquedas en su corpus interno, formado por las páginas web bilingües filtradas con anterioridad, y ofrece una lista de oraciones en el par de lenguas elegido que contienen la palabra o expresión de manera exacta o aproximada separadas en dos grupos: por un lado, los resultados que ya han sido validados y que, por tanto, son fiables aparecen con una señal en verde; y por otro, los resultados que se encuentran a la espera de validación se indican con un triángulo amarillo de aviso. El resultado de la búsqueda permite observar el contexto en el que se emplea una expresión traducida, la frecuencia de dicha traducción y ejemplos de traducciones realizadas con anterioridad para la misma unidad en la lengua origen, siempre partiendo de traducciones realizadas por humanos, es decir, *Linguee* rechaza cualquier traducción que haya sido realizada de forma automática y que encuentre en la Red. Además de las ventajas que hemos mencionado, esta aplicación permite obtener resultados de

⁵²³ <<http://www.linguee.es>>

⁵²⁴ Las combinaciones inglés - español, alemán - inglés, inglés - francés e inglés - portugués son las que se encuentran disponibles en estos momentos, aunque en breve se ofrecerán otras combinaciones con el japonés y el chino.

⁵²⁵ Actualmente, el corpus de trabajo de *Linguee* consta de 100 millones de traducciones en total.

traducciones de términos poco comunes que no encontrará en ningún otro lugar debido a su gran número de entradas y su constante crecimiento.

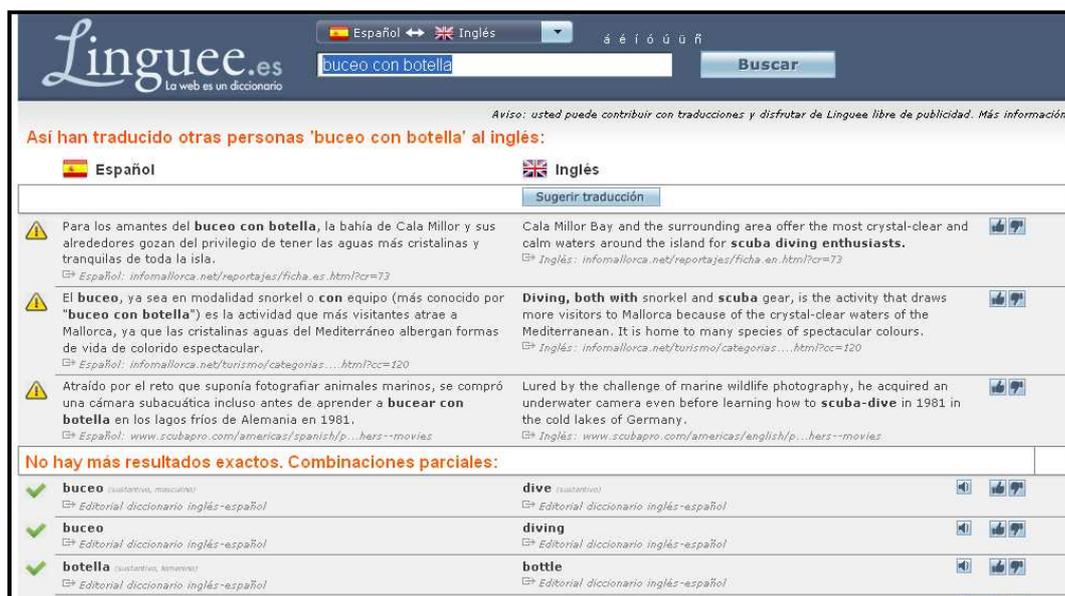


Figura 19. Fragmento de la ventana con los resultados de la búsqueda “buceo con botella” en el par de lenguas español-inglés en *Linguee*.

Como se puede suponer, cualquiera de las aplicaciones que hemos comentado en este apartado son de gran ayuda para los traductores, ya sean profesionales o no, aunque aún consideramos necesario la necesidad de seguir trabajando en esta línea y mejorando este tipo de tecnologías para asistir a este grupo de destinatarios. En esta línea, se están realizando intentos para mejorar los recursos terminográficos dirigidos a traductores con el fin de incrementar la calidad de sus traducciones. Pero, a pesar de haber realizado avances en este ámbito y de empezar a prestar atención a los traductores profesionales como grupo potencial de destinatarios, esto todavía supone muy poco, si tenemos en cuenta que el número de recursos terminográficos disponibles no cubren todos los campos de especialidad, no cubren las necesidades de los traductores, no cubren las combinaciones más comunes y necesarias, no siguen las pautas de los trabajos sistemáticos, entre otros aspectos.

Para fomentar la creación de recursos terminográficos dirigidos a traductores y para homogeneizar la metodología empleada para elaborarlos, presentaremos a

continuación una propuesta de metodología, dividida en diversas fases de trabajo, para llevar a cabo un trabajo ontoterminográfico sistemático basado en corpus dirigido a este grupo potencial de usuarios, los traductores profesionales. Con esta metodología, pretendemos cumplir los requisitos establecidos por los especialistas en terminología y satisfacer, además, las necesidades y expectativas de los traductores profesionales que nos han indicado a través de la encuesta llevada a cabo en este trabajo de investigación.

2.6. METODOLOGÍAS PARA EL TRABAJO ONTOTERMINOGRÁFICO SISTEMÁTICO BASADO EN CORPUS

Not only can terminology support ontology engineering, ontologies can in turn underpin the work of terminographers and lexicographers in different phases of their work, e.g. in websearches during corpus compilation; in corpus analysis and data mining. Tools at the service of terminographers and lexicographers need ontological underpinning (Temmerman y Kerremans, 2003: 4).

Llegados a este punto, y después de haber realizado una revisión exhaustiva de la terminología y de su rama aplicada, la terminografía, así como tras haber llevado a cabo un estudio sobre los diferentes recursos terminográficos disponibles en la actualidad y las necesidades que presentan los traductores profesionales con relación a dichos recursos, pasaremos a exponer una propuesta metodológica para la elaboración de recursos terminográficos dirigidos principalmente a nuestro grupo de destinatarios objeto de estudio: los traductores profesionales.

Antes de presentar nuestra metodología, es necesario destacar que el objetivo principal de esta es la elaboración de recursos terminográficos sistemáticos y descriptivos, es decir, recursos que abarcan un campo de conocimiento determinado en su totalidad (en este caso, el turismo de aventura) y que siguen una serie de criterios para la realización de las diferentes fases de trabajo y que, además, no persiguen una finalidad normalizadora e impositiva, sino más bien de descripción de la realidad objeto

de estudio. Asimismo, nuestra metodología va dirigida a la elaboración de recursos para un grupo de destinatarios específico, los traductores profesionales, aunque también se podría aplicar a cualquier otro grupo siempre y cuando se realizaran las modificaciones pertinentes para adaptarse a las necesidades de estos nuevos grupos.

A fin de delimitar nuestro modelo de metodología, hemos consultado previamente otras propuestas realizadas hasta la fecha, las cuales nos han permitido conocer el trabajo que se ha venido realizando en los últimos años y beneficiarnos, así, de algunas de sus aportaciones en el ámbito terminográfico. De esta manera, llevamos a cabo un escrutinio de diferentes metodologías que, siguiendo la clasificación de Cabré Castellví (1993: 45), se pueden dividir en tres grupos: en primer lugar, la corriente normalizadora,⁵²⁶ que surge a partir de la metodología propuesta por Auger y Rousseau (1987), concebida para equiparar el uso del francés en las áreas técnicas y científicas con el inglés en la región de Quebec, la cual sigue el centro de terminología TERM-CAT,⁵²⁷ para la planificación lingüística en Cataluña; en segundo lugar, la corriente estandarizadora,⁵²⁸ donde encontramos la metodología de trabajo que propone la ISO en las siguientes normas: ISO 10241: 1992 *International terminology standards – Preparation and layout* e ISO 704: 2000. *Terminology Work. Principles and methods* y, también, la utilizada para la elaboración de la base de datos terminológica TERMIUM, de la oficina de traducción del gobierno de Canadá,⁵²⁹ y, por último, la

⁵²⁶ La corriente normalizadora se observa principalmente en los países que llevan a cabo planes de normalización de su lengua propia, que se considera minoritaria, ya sea dentro del mismo país o en el panorama internacional. Surge como iniciativa gubernamental con el objetivo de reconocer la situación de un determinado nivel de la lengua y, así, identificar vacíos gramaticales y lexicográficos mediante la reglamentación oficial.

⁵²⁷ <<http://www.termcat.cat/>>

⁵²⁸ La corriente estandarizadora hace referencia a la terminología que llevan a cabo principalmente en las organizaciones de normalización y en las grandes empresas multinacionales, con el fin de obtener una terminología normativa y estandarizada que muestre la univocidad entre el concepto y el término.

⁵²⁹ <<http://www.btb.termiumplus.gc.ca/site/termium.php?lang=eng&cont=001>>

corriente traduccional,⁵³⁰ que hace referencia a la metodología de trabajo seguida en los organismos internacionales como la ONU, EU, UNESCO, FAO, etc.⁵³¹

Además de estas tres corrientes metodológicas, que se encuentran vinculadas a organismos oficiales y multilingües, encontramos otras propuestas que presentan generalmente menor alcance. Entre ellas, podemos destacar la metodología sistemática y genérica que nos propone Cabré Castellví (1993: 290 y ss.), en la cual se hace patente su teoría terminológica (TCT) a través del interés que muestra hacia la adecuación del trabajo al tema, a los destinatarios, a la función y a los objetivos planteados, así como las propuestas de Sager (1990) o Arntz y Picht (1995).

Todas estas propuestas metodológicas para la creación de recursos terminológicos presentan una serie de puntos en común, aunque también algunas discrepancias. Como puntos en común, podemos citar los siguientes:

- Todas dividen el trabajo terminográfico en tres grandes fases generales, a saber: i) Concepción del trabajo, donde incluyen la definición y preparación del trabajo; ii) Elaboración del trabajo, donde se realizan las labores de recopilación de información y la gestión terminológica; y iii) Revisión/supervisión del trabajo, donde se solventan los posibles problemas y discrepancias, se corrigen los errores e incoherencias, etc.
- Todas basan el trabajo terminográfico en corpus textuales compilados según unos criterios internos y externos específicos.
- Todas recogen el empleo de herramientas informáticas para la realización de las tareas de cada fase en mayor o menor medida (extracción terminológica, búsqueda de concordancias y equivalentes, etc.).

No obstante, como decíamos, también presentan sus divergencias, especialmente por lo que respecta a la función que desempeña el trabajo terminográfico elaborado

⁵³⁰ La corriente traduccional se orienta hacia la fijación de equivalencias terminológicas en las distintas lenguas, con objeto de asistir al traductor durante el proceso de traducción y de contribuir a la calidad del texto traducido.

⁵³¹ Para consultar estas metodologías de forma exhaustiva, véase Gómez González-Jover (2006: 184 y ss.), donde se ofrece una descripción específica de cada una de las metodologías indicadas aquí.

(planificación lingüística, estandarización, traducción, etc.). Por lo que respecta a estas diferencias, consideramos las siguientes como las más destacadas:

- Cada propuesta utiliza su propia ficha terminológica de vaciado, incluyendo más o menos campos según el objetivo del trabajo terminográfico.
- Cada metodología divide las fases generales en diferentes subapartados y, en este sentido, no siempre coinciden todos sus componentes. Por ejemplo, para la primera fase, la relacionada con la preparación del trabajo, unas incluyen el análisis de los recursos terminográficos y lexicográficos disponibles del dominio de especialidad y otras, no.
- Las metodologías utilizadas por organismos internacionales, ya sean de normalización o no, tienden a la búsqueda de la univocidad y a la eliminación de la variación (denominativa, semántica, etc.). Por el contrario, tanto la metodología que propone Cabré Castellví (1993: 291) así como otras consultadas defienden la necesidad de atender a la variación y a la existencia de la falta de univocidad absoluta entre el concepto y el término, así como entre términos de diferentes lenguas.
- De nuevo, las metodologías utilizadas por organismos internacionales tienden a la unificación de los objetivos y de los destinatarios, por lo que no hacen ninguna distinción a la hora de elaborar sus recursos terminográficos. Sin embargo, las demás metodologías tienden a prestar atención a la pragmática del recurso y, por tanto, a adecuar su resultado final a los destinatarios, función, contexto, etc.

De las metodologías mencionadas, la más próxima a nuestra postura es el modelo propuesto por Cabré Castellví (ibid.) (Figura 20), ya que se trata de una metodología para un trabajo descriptivo y enmarcada en la TCT, por lo que tiene en cuenta los puntos más relevantes de la terminología moderna, a saber: el aspecto comunicativo de la terminología, la variación terminológica, el empleo de corpus textuales, la adecuación a las necesidades de los destinatarios y los diferentes objetivos que presenta el trabajo terminográfico.

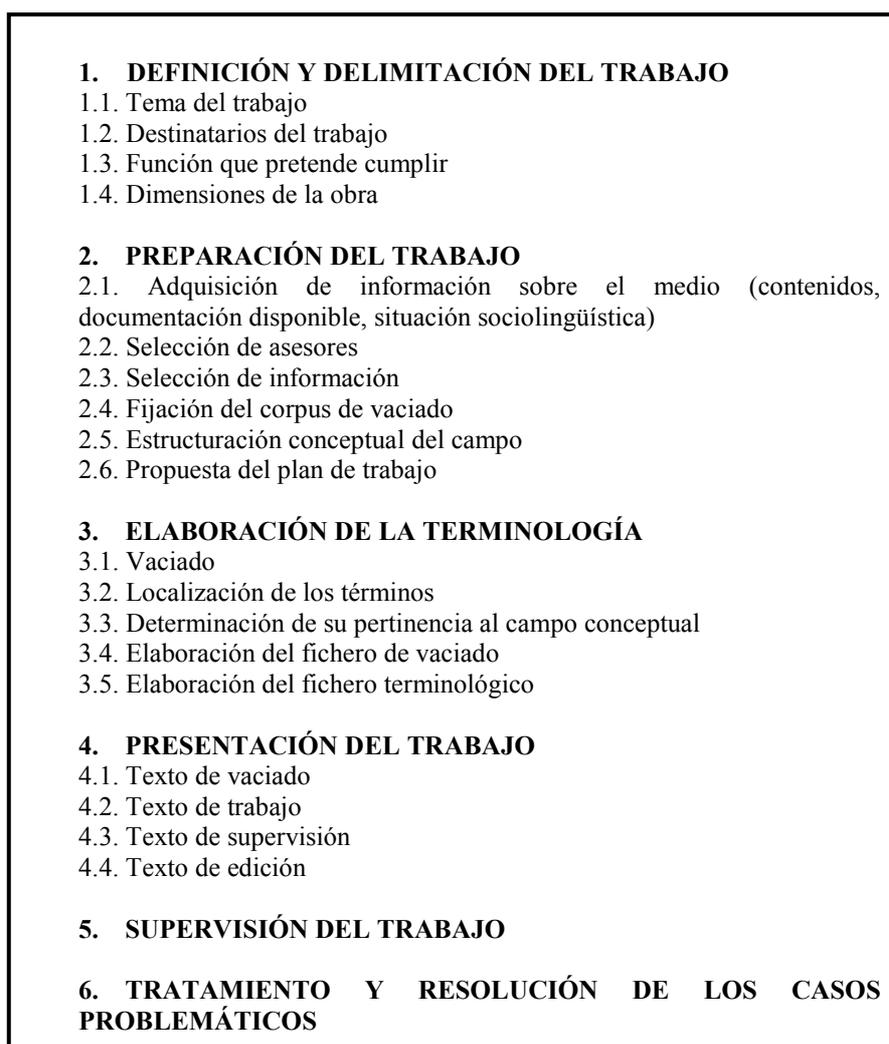


Figura 20. Metodología del trabajo terminográfico sistemático y genérico propuesto por Cabré Castellví (1993: 290 y ss.).

A pesar de tratarse de una metodología muy completa, ya que incorpora todos los aspectos de la terminología moderna y realiza una descripción pormenorizada del trabajo terminográfico, no hace mención al trabajo ontoterminográfico, es decir, no incorpora el uso de las ontologías a la labor terminográfica, sino una estructuración conceptual básica del campo de especialidad en forma de mapa conceptual en la segunda fase. En este sentido, se aleja de la finalidad de nuestra investigación y nos obliga a considerar otras propuestas metodológicas más concretas. No obstante, esta propuesta nos ha servido de guía a la hora de elaborar nuestra propia propuesta y de diseñar algunas fases de nuestro trabajo metodológico.

Las metodologías que sí han ocupado un lugar destacado en nuestra tesis doctoral han sido, por un lado, la *Termonography* (Temmerman y Kerremans, 2003) y, por otro, la propuesta por Faber Benítez y su equipo de investigación para sus proyectos de investigación ONCOTERM, PUERTOTERM y MARCOCOSTA (cf. 1.5.2.) (Faber Benítez y Jiménez Hurtado, 2002).⁵³² Estas dos metodologías emplean el uso de las ontologías en el trabajo terminográfico, a fin de realizar una conceptualización más completa del dominio de especialidad en cuestión a la par que elaborar definiciones más controladas y precisas. A pesar de presentar ciertas coincidencias, también cuentan con algunas diferencias, que las hacen únicas en su aplicación.

En primer lugar, trataremos la *Termonography* (termontografía), una aproximación multidisciplinar que consiste en el desarrollo de una metodología terminográfica que combina los supuestos teóricos y los métodos propuestos en la Terminología Sociocognitiva de la Terminología (TST) (Temmerman, 2000) (cf. 1.3.2.2.3.) con los métodos y directrices de la ingeniería ontológica (Gómez Pérez et ál., 1996; Fernández et ál., 1997; Sure y Studer, 2003).

El enfoque termontográfico (Temmerman y Kerremans, 2003) nació en el seno del proyecto FFPOIROT⁵³³ de la mano de los investigadores colaboradores del grupo de investigación del *Centrum voor Vaktaal en Communicatie* (CVC),⁵³⁴ de la Erasmushogeschool de Bruselas. Su enfoque es eminentemente funcional, ya que el contenido y la estructura del recurso 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.

⁵³² Otra metodología similar a estas dos es la propuesta por el grupo de investigación Condillac, de la Universidad de Saboya, aunque difiere en algunos puntos básicos de estas metodologías y, por ello, no la hemos incluido aquí. Esta metodología defiende un trabajo completamente onomasiológico para la puesta en práctica de la *Ontoterminology* (ontoterminología); en otras palabras, al contrario que estas dos metodologías expuestas en este apartado, que parten de un corpus de textos especializados para realizar el trabajo ontoterminográfico (es decir, parten de un enfoque semasiológico), esta metodología parte de los conceptos para elaborar la ontología de dominio y, posteriormente, identifica los términos adecuados para denotar cada concepto (y, en caso necesario, proponer nuevos términos normalizados) (cf. Roche et ál., 2009). En este sentido, se aproxima a los principales fundamentos de la terminología tradicional.

⁵³³ Véase nota a pie 126.

⁵³⁴ <<http://taalkunde.ehb.be/cvc>>

Este enfoque se puede resumir como sigue: en primer lugar, es necesario contar con una estrecha colaboración con los especialistas del dominio de especialidad en cuestión, a fin de desarrollar un marco de categorías específicas del dominio y de relaciones intercategoriales.⁵³⁵ A continuación, el marco categorial elaborado sirve de medio para la extracción manual y semiautomática del conocimiento del corpus textual compilado para el trabajo terminográfico en cuestión. Posteriormente, se enriquece y se refina la categorización elaborada en el primer paso a partir del conocimiento extraído del corpus textual utilizado y se reflejan las especificidades culturales de cada lengua (ibíd.: 665). Debido a este trabajo circular, esta metodología se considera como un enfoque *middle-out*, puesto que combina los enfoques *top-down* y *bottom-up*.

En Temmerman y Kerremans (ibíd.) y Kerremans et ál. (2003), se presenta el flujo de trabajo de la termontografía, el cual se encuentra organizado en tres grandes fases, que a su vez se dividen en seis fases más concretas (Figura 21).

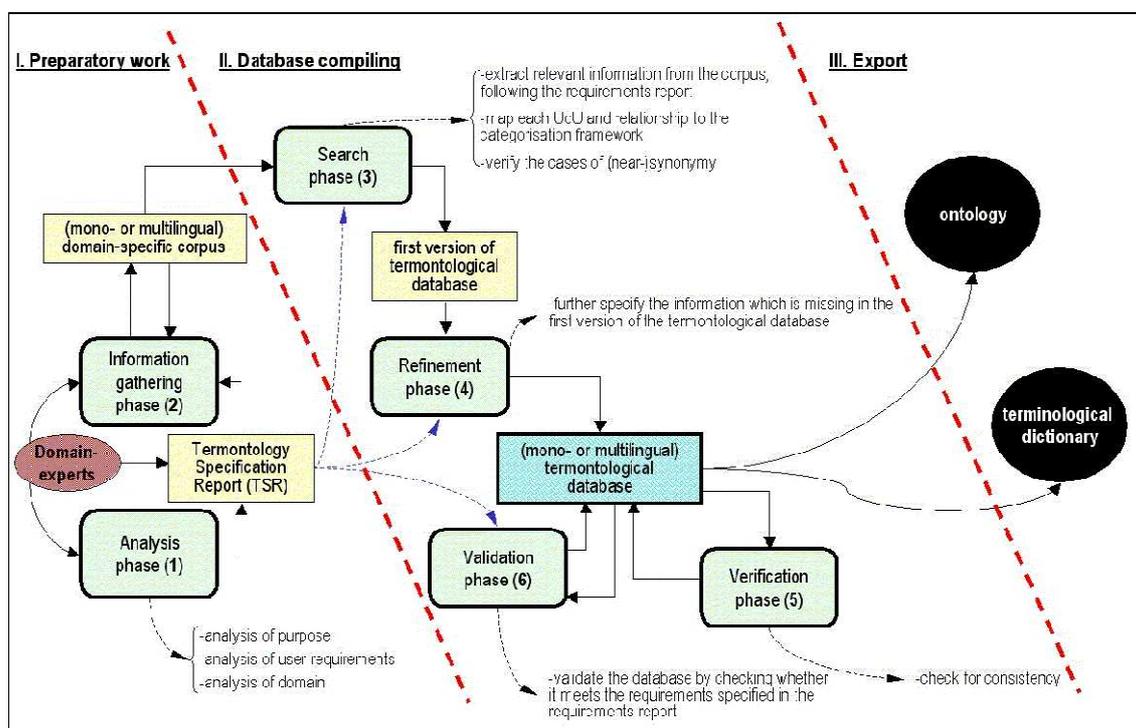


Figura 21. Flujo de trabajo de la metodología termontográfica (Temmerman y Kerremans, 2003).

⁵³⁵ En este contexto, «categoría» hace referencia a los conceptos prototípicos del dominio.

Como decíamos, el trabajo termontográfico se lleva a cabo en tres fases diferentes: en primer lugar, tenemos la fase de preparación del trabajo, que incluye la fase de análisis y recopilación de la información; en segundo lugar, la fase de compilación de la base de datos, que consiste en la búsqueda, refinamiento, verificación y validación de la información extraída e introducida en la base de datos; y, por último, la fase de exportación, que se trata de la fase donde se exporta la información incluida en la base de datos (monolingüe o multilingüe) a una ontología o a un diccionario terminográfico. A continuación, describiremos las diferentes fases:

- Fase de análisis. En esta fase, se debe especificar la finalidad del proyecto, el alcance del dominio de especialidad y los requisitos del grupo de destinatarios. Esta información se especifica en el informe denominado *Termonology Specification Report* (TSR). En esta fase, se elabora también la categorización del dominio de especialidad.
- Fase de compilación de la información. Tanto la categorización establecida como el TSR ayudarán al terminógrafo en su investigación para compilar el corpus textual pertinente.
- Fase de búsqueda. A partir del corpus compilado en la fase anterior, el terminógrafo debe extraer los términos referentes al dominio de especialidad y vincularlos a sus categorías correspondientes, así como las relaciones intercategoriales reflejadas en los patrones verbales. Esta fase da lugar a una primera versión de la base de datos termontográfica.
- Fase de refinamiento. Esta fase tiene como objetivo completar la información contenida en la base de datos termontográfica de la fase anterior. Por ejemplo, con la alineación de equivalentes de traducción, con la especificación de contextos o concordancias en las que aparecen los términos, etc.
- Fase de verificación. En este paso, se revisa la coherencia y la corrección de la base de datos elaborada hasta el momento y se realizan comprobaciones de diferente naturaleza, como por ejemplo si todos los términos tienen asignada una categoría.

- Fase de validación. Como último paso antes de la exportación, el terminógrafo debe comprobar que el trabajo que se ha llevado a cabo cumple todos los requisitos establecidos en el informe TSR inicial.

Después de haber expuesto las diferentes fases en las que se divide el trabajo termontográfico propuesto por Temmerman y Kerremans (2003) y Kerremans et ál. (2003), observamos las diferencias y similitudes existentes con las propuestas metodológicas para el trabajo terminográfico. De las diferencias, destacamos la fase de análisis y la fase de búsqueda, puesto que son las dos fases que comprenden acciones relacionadas con las ontologías: la primera, a la hora de elaborar la categorización, y la segunda, cuando extrae las relaciones intercategoriales y vincula los términos y dichas relaciones con la categorización realizada previamente.

A pesar de estas diferencias, podemos afirmar que en el resto de fases no existen grandes divergencias con las otras metodologías, en concreto con la de Cabré Castellví (1993), ya que ambas incluyen la fase de análisis, de recopilación de información, de validación, etc. y están enfocadas a las funciones y destinatarios de los productos terminográficos en los que trabajan. En otras palabras, la esencia del trabajo sistemático en terminología se mantiene, ya sea con la incorporación o no de las ontologías.

Como indicábamos anteriormente, otra metodología que ha influido en nuestro modelo es la propuesta por Faber Benítez y su equipo de investigación para sus proyectos de investigación ONCOTERM, PUERTOTERM y MARCOCOSTA (cf. 1.5.2.).⁵³⁶ En este caso, esta metodología se denomina «gestión de terminología orientada a procesos» (Faber Benítez et ál., 2005) y surge como una solución para alcanzar una verdadera estructuración del conocimiento dentro de una disciplina. Según estos investigadores, «no basta con obtener unas listas de términos y delimitar sus posibles relaciones jerárquicas, ya que son demasiado estáticas y no ofrecen la verdadera visión dinámica que tiene cualquier área del saber» (Senso et ál., 2007: 594). En este contexto, las categorías conceptuales se representan mediante marcos semánticos (cf. 1.3.2.3.3.1.), que dan lugar a esquemas conceptuales prototípicos más complejos del conocimiento

⁵³⁶ Aunque algunos aspectos de estos proyectos difieren entre sí, todos comparten el mismo modelo metodológico.

especializado y que tienen como objetivo la representación de los procesos dinámicos dentro de cada dominio concreto. En lo que respecta al nivel de la microestructura, las definiciones siguen una plantilla definicional diseñada para cada categoría según los esquemas conceptuales creados previamente (cf. León Araúz et ál., 2009: 47).

En esta metodología no se establece explícitamente una fase previa de análisis y de preparación del trabajo como habíamos observado en las anteriores, aunque sí que se reconoce la necesidad de establecer el grupo de destinatarios y las funciones que desempeñará el trabajo. De esta manera, consideran que el trabajo ontoterminográfico está constituido por las siguientes fases:⁵³⁷

- En la primera fase, se lleva a cabo la recopilación de información, tanto mediante la compilación de un corpus textual como mediante el estudio de los recursos terminográficos existentes del dominio en cuestión. La información recopilada se almacena y comienza su procesamiento.
- En la segunda fase, se emplean diferentes herramientas informáticas para extraer los candidatos a término⁵³⁸ y, a través de una lista de frecuencia y de búsquedas de colocaciones, se infiere el conocimiento especializado del dominio en cuestión, siempre con la colaboración de expertos en la materia.
- En la tercera fase, después de haber identificado los conceptos y sus relaciones semánticas, se lleva a cabo el modelado conceptual de la ontología, que articula todo el conocimiento de la disciplina. Esta representación se ejemplifica en forma de marco semántico, ya que permite describir los procesos o acciones que tienen lugar dentro de cada materia de forma dinámica y completa.

⁵³⁷ No hemos encontrado ninguna fuente donde se especifiquen claramente las diferentes fases en las que se basa esta metodología, por lo que esta división se trata de una propuesta propia basada en diferentes publicaciones de los investigadores de estos proyectos (López Rodríguez et ál., 2001; Faber Benítez y Jiménez Hurtado, 2002; Faber Benítez, 2004; Faber Benítez et ál., 2005; Senso et ál., 2007; León Araúz et ál., 2009).

⁵³⁸ En los tres proyectos mencionados se emplean las herramientas del paquete informático *Wordsmith Tools*, programa desarrollado por M. Scott (1997), de la Universidad de Liverpool, y distribuido por Oxford University Press. Se puede descargar en este enlace: <<http://www.lexically.net/wordsmith/>>.

- En la cuarta fase, se crean las entradas terminológicas en la base de datos y se elaboran las definiciones terminográficas pertinentes, que están basadas en el marco conceptual previamente elaborado y en una plantilla definicional, que servirá de base para elaborar todas las definiciones de los términos de una misma categoría. A la par, se introduce el resto de la información terminológica (equivalentes, contextos, categoría gramatical, concordancias, etc.) y multimedia (imágenes, vídeos, etc.), a fin de completar las entradas de cada término.

Como hemos observado, esta metodología cuenta con varias novedades con respecto a las metodologías que hemos comentado anteriormente. Estas novedades están relacionadas principalmente con la forma en la que se representa el conocimiento especializado mediante marcos semánticos y con el empleo de la representación conceptual para crear sus plantillas definicionales.

No obstante, cabe destacar el empleo en todas las propuestas de los corpus textuales como uno de los medios para adquirir el conocimiento especializado, la utilización de herramientas informáticas para facilitar la labor de los terminógrafos, así como la necesidad de incorporar información pragmática (contexto, concretamente) a las entradas terminológicas.

Una vez que hemos revisado algunas de las metodologías existentes y, concretamente, dos de las más influyentes en el marco de la ontoterminografía, pasaremos a describir nuestra propuesta metodológica y sus puntos más relevantes, a fin de fijar nuestro modelo antes de comenzar con nuestro trabajo ontoterminográfico sobre el turismo de aventura.

2.6.1. PROPUESTA DE METODOLOGÍA PARA EL TRABAJO ONTOTERMINOGRÁFICO SISTEMÁTICO BASADO EN CORPUS DIRIGIDO A TRADUCTORES

En este apartado presentaremos nuestra propuesta metodológica para la creación sistemática de recursos ontoterminográficos dirigidos a traductores. Así pues, podemos

resumir los puntos más relevantes con los que va a contar este modelo de la siguiente manera:

- El trabajo es sistemático, puesto que va a tratar toda la terminología del ámbito del turismo de aventura y, además, va a seguir una serie de pasos establecidos previamente para ello.
- El trabajo se enmarca dentro de la ontoterminografía, es decir, combina los criterios teóricos de la Terminología moderna, especialmente de la Teoría Comunicativa de la Terminología (Cabré Castellví, 1993) y de la Teoría Sociocognitiva de la Terminología (Temmerman, 2000), con los métodos de la ingeniería ontológica.
- El trabajo está dirigido a los traductores profesionales como grupo potencial de destinatarios, puesto que pretende satisfacer sus necesidades y adaptarse a sus expectativas, según la información obtenida en la encuesta realizada dentro de esta investigación (cf. 2.5.3.). Además, se pretende facilitar la consulta de estos usuarios a este tipo de recursos mediante la creación de una aplicación informática que mejore y agilice la búsqueda y consulta de la información terminológica.⁵³⁹

Asimismo, y derivado de la revisión realizada previamente a las diversas metodologías:

- El trabajo presenta un enfoque *middle-out*, es decir, combina el estudio de los términos en contextos comunicativos concretos (enfoque *bottom-up*) con la elaboración de una categorización inicial realizada por expertos en la materia (enfoque *top-down*). Asimismo, al ser un enfoque intermedio, el proceso es circular, para llevar a cabo un refinamiento de la categorización y una búsqueda más exhaustiva de términos.

⁵³⁹ La consulta se supone fuera de la metodología terminográfica y, en efecto, no forma parte de ella. Sin embargo, sirve de poco el hecho de realizar un trabajo (onto)terminográfico exhaustivo si, posteriormente, los usuarios encuentran la consulta a la información lenta, difícil, compleja, etc. En nuestro caso, nos hemos propuesto considerar la consulta como un aspecto importante del trabajo ontoterminográfico y, así, consideramos la fase de edición del recurso esencial para satisfacer las necesidades del grupo de usuarios. Por ello, hemos trabajado en una aplicación informática nueva, que presentamos en el capítulo cuarto (cf. 4.3).

- La representación conceptual del dominio del turismo de aventura se plasma mediante marcos semánticos, que se establecen a partir de la conceptualización inicial y de las relaciones semánticas obtenidas del corpus textual.
- Al tratarse de un trabajo ontoterminográfico bilingüe, se emplean dos tipos de subcorpus: por un lado, un subcorpus comparable, formado por textos originales en las lenguas de trabajo (en nuestro caso, en inglés, alemán y español); y por otro, un subcorpus paralelo, constituido por componentes paralelos de textos originales y sus respectivas traducciones.⁵⁴⁰ Esta combinación de corpus se establece con dos objetivos: en primer lugar, extraer posibles candidatos a equivalentes del subcorpus paralelo; y en segundo lugar, comprobar la validez de esos términos extraídos en sus contextos comunicativos reales y originales y completar el esquema categorial y el marco semántico del dominio con el subcorpus comparable.
- Se realizan estudios de términos *in vivo*, es decir, en sus contextos reales, e *in vitro*, es decir, dentro de recursos terminográficos, legislación, etc.
- Las definiciones terminográficas están basadas en un lenguaje controlado y bien definido a través de las relaciones semánticas detectadas y mediante la utilización de unas plantillas definicionales para cada categoría.

Para cumplir estos puntos básicos de nuestra metodología, hemos dividido el trabajo en seis fases principales: 1) la fase de concepción del trabajo y análisis del dominio de especialidad; 2) la fase de compilación del corpus especializado; 3) la fase de gestión ontoterminográfica; 4) la fase de elaboración de la base de datos ontoterminográfica; 5) la fase de resolución de problemas y validación, y 6) la fase de edición. Estas fases las englobamos a su vez en tres macrofases diferentes: en primer lugar, la fase de preparación del trabajo, que hace referencia a las fases de concepción del trabajo y análisis del dominio de especialidad y de compilación del corpus de trabajo; en segundo

⁵⁴⁰ Dependiendo de los objetivos planteados, este subcorpus paralelo puede ser unidireccional, es decir, contiene textos originales de una misma lengua y sus traducciones en las restantes, o bidireccional, si incluye textos originales en las diferentes lenguas de trabajo y sus respectivas traducciones.

lugar, la fase de realización del trabajo ontoterminográfico, que consta de las demás fases del trabajo a excepción de la fase de edición, que se considera como la tercera macrofase, donde se lleva a cabo la exportación de la base de datos y cualquier trabajo relativo a su publicación o disponibilidad de consulta. El orden establecido para estas seis fases no implica que se deba seguir una secuencia estricta, puesto que, por ejemplo, las fases de gestión ontoterminográfica y de elaboración de la base de datos son a menudo simultáneas o, en ocasiones, algunas fases se realizan con posterioridad o anterioridad.

Estas fases están formadas por tareas más concretas, que comentamos a continuación junto con las fases a las que pertenecen.

2.6.1.1. Fase I: Delimitación del trabajo y análisis del dominio

Esta primera fase⁵⁴¹ hace referencia a la fase preparatoria del trabajo y, en ella, se debe tomar una serie de decisiones y establecer unos criterios que servirán de base para la estrategia de trabajo que se mantendrá a lo largo de todo el proceso. En este esquema de trabajo inicial se fijarán algunos aspectos, como los destinatarios, las lenguas de trabajo, etc., pero habrá otros que puedan sufrir modificaciones o adaptaciones conforme avanza la investigación. Esto podría ocurrir, por ejemplo, en caso de que consideremos necesario optimizar cualquier aspecto que no se haya previsto, tanto en cuestiones relacionadas con las tareas como en otras vinculadas a los recursos humanos y técnicos, el tiempo, el producto, etc.

Esta primera fase está compuesta por dos etapas bien diferenciadas: por un lado, encontramos la parte de concepción o definición del trabajo; y por otro, la fase de análisis del dominio de especialidad en el que se va a trabajar.

⁵⁴¹ En algunas metodologías no se contempla esta fase o no se le otorga la importancia que, en nuestra opinión, tiene y que nos ha hecho considerarla como una parte más del trabajo.

En la primera etapa, es necesario establecer las variables pragmático-lingüísticas⁵⁴² que caracterizarán al recurso y que se mantendrán a lo largo de todo el proceso de trabajo, ajustándose siempre a los objetivos de nuestra investigación. Estas variables son principalmente el tema que se va a tratar, la lengua del recurso, el grupo potencial de destinatarios al que va destinado, la función que desea cubrir, la aplicación que se desea dar al recurso y sus dimensiones. Todas estas variables son claras y deben quedar resueltas antes de comenzar propiamente el trabajo ontoterminográfico. Por ejemplo, es imprescindible determinar si un recurso va a ser monolingüe o multilingüe, si va a ir dirigido a estudiantes o a especialistas de un dominio determinado, si su aplicación se basa en la codificación o descodificación del mensaje, etc. Asimismo, es necesario concretar las siguientes cuestiones (IULA, 2008):

1) La macroestructura del trabajo, que se puede dividir en cuatro bloques:

- Tipos de unidades terminológicas seleccionadas: términos de igual o distinto nivel de especialización (especializados, semi-especializados o divulgativos), información sintagmática o paradigmática, diferentes unidades gramaticales (nombres, adjetivos, verbos, etc.), etc.
- Orden de presentación: alfabético, temático o híbrido; e
- Inclusión de anexos: prólogo, introducción, bibliografía, árbol de campo, índices, etc.

2) Los campos que forman la microestructura de cada entrada terminológica, es decir, la información lingüística, pragmática y semántica que se va a ofrecer en cada registro. Al existir tal cantidad de campos que se pueden incluir en una entrada terminológica,⁵⁴³ es imprescindible que se tengan en cuenta aquí las necesidades de los usuarios potenciales, por lo que se recomienda el análisis de las necesidades de estos usuarios antes de comenzar cualquier trabajo (onto)terminográfico.

⁵⁴² Las variables que se deben fijar en esta fase coinciden con las características manejadas durante la propuesta de clasificación de las fuentes terminológicas presentadas en este mismo capítulo (cf. 2.4.1).

⁵⁴³ En la norma ISO 12620 (1999) se indican los nombres normalizados de todas las categorías de datos que se pueden emplear en un trabajo terminológico.

3) Soporte y tipología de su edición: papel (diccionario, libro, fichero, al final de un libro, etc.) o electrónico (base de datos en CD-ROM, en línea con o sin hipervínculos, etc.).

Una vez que se han fijado las variables anteriores, también es importante definir el equipo de trabajo y la selección de asesores, así como los recursos (materiales, económicos y de tiempo) con los que se cuenta. Con respecto al factor humano, conviene asegurarse de la colaboración de personas con un dominio experto en las lenguas de trabajo (si se trata de un recurso multilingüe), así como en el campo de especialidad en el que se está trabajando, a fin de que el terminógrafo o su equipo pueda realizar consultas a estos expertos y de que el trabajo sea revisado y validado por estos.⁵⁴⁴ En lo que respecta al resto de recursos, como por ejemplo las decisiones relativas a las herramientas informáticas que se van a utilizar, dependerán en gran medida de las posibilidades de cada proyecto, por lo que no se pueden establecer de forma estricta. No obstante, cabe destacar que se deben tener muy en cuenta a la hora de planificar la duración, el alcance y los resultados del proyecto en su conjunto, puesto que, de lo contrario, podría dar lugar a un fracaso del proyecto completo.

Después de finalizar la etapa de delimitación del trabajo, podemos comenzar con el análisis del dominio de especialidad, la segunda etapa de esta primera fase, que consiste en el estudio del dominio de forma amplia y desde diferentes vertientes: sociolingüística, económica, organizativa e incluso traductológica, si se trata de un recurso multilingüe, a fin de que el terminógrafo se aproxime a la situación del dominio en ese momento y adquiera un conocimiento general de las posibles necesidades y dificultades del trabajo que va a realizar. Para ello, deberá consultar a expertos, revistas especializadas, informes sobre la situación, entidades relevantes en el dominio, legislación existente (si fuera pertinente), etc.

Esta fase ha de servirle, además, para concretar la conceptualización inicial del campo de especialidad objeto de estudio y para establecer, así, un esquema categorial

⁵⁴⁴ Aunque lo más aconsejable es que el trabajo sea realizado por un equipo mixto (Termcat, 1990: 16), formado por terminológicos o especialistas del dominio, en la mayoría de los casos no es así y se suele contar con asesores externos que colaboran puntualmente con la resolución de dudas y, sobre todo, en la fase inicial para la conceptualización del dominio y en la fase de revisión del trabajo.

prototípico, es decir, una organización de las categorías principales del dominio en cuestión en forma de esquema.⁵⁴⁵ En caso de que el terminógrafo no sea un experto en la materia, su trabajo deberá estar respaldado por la colaboración de expertos en el dominio, por documentación especializada, por legislación pertinente, así como por los recursos terminográficos disponibles de calidad, si los hubiera. En este sentido, se considera el comienzo del trabajo ontoterminográfico desde un enfoque *top-down*.

Este esquema categorial deberá ser considerado como provisional en esta fase por dos razones principalmente: por un lado, las categorías y las relaciones intercategoriales que lo forman se concretarán y completarán posteriormente mediante el estudio del corpus textual que tiene lugar en la siguiente fase, a través del cual se analizará la comunicación especializada en un contexto real y se extraerán los conceptos y términos relevantes del dominio en cuestión; y por otro, la dinamicidad que caracteriza cualquier ámbito de especialidad obligará a la actualización de dicho esquema a lo largo del proyecto, así como después de su finalización, con el fin de captar las variaciones que se van produciendo en dicho dominio. Asimismo, es imprescindible tener en cuenta que el esquema categorial que se elabore, ya sea en esta fase provisional o más adelante, no se puede considerar como la única opción posible, sino que pueden confeccionarse diferentes estructuras según los autores, según la aproximación, etc. Y, debido a ello, se recomienda tener en cuenta los esquemas que se hayan podido realizar con antelación, ya que pueden resultar de enorme utilidad a la hora de elaborar el correspondiente a un trabajo ontoterminográfico nuevo.

A su vez, el terminógrafo deberá realizar una recopilación y una valoración de los recursos terminológicos disponibles en ese momento para ese dominio de especialidad concreto, a fin de conocer qué existe en el mercado y establecer posible materia de ayuda y conocer sus deficiencias. Es decir, deberá realizar búsquedas y evaluaciones de los diferentes recursos que encuentre, tanto en papel como en formato electrónico, antes de comenzar con su trabajo terminográfico, a fin de conocer las

⁵⁴⁵ Este esquema puede ser de diferentes tipos: lineal, en forma de grafos o de marcos semánticos. En nuestro caso, hemos elegido este último por considerar que se trata, en nuestra opinión, del tipo de esquema más completo y expresivo de la organización de un dominio. Para consultar la puesta en práctica de la metodología propuesta aquí y, por tanto, del esquema categorial elaborado, véase el capítulo cuarto.

necesidades del mercado y saber qué debe ofrecer su recurso para tener éxito. Además, este análisis de los recursos disponibles permitirá realizar más adelante un estudio *in vitro* de la terminología del sector, así como de las definiciones y demás informaciones incluidas en dichos recursos.

Con todo ello, se pretende alcanzar el objetivo de conocer *grosso modo* el dominio de especialidad con el que se va a trabajar, determinar el esquema categorial que servirá de base en las siguientes fases y delimitar claramente la temática del trabajo, especialmente si se trata de materias interdisciplinares. En definitiva, el análisis exhaustivo del dominio y la elaboración del esquema categorial permitirán al terminógrafo llevar a cabo de forma eficiente el resto de fases del proyecto de trabajo y, así pues, obtener unos resultados finales acordes con los establecidos en los objetivos iniciales.

2.6.1.2. Fase II: Compilación del corpus especializado

La segunda fase de nuestro modelo metodológico consiste en la compilación de un corpus textual especializado sobre el tema de trabajo en cuestión, a fin de contar con documentación real y fiable para las siguientes fases que conforman el trabajo ontoterminográfico. Además, permitirá la fijación del esquema categorial elaborado previamente y la elaboración del marco semántico para el dominio objeto de estudio. Esta fase, por tanto, es muy significativa para la calidad del trabajo final, puesto que la documentación recopilada servirá de base para la extracción de información necesaria para la elaboración de la base de datos ontoterminográfica, a saber: información lingüística (términos, contextos definatorios), información semántica (relaciones semánticas, conceptos) e información pragmática (frecuencia de uso, situación comunicativa). Por ello, es necesario prestar especial atención a la compilación del corpus textual durante esta fase, ya que, en el caso de que el corpus no se haya compilado de forma adecuada, todo el trabajo ontoterminográfico adolecerá de rigor metodológico y, por tanto, de buena calidad.

Como indicábamos anteriormente, esta metodología defiende el empleo de dos subcorpus diferentes para el trabajo ontoterminográfico bilingüe: por un lado, un subcorpus comparable, formado por textos originales en las lenguas de trabajo (en nuestro caso, en inglés, alemán y español); y por otro, un subcorpus paralelo, constituido por componentes paralelos de textos originales y sus respectivas traducciones, con objeto de llevar a cabo un trabajo ontoterminográfico sistemático y de calidad y de validar en textos originales las diferentes opciones de equivalentes que se detecten en el subcorpus paralelo.

En el protocolo de compilación de cualquier tipo corpus especializado, ya sea comparable o paralelo, se pueden distinguir cinco etapas básicamente: a) el diseño; b) la búsqueda documental y c) la organización de textos,⁵⁴⁶ que comentamos a continuación.

2.6.1.2.1. Diseño del corpus

En la fase del diseño, es imprescindible tener en cuenta las variables lingüístico-pragmáticas determinadas en la fase anterior sobre la delimitación del trabajo, a saber: las lenguas de trabajo, los destinatarios, la finalidad, la temática, etc., puesto que serán los puntos básicos para establecer el diseño de un corpus especializado. De esta manera, una vez establecidas estas variables iniciales, se continuará con el diseño del corpus.

El siguiente paso consiste en el establecimiento de los criterios internos y externos, que deben cumplir todos los textos que se seleccionen en la etapa de búsqueda documental, a fin de perseguir la representatividad del discurso especializado objeto de estudio. Como veíamos en el capítulo primero (cf. 1.4.2.3.), los criterios externos son esencialmente criterios no lingüísticos que sirven para determinar todos los posibles aspectos que influyen en los textos que se van a incorporar al corpus y delimitar las posibles variedades de la lengua, tanto en lo que respecta al discurso especializado como a otros aspectos relevantes; mientras que los criterios internos o lingüísticos, son los que permiten identificar los textos en función de una serie de categorías o atributos

⁵⁴⁶ Esta división del protocolo de compilación está basada en las propuestas establecidas en *Corpas Pastor* (2001, 2002, 2004a, 2004b), Seguiri Domínguez (2006, 2008) y Atkins y Rundell (2008).

lingüísticos, como por ejemplo la formalidad o informalidad de un texto según su léxico.

Los principales criterios externos que han de tenerse en cuenta a la hora de compilar cualquier corpus textual, y especialmente un corpus textual especializado son los siguientes:

- Criterio temporal, que está basado en la variación diacrónica de la lengua. Debido a los cambios que se producen en la lengua, tanto general como en cualquier lenguaje especializado, es necesario fijar este criterio, a fin de buscar la mayor representatividad del discurso objeto de estudio. En este sentido, dependiendo de la finalidad del trabajo (onto)terminográfico, se establecerá un criterio temporal diacrónico, si se persigue el estudio de un lenguaje de especialidad a lo largo de un periodo de tiempo, o sincrónico, si se pretende analizar un lenguaje de especialidad durante un momento determinado. Por lo general, los trabajos (onto)terminográficos buscan el estudio de los lenguajes de especialidad en el momento presente, por lo que serán eminentemente estudios sincrónicos de una duración limitada y lo más actual posible, a fin de recoger la terminología actual de un dominio. No obstante, es necesario recordar aquí la dinamicidad característica de la lengua, y concretamente de los lenguajes de especialidad, por lo que el estudio sincrónico podría incluir algunos tintes de diacronía según la duración del proyecto.
- Criterio geográfico, que hace referencia a la variación diatópica de la lengua. Debido a las variaciones geográficas que existen de una misma lengua como por ejemplo el español (español de España frente al español de México), es necesario delimitar la procedencia de los textos que se van a incluir en el corpus especializado, a fin de buscar la coherencia terminológica y de evitar la posible ambigüedad entre una variedad y otra. Esto puede resultar especialmente importante si se trata de recursos multilingües, puesto que el tamaño de este puede ser excesivo. Asimismo, es imprescindible esta delimitación si se trata de recursos dirigidos a determinados grupos de

destinatarios, como son los traductores, puesto que estos trabajan de una variedad lingüística determinada a otra variedad lingüística concreta y, por tanto, no deben encontrar términos de diferentes variedades lingüísticas al realizar sus búsquedas. No obstante, se podrían considerar textos de distintas procedencias, siempre y cuando se indicaran sus diferencias y se tuvieran en cuenta a la hora de seleccionar la terminología y de elaborar el producto final.

- Criterio formal, que se refiere a las variedades diastrática y diafásica de la lengua. La primera hace referencia a los diferentes niveles de abstracción que se pueden encontrar en los lenguajes de especialidad, a saber: altamente especializado, medio-especializado y divulgativo (cf. Tabla 1). Por ello, es necesario determinar aquí el nivel de abstracción al que van a pertenecer los textos que se seleccionen para la compilación del corpus, es decir, qué situación comunicativa van a representar estos textos. Dependiendo de los objetivos del proyecto, el corpus especializado estará formado por textos pertenecientes a los diferentes niveles de especialización o, por el contrario, pertenecerán a algún nivel en concreto. La elección de estos niveles y, por tanto, de los textos, influirá en la terminología que se extraiga y, por tanto, será necesario introducir una serie de indicaciones en el producto final, con objeto de especificar el uso de cada unidad terminológica o las diferentes opciones según el nivel. Con respecto a la variedad diafásica, se deberán establecer los siguientes puntos: el canal utilizado (oral, escrito, mixto), el medio empleado (páginas web, manuales, revistas, entrevistas orales, multimedia, etc.), la relación entre el emisor y el receptor, entre otros. De esta manera, se caracteriza la situación comunicativa y los actores que intervienen en ella y, así, se determinarán los tipos de textos que se seleccionarán para el corpus textual especializado.
- Tipología textual. Debido a la variedad de géneros textuales que existen en los campos de especialidad, es necesario que el terminógrafo determine cuáles serán los géneros seleccionados para la realización del producto final, que dependerá de los objetivos planteados al comienzo del proyecto. Según estos

objetivos, se seleccionarán los textos pertenecientes a un único tipo textual dentro de un dominio o, por el contrario, se ampliará el abanico y se compilarán textos de diferentes géneros. Asimismo, según la tipología textual elegida, se deberán establecer también las instrucciones para la compilación de textos procedentes de diferentes fuentes, es decir, se determinará si los textos pueden compilarse a partir de fuentes personales (*blogs*, páginas web, etc.), publicaciones (revistas, manuales, libros de consulta, etc.), estatales (organizaciones, fuentes gubernamentales, institutos oficiales, etc.), entre otras, con objeto de diseñar homogéneamente la búsqueda documental.

Por su parte, los criterios internos desempeñan un papel menor en la compilación de corpus con una finalidad (onto)terminográfica, puesto que se basan en aspectos lingüísticos que predominan en los textos, como son los tipos de oraciones (oraciones pasivas, preposicionales, etc.), los tipos de sintagmas (nominal, adverbial, etc., entre otros aspectos. Es decir, son criterios que adquieren importancia en estudios discursivos, sintácticos o incluso morfológicos, pero que carecen de gran importancia en nuestro contexto y que, además, se pueden considerar incluidos en el momento en el que los criterios anteriormente mencionados delimitan textos que presentan una serie de características lingüísticas en común al pertenecer a un mismo dominio, nivel de especialización, área geográfica y temporal, etc. Por ello, consideramos que no se deben tener en cuenta a la hora de compilar un corpus especializado.⁵⁴⁷ De esta manera, basamos el diseño del corpus únicamente en criterios externos, ya que, como nos indican Atkins y Rundell (2008: 68), «The more usual corpus model is one based on external criteria».

Los criterios establecidos durante esta fase pretenden establecer unos límites para la compilación de un corpus, a fin de que este sea representativo de la parcela lingüística que se pretende analizar y no presente ambigüedades o incoherencias. Sin embargo, la representatividad no se consigue simplemente con el seguimiento de estos criterios y se trata, además, de un aspecto complicado a la hora de compilar un corpus debido a la dificultad de realizar una medición objetiva (cf. 1.4.2.2.1). Esta propuesta

⁵⁴⁷ Para consultar la compilación de corpus textuales basados en criterios internos, véase Biber (1993).

metodológica defiende el empleo de un programa, el programa *ReCor* (Seghiri Domínguez, 2006; Corpas Pastor y Seghiri Domínguez, 2006, 2007a, 2007b, 2010), para llevar a cabo la medición cuantitativa de la representatividad de un corpus, una vez compilado, como veremos al final de la fase de la búsqueda documental.

Una vez que se han establecido los criterios externos, continuamos con la siguiente fase: la búsqueda documental.

2.6.1.2.2. Búsqueda documental

Para la búsqueda documental, el criterio acerca del medio en el que se encuentran los textos adquiere gran importancia, puesto que, dependiendo de dónde haya que buscar, se diseñará una búsqueda u otra. En nuestra opinión, la presteza que se requiere a la hora de examinar un dominio y de extraer la terminología actualizada se consigue principalmente con la búsqueda en la Red, es decir, a través de páginas web, de bibliotecas digitalizadas, de publicaciones electrónicas, de sitios web, etc., especialmente si tenemos en cuenta el volumen de información que existe actualmente. Así pues, se evita la necesidad de buscar información en papel, escanearla, grabarla con un formato que permita el manejo de su contenido,⁵⁴⁸ etc. No obstante, para llevar a cabo las búsquedas en Internet, se necesitan tener muy presentes unos criterios de evaluación de los recursos electrónicos, a fin de conseguir contenido fiable y adecuado para nuestro proyecto. De las numerosas plantillas de evaluación que se han elaborado para este tipo de recursos (Kirk, 1996; Codina, 2000; Larsen, 2001; Palomares Perraut et ál., 2004; Gonzalo García et ál., 2004; Seghiri Domínguez, 2006, entre otros),⁵⁴⁹ podemos extraer los criterios de evaluación más relevantes para nuestro modelo metodológico:

- Autoría. El recurso debe estar avalado o pertenecer a un emisor fiable (entidad pública o privada, autor conocido, revista especializada, etc.) y los autores de

⁵⁴⁸ Por lo general, el formato con el que se graban el material escaneado es de tipo imagen (.TIFF, .JPEG, etc.), que no permite su procesamiento.

⁵⁴⁹ La plantilla de evaluación propuesta en este trabajo de investigación (cf. 2.4.2.1.) tiene como objetivo la evaluación de recursos terminográficos, por lo que no nos serviría para llevar a cabo esta búsqueda.

los textos deben ser nativos de los países seleccionados (para los textos originales). Asimismo, tanto los autores como los recursos en sí deben proceder de los países en los que se ha delimitado el estudio.

- Actualización. El recurso debe estar actualizado o ser de reciente creación, para que contenga terminología actualizada del dominio objeto de estudio.
- Contenido preciso y objetivo. El recurso debe tratar sobre el campo de especialidad en cuestión y debe ofrecer un contenido preciso y libre de subjetividad en lo posible.
- Organización de la información. La información debe estar organizada de forma lógica y debe estar estructurada de manera que sea fácil de consultar.

Una vez establecidos los criterios de evaluación, comenzamos la búsqueda documental. Según Autermühl (2001: 52 y ss.), se pueden distinguir tres tipos diferentes de búsqueda en Internet, a saber: *búsqueda institucional*, que hace referencia a la búsqueda que se realiza en sitios web de instituciones u organizaciones internacionales, como la Organización Mundial del Turismo (OMT)⁵⁵⁰ o la Web oficial del turismo en Andalucía;⁵⁵¹ *búsqueda temática*, normalmente mediante directorios, como *Turistea*⁵⁵² o *Turismoactivo.com*,⁵⁵³ y, por último, *búsqueda por palabra clave*, a través de buscadores generales, como *Google*,⁵⁵⁴ o específicos, como *Travigator*.⁵⁵⁵ Dependiendo de los objetivos del proyecto, serán más eficientes un tipo de búsqueda u otro; no obstante, para proyectos (onto)terminográficos, las búsquedas institucionales y

⁵⁵⁰ <http://www.unwto.org/index_s.php>

⁵⁵¹ <<http://www.andalucia.org/>>

⁵⁵² <<http://www.turistea.com/>>

⁵⁵³ <<http://www.turismoactivo.com/>>

⁵⁵⁴ <<http://www.google.com>>

⁵⁵⁵ <<http://www.travigator.com/>>

temáticas serán las más adecuadas y fructíferas, puesto que evitan la «basura virtual» que se puede encontrar con la *búsqueda por palabra clave*.⁵⁵⁶

Durante esta fase es imprescindible tener en cuenta el aspecto de la representatividad del discurso especializado en cuanto al tamaño del corpus. Como veíamos anteriormente, la medición de esta representatividad ha sido un tanto subjetiva hasta hace algunos años, cuando se presentó un método objetivo y cuantificable para medirla en corpus compilados (Seghiri Domínguez, 2006; Corpas Pastor y Seghiri Domínguez, 2006, 2007a, 2007b, 2010) (cf. 1.4.2.2.1.). Este método consiste en la medición del umbral mínimo de representatividad de un corpus mediante el algoritmo N-Cor, que realiza el análisis de la densidad léxica de un corpus compilado.⁵⁵⁷

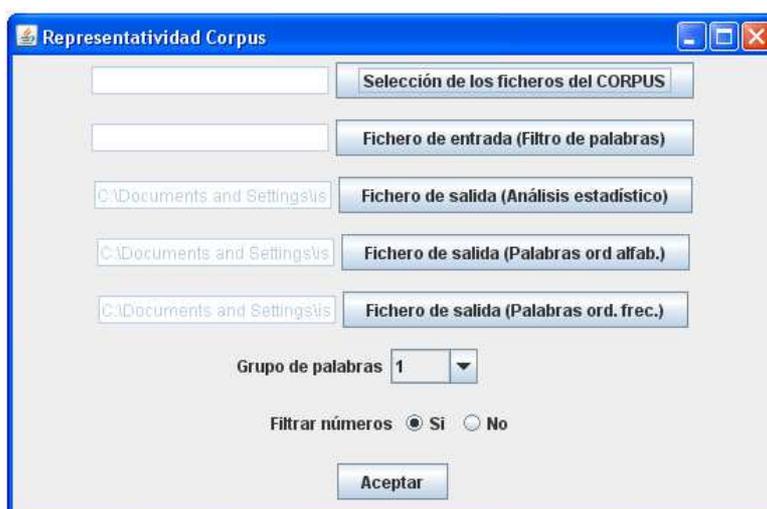


Figura 22. Interfaz del programa *ReCor*.

Como resultado, este programa ofrece unos gráficos en forma de hipérbola, representando la relación invertida que tienen las dos variables: en la primera gráfica, el número de documentos analizados y los *types*, el número de unidades repetidas; y en la segunda gráfica, los *tokens*, el número de unidades totales, y los *types*, de nuevo. Cuanto

⁵⁵⁶ Mediante los criterios de búsqueda avanzada que ofrecen algunos buscadores se puede delimitar la búsqueda y obtener mejores resultados. No obstante, siempre se recupera más información carente de pertinencia que con las otras dos búsquedas.

⁵⁵⁷ Este método se ha visto implementado en la aplicación *ReCor*, que se encontrará disponible en breve en la dirección URL: <http://www.recorweb.com>.

más se acerque la gráfica a la línea horizontal (número de documentos o *tokens*), mayor será la representatividad del corpus.

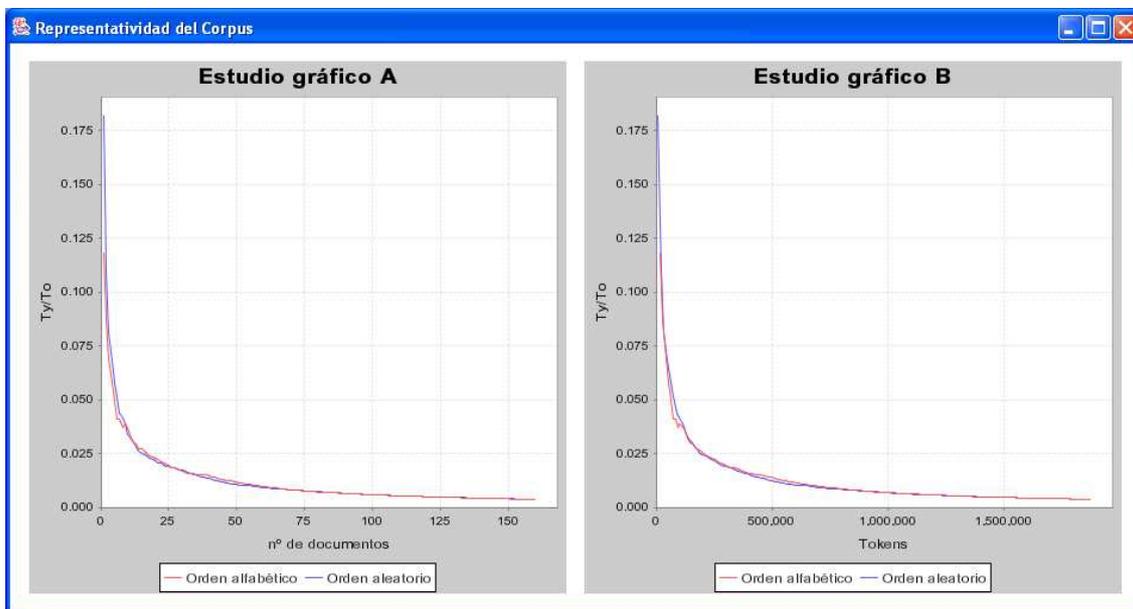


Figura 23. Ejemplos de resultados gráficos de *ReCor* después de analizar un corpus representativo.

El programa ha demostrado su eficacia para lenguajes muy acotados y restringidos, como es el lenguaje jurídico (cf. Seghiri Domínguez, 2006), puesto que se trata de lenguajes con una densidad léxica muy alta y un nivel de repetición de terminología y de estructuras muy elevado. Para lenguajes con mayor variabilidad y flexibilidad, como es el caso del lenguaje del turismo, la eficacia de este programa disminuye, debido al hecho de que el número de documentos (y, por ende, de *tokens*) debe ser extremadamente elevado para conseguir la representatividad. Por tanto, a pesar de demostrar en sus resultados cierta representatividad, los corpus de lenguajes abiertos y flexibles no arrojan resultados tan positivos como otros lenguajes más delimitados y necesitan, por tanto, estar basados en criterios externos más específicos para alcanzar la calidad deseada (Castillo Rodríguez, 2010: 249 y ss.; cf. 3.4.).

Para terminar, otro aspecto importante que se ha de tener en cuenta a la hora de compilar un corpus es el criterio del equilibrio (cf. 1.4.2.2.1.). Con el cumplimiento de este criterio, se obtendrá un corpus formado por textos distribuidos de forma equitativa (con relación al género textual, a la lengua, a los autores, a los

subcampos de la especialidad, etc.) que dará lugar a una representación textual de la realidad del dominio. Como en otras ocasiones, este criterio dependerá de los objetivos planteados y, según estos, el terminógrafo deberá compilar el mismo número de textos (aproximadamente) para las variables determinadas. Por ejemplo, si el proyecto consiste en el análisis de varios subdominios de un mismo campo de especialidad, el corpus deberá estar formado por un número similar de textos pertenecientes a los diferentes dominios.

Por último, llegamos a la etapa de organización de la información encontrada, que presentamos a continuación.

2.6.1.2.3. Organización de textos

Una vez que hemos realizado las búsquedas de textos pertinentes para la compilación de nuestro corpus especializado, debemos organizar la información encontrada mediante su almacenamiento y su etiquetado identificativo.

En primer lugar, es importante guardar los textos seleccionados en un formato adecuado y homogéneo, a fin de preparar los textos para las fases siguientes del trabajo. Para ello, se elige normalmente el formato ASCII o de texto plano para convertir todos los demás formatos que se pueden encontrar en Internet (.pdf, .html, .rtf., etc.) y, así, eliminar las posibles marcas que incorpora el texto. De esta manera, seguimos la *clean-text policy*, expuesta por Sinclair (1991: 21): «The safest policy is to keep the text as it is, unprocessed and clean of any other codes».

A la par que se van convirtiendo los documentos en formato estándar, se deben ir almacenando de forma ordenada e inequívoca. Para ello, se recomienda establecer unas marcas identificativas homogéneas para los textos, así como un documento, por ejemplo un documento Excel, donde se introduzcan las referencias de los documentos creados (autor, URL, dominio, texto original o traducción, etc.). Un ejemplo de etiqueta identificativa podría ser la siguiente: CES001, para un texto en español que forma parte

de un corpus comparable. Esta identificación haría referencia al documento que se ha creado en formato .txt, así como a la entrada pertinente en el documento .xls.⁵⁵⁸

Una vez que hemos compilado y almacenado adecuadamente el corpus especializado, pasamos a la fase de gestión ontoterminográfica, donde se procederá a gestionar el corpus previamente compilado y se indicará el procedimiento que se debe seguir para extraer la información pertinente.

2.6.1.3. Fase III: Gestión ontoterminográfica

Llegados a este punto, ya contamos con la base para comenzar el trabajo ontoterminográfico propiamente dicho, es decir, ya tenemos establecidos todos los objetivos y variables del proyecto, así como el análisis del dominio objeto de estudio (Fase I) y compilado el corpus de trabajo (Fase II). Esta tercera fase de nuestra propuesta se divide a su vez en varias etapas: en primer lugar, se lleva a cabo la extracción de candidatos a término y su consiguiente validación; en segundo lugar, se procede a la búsqueda de equivalentes de traducción de los términos extraídos en la etapa anterior; posteriormente, se realiza la extracción de relaciones semánticas; y por último, se elabora la ontología de dominio, como resultado de todos los procesos elaborados previamente. La división en etapas de esta fase no es estricta, puesto que frecuentemente estas tareas se llevan a cabo de forma simultánea o incluso en diferente orden, según las preferencias del terminógrafo y la herramienta empleada.

Una vez finalizada esta fase, contaremos con una lista de los términos del dominio objeto de estudio, de relaciones semánticas propias de dicho campo de especialidad, así como con los equivalentes de los términos extraídos previamente en las demás lenguas del proyecto. Toda esta información se recogerá en la ontología que se elabore en la última etapa, la cual vinculará los conceptos del dominio con sus

⁵⁵⁸ En la mayoría de los proyectos que realizan estudios con corpus se cuenta con una herramienta que permite la organización del corpus mediante la creación de etiquetas identificativas de forma automática al guardar el documento y rellenar los datos referenciales (autor, procedencia, etc.). De esta manera, se facilita la fase de almacenamiento de documentos. No obstante, no siempre se cuenta con aplicaciones de este tipo y es necesario un almacenamiento manual como este.

respectivos términos en cada lengua y las relaciones semánticas establecidas entre ellos. En nuestro caso, proponemos que la ontología resultante tenga forma de marco semántico (cf. 1.3.2.3.3.1.), a fin de que represente claramente el funcionamiento del dominio en cuestión y su dinamicidad. Este marco semántico estará basado en el esquema categorial prototípico que se había creado en la Fase I y que, después de haber realizado el análisis del corpus y de haber adquirido un conocimiento más profundo del dominio, se podría modificar y fijar de manera más estable.⁵⁵⁹

Antes de comenzar con la descripción de las siguientes etapas, cabe destacar que, según las herramientas con las que se cuente para el proyecto ontoterminográfico, estas etapas se realizarán en una misma aplicación o mediante el uso de varias herramientas independientes (buscador de concordancias, extractor terminológico, editor de ontologías, etc.) (cf. 1.4.2.4. y 1.4.3.4.), puesto que no todas ofrecen las mismas posibilidades.

2.6.1.3.1. Extracción de candidatos a término y validación

La eficacia de la extracción de candidatos a término depende en gran medida de los recursos informáticos con los que cuente el terminógrafo. Así, podemos encontrar proyectos que no dispongan de extractores automáticos de terminología y, por tanto, esta tarea se deba realizar de forma manual; o por el contrario, podemos encontrar otros que sí cuenten con aplicaciones informáticas que faciliten y sistematicen la extracción terminológica. Actualmente, la gran mayoría de proyectos cuentan con extractores de terminología, a fin de agilizar este trabajo y buscar la mayor eficiencia. No obstante, como veíamos en el capítulo primero (cf. 1.4.2.4.), existen diferentes tipos de extractores y no todos ofrecen la misma calidad en el resultado.⁵⁶⁰ De esta manera,

⁵⁵⁹ “De manera estable” no implica que el esquema sea fijo e inamovible, ya que esto sería imposible debido a la dinamicidad de los campos de especialidad, sino que deja de ser provisional para ser más definitivo. No obstante, contando siempre con la posibilidad de modificación y actualización.

⁵⁶⁰ La calidad de un sistema de extracción de términos se suele medir según los parámetros de cobertura (*recall*) y precisión (*precision*), así como según el ruido y el silencio que ofrecen en sus resultados (Cabré *et ál.*, 2001; Zielinski y Ramírez Safar, 2005; Porta Zamorano, 2006, entre otros).

según la calidad que ofrezca el extractor utilizado, los resultados serán mejores o peores y necesitarán de una mayor o menor revisión por parte del terminógrafo.

Comparando las tres técnicas existentes, a saber: lingüística, estadística e híbrida, encontramos que la tercera es la más eficaz y los resultados son mejores, ya que, mediante la combinación de ambas técnicas, reduce el ruido y el silencio y la recuperación de términos es mayor. Sin embargo, a menudo la más utilizada es la estadística, ya que es independiente de la lengua y del dominio de trabajo y, además, hay mayor disponibilidad de aplicaciones informáticas basadas en este enfoque, tanto comerciales (*WordSmith Tools*) como de libre acceso (*TermoStat*).

Como indicábamos anteriormente, el uso de un extractor de términos u otro dependerá de las posibilidades de cada proyecto, por lo que no es posible realizar ninguna propuesta acerca de estas técnicas. No obstante, es necesario tener en cuenta que, independientemente de la herramienta que se utilice, que el resultado que ofrece el extractor es simplemente una lista de candidatos a término y que, posteriormente, el terminógrafo deberá realizar una revisión de los resultados obtenidos y proceder a una validación de los que sí considera términos del dominio objeto de estudio. A pesar de haber adquirido gran conocimiento sobre el campo de especialidad sobre el que versa el proyecto, es recomendable que el terminógrafo cuente con la colaboración de especialistas en la materia para llevar a cabo esta validación. Así pues, proponemos dos fases de validación para los resultados de candidatos a término: una primera fase en la que el terminógrafo lleva a cabo una revisión y su consiguiente eliminación de unidades que pertenecen a la lengua general, posibles unidades gramaticales sin valor (preposiciones, artículos, etc.) que no se han filtrado de forma adecuada, etc.; y, posteriormente, una segunda fase apoyada por la colaboración de los especialistas, donde se discernirá sobre el resto de resultados y se validarán las unidades terminológicas propias del ámbito de especialidad del proyecto.

Con la validación de los candidatos a término, se obtendrá una lista con los términos de ese dominio y se podrá seguir con la gestión ontoterminográfica.

2.6.1.3.2. Búsqueda de equivalentes de traducción

Esta etapa está caracterizada por una mayor complejidad que las anteriores, puesto que la extracción multilingüe y automática de equivalentes no ofrece hasta ahora tan buenos resultados como la anterior y la disponibilidad de aplicaciones informáticas para llevar a cabo esta tarea es más limitada.

La mayoría de las propuestas para la extracción de terminología multilingüe están basadas en la utilización de corpus paralelos, que se alinean como paso previo a su posterior extracción bilingüe. Las alineaciones de estas aplicaciones pueden ser principalmente de párrafos, de frases o de unidades léxicas, y su objetivo principal consiste en la identificación de un candidato equivalente basándose en la identificación previa de un término. Existen diferentes tipos de técnicas, que pueden ir de menor a mayor complejidad. En un nivel bajo, encontramos la técnica manual, que consiste en la identificación manual de equivalentes una vez que se han alineado los textos y se han determinado los términos del dominio en la lengua origen. Y en los niveles superiores, observamos de nuevo la presencia de tres tipos de técnicas diferentes (Martínez Unanue, 1999):

- Enfoque estadístico. Esta técnica se basa en la similitud de algunos rasgos cuantitativos identificados en el corpus, como la longitud de oraciones, el número de palabras o de caracteres, etc. Presenta una gran desventaja y es el hecho de que si la traducción no mantiene una alta semejanza estructural con el texto original, el resultado de la alineación no será satisfactorio.
- Enfoque lingüístico. Aquí la alineación se basa en el emparejamiento previo de unidades sintagmáticas o de estructuras con cierta dependencia y necesita un marcado morfosintáctico previo. El inconveniente de esta técnica consiste en la dependencia que sostiene con la lengua y, además, la necesidad de un preprocesamiento del texto con un anotador morfosintáctico.
- Enfoque mixto. Este método aprovecha la identificación de categorías gramaticales como apoyo para la alineación estadística. En este caso, se sigue

presentando la desventaja acerca de la dependencia de la lengua, pero se pretende evitar el inconveniente del método estadístico.

Como afirma Rapp (1995: 320), los resultados obtenidos por estas técnicas han sido muy útiles para la comprobación de la coherencia terminológica en traducciones y como recurso para el análisis de traducciones, entre otras cosas; sin embargo, la disponibilidad de textos para la compilación de este tipo de corpus para un dominio concreto y para un par de lenguas determinado es a menudo muy limitada y larga. Por este motivo, este autor aboga por el uso de corpus comparables, que presentan una mayor facilidad a la hora de su compilación, para extraer terminología bilingüe.

Since the acquisition of non-parallel texts is usually much easier, it would be desirable to have a program that can determine the translations of words from comparable or even unrelated texts.

A este respecto, también es necesario añadir que los corpus comparables están formados por textos originales y, por tanto, más fiables en cuanto a contenido y forma, ya que no siempre las traducciones han sido elaboradas por traductores profesionales o revisadas por expertos y, a veces, presentan errores de diferente naturaleza (léxicos, semánticos, ortográficos, etc.) (cf. 3.3.5.1.).

En este contexto, existe actualmente una tendencia a la utilización de corpus comparables para la extracción de equivalentes de traducción, aunque sus resultados son todavía bastante tímidos. Entre estas propuestas, cabe destacar aquellas que utilizan la identificación de cognados para el establecimiento de equivalencias (Simar et ál., 1992; McTait, 2001; Mitkov et ál., 2008); las que utilizan pequeños corpus paralelos para extraer «anclas significativas» («*meaningful anchors*») (Gamallo Otero, 2007; Gamallo Otero y Pichel Campos, 2005); las que combinan el uso de diccionarios bilingües externos y cognados (Gamallo Otero y Pichel Campos, 2007); y las que identifican contextos o patrones combinatorios similares en la lengua origen y en la lengua meta (Marinai et ál., 1991; Fung, 1995; Rapp, 1995; Tognini-Bonelli, 2001; Teubert, 2004; Teubert y Čermáková, 2004; Sharoff, 2004, 2006).⁵⁶¹ En definitiva, se dispone actualmente de diferentes aproximaciones para la extracción de equivalentes a partir de

⁵⁶¹ Esta última metodología se trata con profundidad en Corpas Pastor (2008: 57 y ss.).

corpus comparables y, en cada caso y según los recursos informáticos y humanos, se deberá elegir una opción u otra.

Desde nuestro punto de vista, la mejor opción para llevar a cabo la extracción de equivalencias consiste en una combinación de ambas, es decir, en la utilización de un corpus paralelo y un corpus comparable, a fin de valerse de las ventajas que una y otra opción ofrecen, a saber: por un lado, los corpus paralelos son una fuente de equivalencias de traducción que permiten la extracción de estas de forma rápida y sencilla; y por otro, los corpus comparables ofrecen la posibilidad de observar los términos en su contexto real y original y exentos de ambigüedades y errores semánticos. Asimismo, el uso de los marcos semánticos, que proponemos en esta metodología, posibilita el estudio de los contextos similares en el corpus comparable. Por su parte, con la combinación de estos dos corpus reducimos los inconvenientes de ambos: por un lado, se reduce la problemática surgida por la dificultad que se presenta a la hora de elaborar un corpus paralelo de grandes dimensiones para llevar a cabo un trabajo ontoterminográfico, ya que sería suficiente con un corpus de pequeñas dimensiones; y por otro, la desventaja que tienen los corpus comparables a la hora de medir la cobertura de sus resultados se reduciría gracias a la obtención de información a través del corpus paralelo y de la creación de marcos semánticos.

En definitiva, cada terminógrafo deberá decidir en esta etapa la metodología que emplea (según los recursos disponibles en cada caso), pero en nuestra propuesta metodológica defendemos la combinación de dos corpus (paralelo y comparable) para la extracción de equivalentes, así como del resto de información lingüística, pragmática y semántica necesaria.

2.6.1.3.3. Extracción de relaciones semánticas

Un aspecto imprescindible para la elaboración de la ontología de dominio y, por tanto, la base de datos ontoterminográfica, consiste en la detección de las relaciones semánticas propias del dominio objeto de estudio, las cuales permiten obtener

información, tanto genérica como específica, sobre la organización del campo de especialidad y poder, así, establecer una representación de este fiel a la realidad.

Como veíamos en el primer capítulo (cf. 1.4.2.4.), el primer paso para extraer las relaciones de un dominio determinado a partir de un corpus textual requiere la extracción de los términos de dicho dominio. Una vez que hemos realizado esta tarea en la fase anterior y contamos con el corpus especializado compilado, podemos comenzar la búsqueda de las relaciones semánticas que vinculan a estos términos entre sí. De esta forma, mediante un enfoque *bottom-up*, se obtienen las relaciones semánticas propias de un dominio a partir del análisis de corpus textuales y de los términos extraídos de estos.

De los posibles métodos existentes para la adquisición de conocimiento a partir de corpus, el más utilizado actualmente por los investigadores se basa en la extracción de patrones de conocimiento (*knowledge patterns*),⁵⁶² del cual fue pionero Hearst (1992).⁵⁶³ Según este autor, estos patrones de conocimiento son «easily recognisable, occur frequently and across text genre boundaries, and indisputably indicate the lexical relation of interest». A partir del trabajo de este autor, ha surgido gran interés dentro de la comunidad científica y ha dado lugar desde entonces a numerosos estudios. Como ejemplo, podemos citar a los que siguen: Meyer et ál. (1992), Auger (1997), Marshman et ál. (2002), Barrière (2001), Condamines y Rebeyrolle (2001); Meyer (2001), Feliu (2004), Malaisé et ál. (2005), Pantel y Pennacchiotti (2006), Marshman y L'Homme (2006) o Soler Puertes y Alcina Caudet (2008). Según esta aproximación, los patrones de conocimiento permiten al terminógrafo detectar, de una manera rápida y eficaz, las relaciones semánticas y las estructuras lingüísticas que le ayudarán en su trabajo.

La relación más importante y, por tanto, la más estudiada en este ámbito es la relación jerárquica lógica, es decir, la relación de hiperonimia (IS-A). Esta relación se ha considerado generalmente como la fundamental para el desarrollo de estructuras de

⁵⁶² Como veíamos en el capítulo anterior, los patrones de conocimiento son patrones lingüísticos que se repiten y que expresan conocimiento acerca de las relaciones existentes entre los términos del campo de especialidad objeto de estudio.

⁵⁶³ Además de este método basado en los patrones de conocimiento, también se encuentran aproximaciones de tipo estadístico (cf. Agirre et ál., 2000; cf. Faatz y Buitelaar, 2002) y de búsqueda automática basada en el aprendizaje por inferencia (cf. Cimiano, 2006), aunque el más extendido es el basado en los patrones.

conocimiento, puesto que permite la construcción de jerarquías basadas en los conceptos generales del dominio y sus subordinados. Asimismo, se trata de la parte nuclear del tipo de definición aristotélica (*definiens*), lo que hace que su identificación sea extremadamente importante para los terminógrafos.

Aunque no se ha considerado al mismo nivel de importancia que la hiperonimia, la relación jerárquica ontológica, es decir, la relación de metonimia (parte-todo) también ocupa un lugar preferente en la clasificación y estructuración de algunos conceptos. Esta relación, a su vez, ha sido dividida en diferentes subtipos (Winston et ál., 1987), a saber:

- Componente-objeto integral. La parte posee una función con respecto al todo y del que puede desconectarse físicamente. Por ejemplo, los pares de conceptos *asa-taza* y *rueda-coche*.
- Miembro-colección. Las partes se encuentran interconectadas mediante una proximidad espacial o una relación social. Es el caso de *árbol-bosque* o *libro-biblioteca*.
- Porción-masa. Las partes son similares entre sí y con el todo que forman y, además, pueden separarse de él de forma arbitraria. Por ejemplo, los *trozos* de una *tarta*.
- Materia-objeto. La parte se entiende como el ingrediente que compone al todo y del que no puede separarse físicamente sin alterar su identidad, como es el caso de *crystal-ventana* o *oxígeno-agua*.
- Característica-actividad. Las partes reflejan las fases, características o etapas que tienen lugar en un todo o proceso. Por ejemplo, *pagar-comprar* o *patinaje-patinar*.
- Lugar-área. Las partes carecen de una función específica con respecto al todo, pero no se pueden separar de él, como ocurre con *oasis-desierto*.

Por último, encontramos las relaciones no jerárquicas (o asociativas), que han sido ignoradas prácticamente por la terminología tradicional a favor de las relaciones lógicas y ontológicas, pero que en la terminología moderna han adquirido gran importancia. Estas relaciones pueden ser de diferentes tipos: entidad-función, causa-efecto,

productor-producto, instrumento-uso, etc. (cf. Sager, 1990). Además de las relaciones generales que recoge este autor, cada proyecto ontoterminográfico podrá crear sus propias relaciones *ad hoc*, a fin de representar el dominio objeto de estudio y adaptarlas, por tanto, a sus necesidades.⁵⁶⁴

Con la combinación de los tipos de relaciones semánticas, las jerárquicas y no jerárquicas, a la hora de representar el conocimiento de un campo de especialidad, se obtendrá como resultado una representación formada por relaciones verticales (jerárquicas) y horizontales (no jerárquicas) (León Aráuz et ál., 2009: 51).

La detección y extracción de las relaciones semánticas basadas en patrones lingüísticos se divide principalmente en tres pasos: a) definición de las relaciones semánticas de interés; b) descubrimiento de los patrones que expresan tales relaciones en el texto, y c) la búsqueda de instancias para las relaciones mediante el uso de patrones (Auger y Barrière, 2008: 3-7).⁵⁶⁵

(a) Definición de las relaciones semánticas de interés

El terminógrafo debe determinar cuáles son las relaciones que desea analizar y, por tanto, detectarlas en el corpus de trabajo según los objetivos planteados inicialmente para el proyecto. Así, debe decidir si desea extraer únicamente relaciones de hiperonimia, de metonimia, no jerárquicas, etc. o una combinación de todas. Para llevar a cabo esta tarea, el terminógrafo puede partir de un estudio inicial *in vitro* de los términos y las definiciones incluidos en los recursos terminológicos que haya seleccionado durante la fase de análisis del dominio, a fin de observar la organización conceptual a través de entradas terminológicas (Dubuc y Lauriston, 1997: 87).

Este estudio consiste en la extracción de las definiciones pertinentes que se encuentran en los diferentes recursos terminológicos seleccionados para los términos

⁵⁶⁴ Para profundizar en los diferentes estudios que se han llevado a cabo basados en tipos específicos de estas relaciones, recomendamos la consulta de Montiel Ponsoda (2011: 51-61), que realiza una revisión de los estudios más recientes, y, por supuesto, la consulta de los estudios reseñados en dicha obra.

⁵⁶⁵ Estos autores incluyen en la metodología un cuarto paso, a saber: la estructuración de las nuevas instancias como parte de una ontología nueva o existente. En nuestro caso, consideramos este paso como parte fundamental de la gestión ontoterminográfica (Fase III de nuestra metodología) y como resultado final de las tareas realizadas con anterioridad y, por tanto, la consideraremos como un apartado independiente.

extraídos del corpus, que forman parte de la nomenclatura del producto que se está elaborando, y, a continuación, realizar un análisis de la red léxico-conceptual que contienen. De esta manera, el terminógrafo puede visualizar claramente las relaciones semánticas que establecen los términos con otros conceptos y, así, decidir las relaciones que pretende incluir en su obra mediante el estudio de los componentes de la definición (*genus + differentiae*) (cf. 2.6.1.4.2.).

(b) Descubrimiento de los patrones que expresan tales relaciones en el texto

Tradicionalmente, se utilizaban únicamente los diccionarios existentes para el estudio de las relaciones semánticas a partir de las definiciones de sus entradas, pero actualmente se prefiere utilizar los corpus textuales para recuperar y describir empíricamente las relaciones semánticas, debido a las ventajas que presenta frente a los diccionarios (comunicación real, información contextual, mayor variedad, etc.). No obstante, como veíamos en el paso anterior, se siguen utilizando los recursos terminológicos como base para detectar relaciones y patrones semánticos.

Una de las estrategias más empleadas por la metodología basada en patrones lingüísticos para la extracción de relaciones semánticas de corpus textuales consiste en la aplicación de los patrones más recurrentes de cada tipo de relación identificados previamente en otros estudios y trabajo, como por ejemplo los siguientes:

- Hiperonimia: es un tipo de, es un, incluye...
- Metonimia: es una parte de, constituido por, contiene, compuesto de, consta de...
- Causal (causa-efecto): causa, resulta en, provoca, es el resultado de...
- etc.

Una vez establecidos, se adaptan en caso de que sea necesario y se utilizan para encontrar instancias en textos. Estos pasos se realizan de manera cíclica, a fin de mejorar gradualmente la cobertura de las ontologías y la detección de patrones, lo que dota a esta metodología de un carácter dinámico.

c) Búsqueda de instancias para las relaciones mediante el uso de patrones

En su forma más básica, las relaciones semánticas basadas en patrones lingüísticos incluyen un término X, un término Y y una unidad lingüística que expresa una relación semántica entre los términos X e Y. Con objeto de detectar los patrones léxicos⁵⁶⁶ de cada dominio, se pueden emplear dos tipos de metodología principalmente: por un lado, una metodología con una aproximación orientada al término, es decir, basada en la búsqueda de relaciones semánticas a través de oraciones que contienen los términos propios del dominio (Meyer, 2001; Pearson, 1998; Condamines y Rebeyrole, 2001); y por otro, una metodología que defiende la búsqueda de relaciones semánticas en todo el corpus a través de patrones léxicos preestablecidos y sin utilizar los términos como impulsores para la detección de estas relaciones, con objeto de descubrir posibles conceptos nuevos no expresados por términos pero sean de utilidad para la organización del dominio (Barrière, 2004).

En nuestra propuesta de metodología, defendemos una combinación de las dos, ya que el hecho de limitarse a las búsquedas de los patrones vinculados a los términos propios del dominio puede que reduzca las posibilidades del terminógrafo a la hora de profundizar en la estructuración de dicho dominio. No obstante, la segunda se empleará como complementaria a la primera y no al contrario, ya que el objetivo de esta tarea consiste en el descubrimiento de las relaciones semánticas contenidas en el corpus.

Una vez extraídas las relaciones semánticas del campo de especialidad, pasaremos a la siguiente etapa, la elaboración de la ontología de dominio, que consiste en la combinación de las tareas anteriores de esta fase, a saber: la extracción terminológica, la extracción de equivalentes y la extracción de relaciones semánticas.

⁵⁶⁶ Recordemos que se distinguen tres tipos de patrones lingüísticos, a saber: los patrones léxicos, los patrones gramaticales y los patrones paralingüísticos. Los primeros consisten en palabras o grupos de palabras que indican las relaciones conceptuales; los segundos implican combinaciones de estructuras gramaticales que ofrecen relaciones semánticas entre conceptos aunque de un modo más limitado que los patrones léxicos; mientras que los terceros hacen referencia a los elementos textuales como por ejemplo comas, paréntesis, puntos, etc.

2.6.1.3.4. Elaboración de la ontología de dominio

Llegados a esta última etapa de la fase de gestión ontoterminográfica, procederemos a la representación completa del dominio de especialidad a partir de sus conceptos propios y las relaciones semánticas extraídas. Asimismo, podremos realizar la concreción del esquema prototípico del dominio, que se estableció en la primera fase con el análisis del campo de especialidad, en caso de que lo consideremos necesario, es decir, en caso de que durante el proceso de trabajo realizado hasta el momento se hayan detectado carencias, errores o se pueda concretar aún más el marco semántico elaborado. Una vez modificado o ampliado, nos basaremos en él para llevar a cabo la organización del dominio a nivel conceptual. Cabe recordar que no existe un único modo correcto de modelar un dominio, por lo que la estructuración que se realice en esta fase será únicamente una opción posible, resultado de los objetivos propuestos en proyecto determinado.

Esta representación conceptual, es decir, la ontología, se puede llevar a cabo de diversas formas, según la aproximación y las posibilidades que ofrezca el editor que se elija y los conocimientos en ingeniería del conocimiento de los que disponga el terminógrafo. Los principales tipos serían los siguientes, a saber: de forma lineal, es decir, con una ordenación jerárquica de los conceptos (cf. Figura 24); mediante grafos, es decir, mediante una representación gráfica formada por grafos que incluyen los conceptos correspondientes (cf. Figura 25); o mediante marcos semánticos, esto es, con el empleo de esquemas en forma de marcos que relacionan los conceptos que pueden aparecer juntos en una misma situación comunicativa (cf. Figura 25). En este sentido, cabe destacar que, sea cual sea la opción elegida, todas las ontologías estarán formadas por los conceptos propios del dominio objeto de estudio y las relaciones que los vinculan entre sí y que ninguna debería considerarse mejor que la otra, sino como más o menos recomendable según los objetivos del proyecto, según las posibilidades y según la aproximación que se realice.

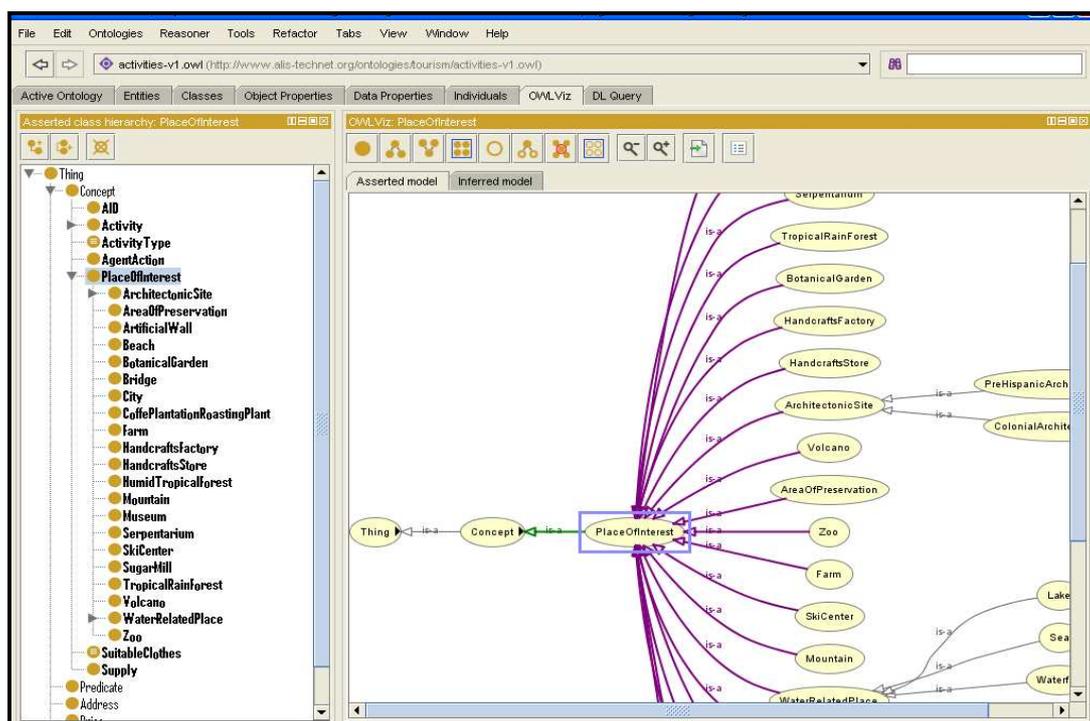


Figura 24. Representación lineal (a la derecha) y en grafos (a la izquierda) con el programa *Protégé*.

Como enunciábamos anteriormente, nuestra metodología está basada en la Semántica de Marcos aplicada a la terminología, ya que permite una estructuración clara y ordenada de los conceptos a la par que posibilita la observación del contexto en el que aparecen los diferentes conceptos (y sus términos). Asimismo, esta forma de estructuración conceptual se asemeja a la representación de la realidad que realizamos en nuestra mente según la neurociencia cognitiva (cf. Givón, 1995), ya que en esta representación se establecen relaciones entre los conceptos que aparecen normalmente en la misma situación comunicativa, al igual que ocurre con la estructuración conceptual de nuestra mente. Por ejemplo, en el caso de la situación comunicativa de «comprar un producto», normalmente asociamos en nuestra mente una serie de conceptos a dicha situación como «vendedor», «comprador», «vender», «comprar», «producto», «dinero», etc., los cuales estarían recogidos en un marco semántico de la misma manera que organizamos dicho conocimiento en nuestra mente. Por ello, consideramos la representación conceptual mediante marcos semánticos como la forma más adecuada para llevar a cabo la elaboración de la ontología, aunque, como complemento, se puedan utilizar otro tipo.

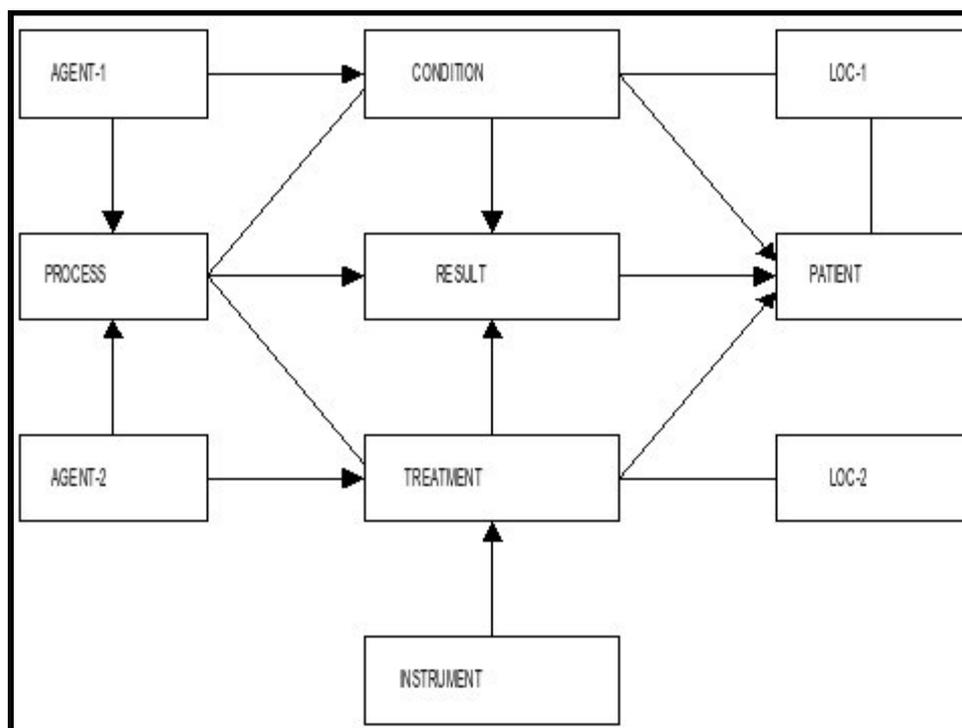


Figura 25. Representación del esquema categorial prototípico en el que se basan los marcos semánticos específicos. Ejemplo extraído del proyecto de investigación ONCOTERM (Faber Benítez, 2002: 7)

En el ámbito ontoterminográfico, la elaboración de ontologías debe cumplir una serie de requisitos generales y específicos, a fin de que el resultado sea adecuado y fiable. Como comentamos en el primer capítulo (cf .1.4.3.), los principales criterios específicos dentro del marco ontoterminográfico son los siguientes: a) flexibilidad a la hora de representar el conocimiento de cualquier dominio, con objeto de adaptarse a su dinamicidad y multidimensionalidad, así como a los objetivos del proyecto; b) la complejidad necesaria según los objetivos planteados para el proyecto; c) la utilización de las relaciones jerárquicas y la posibilidad de creación de relaciones no jerárquicas *ad hoc* para la representación del dominio; d) la facilidad de manejo para el terminógrafo, y e) la compatibilidad de formato con otros sistemas y aplicaciones. Así pues, la herramienta que se elija para llevar a cabo la elaboración de la ontología deberá cumplir los requisitos anteriores, a fin de que el terminógrafo pueda realizar su trabajo de forma apropiada.

Además de la representación del conocimiento del dominio objeto de estudio, la ontología servirá como base para la elaboración posterior de las definiciones, que se consideran como «*mini-knowledge representations or frames*» (León Aráuz et ál., 2009: 51).

Actualmente, se dispone de diferentes metodologías generales para la elaboración de ontologías, con diverso grado de dificultad y especificación en su aplicación. Entre ellas, cabe destacar METHONTOLOGY, Cyc, On-to-Knowledge, Uschold y King, Grüninger y Fox, Kactus, entre otras.⁵⁶⁷ Todas estas metodologías son propuestas que se incluyen dentro de la Ingeniería del Conocimiento, por lo que sus objetivos no son completamente compatibles con los que tenemos en la ontoterminografía y, además, el nivel de complejidad para la elaboración de ontologías mediante estas metodologías es bastante elevado. Con objeto de facilitar la labor del terminógrafo, se han creado diferentes herramientas para la elaboración de ontologías dentro de un marco ontoterminográfico, como son *Termontology Tools*, *Ontoterm*, *Spaterm*, *Protermino*, etc.,⁵⁶⁸ que guían la labor ontoterminográfica y facilitan en gran medida la elaboración de la ontología de dominio.

No obstante, cada proyecto ontoterminográfico dependerá de las posibilidades de utilización de herramientas propias para la realización de ontoterminografía u otras herramientas más específicas para la creación de ontologías, con lo cual sería recomendable que contara con un apoyo informático importante.

En nuestra propuesta metodológica, procederemos de la siguiente manera:

Una vez establecido el marco semántico prototípico del campo de especialidad en la fase inicial, que estará formado por las categorías y relaciones semánticas básicas detectadas después de una aproximación inicial al dominio en cuestión, se procederá a revisar dicha organización y a llevar a cabo las modificaciones y ampliaciones que sean pertinentes en cada caso. De esta manera, se podrán incluir, eliminar o mover

⁵⁶⁷ En Hernández Ramírez y Saiz Noeda (2007) se puede consultar una revisión de estas y otras metodologías de desarrollo de ontologías.

⁵⁶⁸ Estas herramientas son sistemas de gestión ontoterminográfica y serán tratadas en profundidad en el siguiente capítulo (cf. 3.2.1.)

relaciones, así como especificar las posibles subcategorías que se encuentran dentro de las categorías principales determinadas al comienzo del proyecto o, en su caso, modificarlas, siempre que se considere oportuno. Con posterioridad, se utilizará dicho esquema inicial (ya sea en forma lineal, de grafo o de marco semántico) como base para elaborar la ontología del dominio, es decir, para concretar dicho esquema categorial con los conceptos propios del dominio de trabajo en cuestión. En este paso, el terminógrafo deberá determinar los conceptos propios del campo de especialidad e ir vinculándolos a las categorías principales y subcategorías que forman el esquema inicial con objeto de ir modelando y organizando el conocimiento de dicho dominio a través del nivel conceptual. Dentro de dichas categorías, los conceptos se organizarán, a su vez, a través de relaciones jerárquicas y no jerárquicas. Posteriormente, el terminógrafo procederá a vincular los diferentes términos encontrados dentro de corpus de trabajo a través de la extracción terminológica (ya sea manual, semiautomática o automática) con los diferentes conceptos de este esquema prototípico, es decir, pasará del nivel puramente conceptual (independiente de la lengua) a un nivel lingüístico (dependiente de cada lengua). Durante este proceso, se deberá presentar especial atención a las posibles diferentes variantes de un mismo concepto (variantes geográficas, ortográficas, denominativas, etc.) que se vayan encontrando y que deberán tenerse en cuenta para la posterior elaboración de la base de datos ontoterminográfica. Una vez que se haya realizado esta vinculación entre el nivel conceptual y el nivel lingüístico de una lengua determinada y se hayan corregido los posibles vacíos o errores encontrados, se pasará a la realización del mismo proceso en las demás lenguas de trabajo. De esta manera, se irán observando las posibles diferencias entre el esquema categorial, independiente de la lengua, y las posibles diferencias entre las distintas lenguas, lo que también se apuntará para la siguiente fase.

Así pues, se consigue la simplificación de la elaboración de ontologías de dominio para el trabajo ontoterminográfico, haciéndolo más accesible para los que carecen de conocimientos de la Ingeniería del Conocimiento, y se ha adaptado a las necesidades que requiere un trabajo ontoterminográfico para conseguir sus resultados. Asimismo, permite la consecución de los objetivos de las ontologías en este marco de trabajo, a saber:

- Estructurar el conocimiento especializado de forma clara y concisa mediante la representación de conceptos y las relaciones que los vinculan entre sí;
- Detectar las diferencias culturales y lingüísticas de las lenguas de trabajo; y
- Servir de base para la elaboración de definiciones en una fase posterior.

2.6.1.4. Fase IV: Elaboración de una base de datos ontoterminográfica

Durante décadas, traductores, lexicógrafos y terminólogos han almacenado la información terminológica en fichas de papel (las denominadas «fichas de vaciado» o «fichas terminológicas»),⁵⁶⁹ lo que daba lugar a un trabajo manual que a menudo era largo y tedioso. Actualmente, gracias a los avances de las aplicaciones informáticas y de su difusión, esta situación ha cambiado y, hoy en día, la forma más frecuente de almacenamiento de la información terminológica se lleva a cabo en formato electrónico, ya sea mediante un procesador de texto, un programa de creación de bases de datos genérico o un sistema gestor de bases de datos terminográficas u ontoterminográficas, lo que se conocen como sistemas de gestión terminológica (SGT). Estos últimos son los más recomendados a la hora de realizar cualquier trabajo terminográfico, ya que permiten introducir, manipular, relacionar y extraer datos de manera rápida, sencilla y ordenada.

Dentro de esta fase, son dos las decisiones generales que se han de tomar, a saber: 1) qué aplicación se va a utilizar para llevar a cabo esta tarea y, por tanto, crear y gestionar los registros de entrada y 2) qué campos va a contener cada una entrada, según los objetivos de la entrada y, concretamente, según el grupo de destinatarios.

Para la primera cuestión, será necesario evaluar las diferentes herramientas existentes, a fin de seleccionar la más adecuada para alcanzar los objetivos del proyecto. De esta manera, según estos objetivos y las posibilidades económicas, informáticas y humanas de cada proyecto, será necesario utilizar una u otra aplicación. El grupo

⁵⁶⁹ En Casas Gómez (2006), se realiza una revisión de diferentes propuestas de fichas de vaciado. Véase también Cabré Castellví (1993: 308 y ss.).

EAGLES (1995) destaca los siguientes criterios para realizar la evaluación de los diferentes sistemas, a fin de seleccionar el más adecuado en cada situación y según los objetivos planteados:

1. Descripción técnica: el hardware y software disponibles en el proyecto para que la herramienta funcione correctamente.
2. Compatibilidad con trabajos y estudios anteriores y con las posibles actualizaciones, así como posibilidad de impresión (ya sea con la selección de todos los datos o solo de algunos), de importación/exportación en formatos estándar (por ejemplo, TMX) y de interacción con otras aplicaciones (procesadores de textos, otras bases de datos y aplicaciones, etc.).
3. Interfaz de usuario y características del sistema: lenguas en las que se visualiza, tipo de ayuda a disposición del usuario, modo de representación de la información en pantalla, posibilidades de manipulación de los menús, iconos, botones, de la visualización de la información en pantalla, etc.
4. Aspectos terminológicos, en relación con (a) la gestión de la terminología: lenguas que admite, tipo de datos que se pueden introducir (textuales, gráficos, multimedia), posibilidad de cambiar la dirección lingüística, etc.; y (b) el modelo y estructura de la entrada: si cumple los estándares, si se trata de una estructura de ficha fija o modificable (añadir o eliminar campos, cambiar el nombre, etc.), si admite categorías de datos administrativos, el número máximo de caracteres en los campos, posibilidad de crear referencias cruzadas entre entradas, número total posible de registros para cada base de datos, etc.
5. Consulta de información: posibilidades de consulta, respuestas del sistema ante consultas, etc.
6. Revisión y corrección: validación y control ortográfico, aviso de entradas duplicadas o de la omisión de un campo obligatorio, etc.
7. Operaciones necesarias para su mantenimiento.

8. Aspectos comerciales: fabricante, distribuidor, precio, disponibilidad y servicio técnico en el territorio nacional, personas/instituciones conocidas que lo utilicen, etc.

Principalmente, estos serían los criterios de evaluación más significativos para ayudar a elegir el sistema de gestión más apropiado según las necesidades y posibilidades del proyecto.

Con respecto a la segunda cuestión, es decir, la selección de los campos terminológicos para cada entrada, la trataremos en el siguiente apartado de forma independiente para profundizar en este aspecto.

2.6.1.4.1. Selección de campos

Tradicionalmente en terminografía se ha realizado una clasificación de los diferentes datos que se pueden encontrar en cualquier sistema de gestión, a saber: información administrativa, lingüística, pragmática y conceptual. En esta línea, Sager (1990: 145) establece siete grandes bloques de categorías de datos:

- Entrada terminológica, que indica la unidad terminológica a la que nos referimos.
- Información sobre las fuentes, que incluye las referencias de las fuentes de las que se han extraído la definición, el término, el contexto u otras informaciones asociadas.
- Especificaciones conceptuales del término, que se componen normalmente de la pertenencia a un campo de especialidad determinado, definiciones, sinónimos y enlaces a otros conceptos expresados como relaciones terminológicas.
- Especificaciones lingüísticas de término, que hacen referencia a las indicaciones gramaticales (género, número, categoría), forma abreviada, etc.
- Especificaciones pragmáticas, que indican los usos de los términos y pueden expresarse mediante ejemplos contextuales y notas de uso.

- Información de gestión o administración, que suele incluir el nombre del terminógrafo responsable y las fecha de creación o modificación del registro.
- Especificaciones sobre equivalencia en otras lenguas, en el caso de las bases de datos multilingües.

La mayoría de los sistemas de gestión contemplan la inclusión de estos siete bloques de categorías de datos; sin embargo, a menudo se encuentran diferencias en la definición y estructuración de las categorías que componen cada bloque, lo que provoca que el resultado final difiera en cuanto a su estructura, además de que obstaculiza, o imposibilita el intercambio, la reutilización y fusión de recursos terminográficos entre instituciones, centros de investigación, empresas, proyectos, etc. Por este motivo, se han llevado a cabo iniciativas desde organizaciones internacionales con el objetivo de desarrollar estándares que permitan el intercambio y homogeneización de las bases de datos terminológicas. Una de estas iniciativas ha sido el trabajo desarrollado por el comité técnico 37 de la ISO, que después de varios años de documentación y consultas extensivas, presentaron en 1999 un amplio listado de las categorías de datos usadas con mayor frecuencia en la gestión terminográfica. Estas categorías de datos forman la base de la norma ISO 12620 (1999) y se dividen en diez secciones diferentes agrupadas en cuatro categorías:

ISO 12620: GRUPOS DE CATEGORÍAS DE DATOS

- Categoría de término

Sección 1: campo que hace referencia a la unidad terminológica de un dominio especializado, esté formada por una única unidad o por varias.

- Categoría de información relacionada con el término

Sección 2: campos que contienen información sobre el término (sinónimos, nombre común, nombre científico, forma completa, abreviatura, variante, fórmula, símbolo, pronunciación, etimología, silabificación, grado de normalización, etc.).

Sección 3: campos de datos sobre la equivalencia entre términos asignados al mismo concepto o a conceptos similares (falsos cognados, grado de equivalencia entre conceptos cercanos, etc.).

- Categoría de datos descriptivos

Sección 4: campos de datos que clasifican los conceptos en áreas o subáreas temáticas.

Sección 5: campos de datos que sirven para describir el concepto (definición, contexto, ejemplos, imágenes, audio, etc.).

Sección 6: campos de datos que indican las relaciones semánticas entre dos conceptos.

Sección 7: campos de datos que se emplean para indicar la posición de los conceptos en un sistema conceptual.

Sección 8: campo «Nota», donde se puede indicar cualquier información complementaria para explicar cualquier aspecto relevante. Esta categoría puede aparecer en cualquier sección de la BDT, si se considera necesario realizar algún apunte.

- Categorías de datos administrativos

Sección 9: campos de datos para especificar los lenguajes documentales y códigos de tesauros utilizados.

Sección 10: campos de datos relativos a la gestión administrativa (autores, fecha de creación y de modificación, responsable, etc.).

Tabla 12. Categorías y secciones de los campos recogidos en la norma ISO 12620: 1999 ordenados según la *CLS Framework*.⁵⁷⁰

Las categorías descritas en esta norma no deben considerarse exhaustivas ni obligatoria su completa inclusión en cualquier trabajo terminográfico que se realice

⁵⁷⁰ El *CLS Framework* (*Concept-oriented with Links and Shared references Framework*) es el resultado de la unión del grupo de investigación de traducción de la Universidad Brigham Young (BYU TRG) y el Instituto de Lingüística Aplicada de la Universidad de Kent (KSU IAL), cuyo principal objetivo consiste en la propuesta de estructura y contenido de las bases de datos terminográficas. La metodología propuesta se puede emplear tanto para la representación de los sistemas de gestión existentes, para el diseño de nuevos sistemas y para compartir información terminológica. Para obtener más información acerca del *CLS Framework*, consúltese su página web. URL: <<http://www.ttt.org>>.

(Wright, 2001: 567). De hecho, cada proyecto deberá seleccionar las categorías que necesite según sus objetivos y sus usuarios y establecer, así, su propio inventario de categorías.

Asimismo, además de seleccionar los campos que se incluirán en cada registro terminológico, el terminógrafo deberá determinar el orden que estos presentarán en el sistema. Debido a la carencia de estas especificaciones en la norma ISO, el *CLS Framework* elaboró una propuesta para estructurar y homogeneizar el orden de aparición de los campos de las diferentes bases de datos, que se puede consultar en su página web. Según esta propuesta, cualquier sistema de gestión se debe estructurar en tres partes diferenciadas y ordenadas de la siguiente manera:

1. Información global sobre la base de datos. En esta parte se especifica la información general acerca de la base de datos, como son el nombre, la versión, los derechos de autor, los creadores, las lenguas, etc.
2. Entradas terminológico-conceptuales. Esta parte está compuesta por los conceptos en un campo de especialidad y los términos que les son asignados como designación específica en una o varias lenguas. A cada uno de estos términos se le asigna, a su vez, una serie de informaciones específicas (descriptivas, administrativas, lingüísticas, de uso, etc.).
3. Grupo de referencias que pueden compartir varias entradas o partes de una entrada. Esta parte se compone de una serie de enlaces que relacionan unas entradas (o elementos de una entrada) con otras entradas o elementos de otras entradas, o bien con un conjunto de referencias que se consideran compartidas (referencias bibliográficas, gráficos, tablas, imágenes y archivos de sonido, etc.).

En definitiva, mediante la selección de los campos propuestos por la norma ISO y siguiendo la propuesta de estructuración del *CLS Framework* se conseguiría una homogeneización de las bases de datos terminológicas, tanto en lo referente a su estructura como a su contenido, por lo que facilitaría su reutilización posterior para proyectos diferentes y su intercambio entre centros, empresas, aplicaciones, etc., lo cual

beneficiaría tanto a los terminógrafos como a los usuarios y propietarios de este tipo de sistemas de gestión, aplicaciones, etc.

Una vez presentados los campos normalizados que recoge la norma ISO y la estructura recomendada del *CLS Framework* solo falta tener en cuenta a los destinatarios del producto en el que se está trabajando, a fin de seleccionar los campos que se incluirán en el producto final. Para ello, es imprescindible tener en cuenta sus necesidades y expectativas, así como las diferentes carencias que presentan otros recursos del mismo campo de especialidad. En nuestra propuesta metodológica, planteamos la necesidad de realizar un estudio específico de esas necesidades y expectativas de los usuarios potenciales, bien a través de estudios realizados directamente a ellos mediante encuestas, entrevistas personales, etc. (cf. 2.5.) o bien mediante la consulta de estudios similares realizados con anterioridad. De esta manera, se consigue una visión general acerca de cómo debería ser el contenido del recurso para ser utilizado por sus destinatarios de forma satisfactoria y se pueden seleccionar los diferentes campos de las entradas según los requisitos de los usuarios finales.

En definitiva, según las variables pragmático-lingüísticas, especialmente según el grupo potencial de destinatarios, será necesario incluir una información u otra, por lo que es imprescindible conocer sus necesidades antes de finalizar el recurso.

Recommending a single model for entry layout is not an easy task, because there are as many different models as there are working terminologists, and these various models reflect diverse needs with respect to the objectives of the individual terminology projects, the theoretical orientations of the people creating and using the terminology, and the limitations or the flexibility provided by the software they are using (Wright, 1993: 179).

Por tanto, a la hora de trabajar con un sistema de gestión de este tipo habrá que tener en cuenta las necesidades que se nos presentan en cada proyecto y seleccionar un sistema que nos permita incluir los campos que se hayan determinado según los objetivos planteados.

En nuestro caso, al tratarse de una metodología de trabajo dirigida a los traductores profesionales tendremos en cuenta los resultados obtenidos en la encuesta

realizada *ad hoc* para conocer sus necesidades y expectativas (cf. 2.5.3.3.). En este contexto, hemos establecido la microestructura tipo de cada entrada en la base de datos, que incluye los siguientes campos:

- Información lingüística: término, categoría gramatical, género, número, abreviaturas y acrónimos, variantes denominativas u ortográficas y colocaciones.
- Información conceptual: área temática, definición, equivalentes de traducción y sinónimos.
- Información pragmática: ejemplos contextuales, marcas de uso y marcas de variación lingüística.

De esta manera, la información administrativa y las referencias no se incluirán en la microestructura de cada entrada, sino que se hará referencia a ellas en un apartado específico, junto con las instrucciones de uso del recurso final, una vez que se haya finalizado la elaboración de la base de datos terminológica.

En el diseño de la microestructura del diccionario no solo se ha de determinar qué tipo de información se va a incluir en cada entrada, sino también cuál es el formato seleccionado para cada campo y cuál va a ser el orden de aparición. En este caso, cada proyecto deberá decidir sobre estos asuntos, según sus objetivos, preferencias del terminógrafo y de las posibilidades que permita el sistema de gestión. Un ejemplo de entrada en un diccionario bilingüe puede ser el siguiente:

término ES *categoría gramatical*. DOMINIO. definición ES. *Contexto ES*;
remisión ES. término EN, sinónimo(s) EN. [Observación ES].⁵⁷¹

No obstante, como decimos, cada sistema de gestión permitirá el formato de cada entrada según sus posibilidades y las preferencias de cada terminógrafo.

En el capítulo cuarto, presentaremos el trabajo llevado a cabo en el dominio del turismo de aventura según la metodología propuesta en estos apartados, por lo que

⁵⁷¹ EN y ES hacen referencia a los códigos normalizados por la norma ISO 639-1 (2002) para hacer referencia a las lenguas español (ES) e inglés (EN).

trataremos en ese momento y de manera específica los diferentes campos que hemos seleccionado para la microestructura de nuestra base de datos. No obstante, nos parece relevante tratar aquí la elaboración de las definiciones y la selección de los ejemplos contextuales, puesto que consideramos que son los campos más complejos a la hora de proceder a la elaboración de la base de datos.

2.6.1.4.2. Elaboración de definiciones

En Terminografía, el papel que ocupan las definiciones es fundamental, puesto que proporcionan el vínculo entre el término y el concepto que designa. Además, es a través de las definiciones que los conceptos pueden ser identificados, formados, fijados y delimitados (Pozzi, 2001: 27). Por este motivo, vamos a tratar la elaboración de las definiciones como un apartado independiente dentro de la elaboración de la base de datos ontoterminográfica.

Según la norma ISO 1807 (1990: 4), el término «definición» se define como:

A statement which describes a concept and permits its differentiation from other concepts within the system of concepts.

Según esta norma, cada definición sirve para clarificar el concepto y las relaciones que este establece con el resto, por lo que la definición permite conocer la posición del concepto definido dentro de un sistema conceptual. Así pues, cualquier modificación en el sistema conceptual supondrá también un cambio en la definición.

Cabré Castellví (1993: 209) distingue tres tipos de definiciones, que se diferencian entre sí tanto por el objeto que describen como por los contenidos que expresan:

- la definición de tipo lingüístico;
- la definición de tipo ontológico;
- la definición de tipo terminológico.

Estos tres tipos de definición reflejan una misma realidad desde puntos de vista diferentes. Así pues, en el primer caso, que corresponde normalmente a las definiciones

lexicográficas, el objeto de la definición es el signo lingüístico; el de la definición de tipo ontológico, es la realidad; y el de la definición de tipo terminológico, el concepto del sistema conceptual de un campo de especialidad.

En consecuencia, estos tres tipos de definiciones también expresarán contenidos diferentes: en la definición lingüística, a diferencia de la definición terminológica, no se suelen incluir todas las características de un concepto, sino aquellas que son necesarias para distinguirlo de otro dentro del sistema de la lengua; la definición ontológica, por su parte, incluye todos los aspectos peculiares de un concepto, tanto si son o no relevantes para definirla como categoría; y, finalmente, una definición terminológica describe la noción en referencia exclusiva a un dominio de especialidad, y no en relación al sistema lingüístico.

Además, las definiciones terminológicas tienen la importante finalidad de fijar la relación entre un concepto y el término que lo designa. En palabras de Sager (1990: 21):

Through the activity of definition we fix the precise reference of a term to a concept, albeit by linguistic means only; at the same time it creates and thereby declares relationships to other concepts inside a knowledge structure.

De esta manera, las definiciones no solo establecen la relación entre el término y el concepto que designan, sino que, además, plasman las relaciones entre los conceptos de un sistema conceptual. Para esto, la definición terminológica tiene que cumplir con dos requisitos fundamentales: por un lado, hacer referencia al concepto genérico o superordinado del sistema conceptual y, por otro, proporcionar las características que lo distinguen del resto de conceptos dentro del mismo grupo (Sager y Ndi-Kimbi, 1995). Por tanto, la descripción del concepto debe estar formada por el concepto genérico o hiperónimo y el conjunto de características que lo distinguen. En 1916 Josephs, citado en Ndi-Kimbi (1994: 328), estableció la estructura de la definición terminológica mediante la siguiente fórmula:

$$\text{Definiendum} = \text{Genus} + \text{Differentiae}$$

donde el *definiendum* es el concepto que queremos definir; el *genus* es el concepto genérico o hiperónimo, y el *differentiae* son las características que lo distinguen del resto de conceptos englobados en la extensión del descriptor.

De Bessé (1997: 65), por su parte, establece una distinción entre definiciones de naturaleza prescriptiva y definiciones de naturaleza descriptiva basada en las distintas aproximaciones que normalmente poseen estas actividades: normativa en el primer caso y descriptiva en el segundo. Este autor señala que las prescriptivas, denominadas definiciones terminológicas, crean conceptos nuevos, los nombran y los normalizan; mientras que las segundas, llamadas definiciones terminográficas, solo los describen en referencia a un sistema conceptual preexistente. Consecuentemente, los encargados de formular las definiciones terminológicas serán los propios expertos de la materia y los organismos normalizadores correspondientes, mientras que las definiciones terminográficas serán elaboradas por terminógrafos que realicen trabajos descriptivos.

Las normas ISO, siguiendo los fundamentos de la Escuela de Viena, reconoce dos tipos de definiciones terminográficas: a) la definición intensional y b) la definición extensional. Además, reconocen un tercer tipo, la definición partitiva, aunque ocupa un lugar menor con respecto a las otras dos.

Las definiciones intensionales describen las características de un concepto, es decir, su intensión. Se parte del genérico más próximo y se enumeran las características que identifican el concepto y lo diferencian de los demás conceptos en la misma posición jerárquica.

Intensional definitions shall indicate the superordinate concept, either immediately above or at a higher level, followed by the characteristic(s) that distinguish the concept from other concepts. The superordinate concept situates the concept in its proper context in the concept system (ISO 704, 2000: 15).

Por su parte, las definiciones extensionales enumeran la extensión de un concepto, es decir, la totalidad de los objetos a los que ese concepto se refiere. Según la norma ISO (ibíd.: 17), este tipo de definición se debería utilizar únicamente cuando el número de conceptos que se debe enumerar es limitado. El ejemplo clásico de una definición extensional es la que recoge de Bessé (1997: 68) para el concepto <planet>:

planet

The planets of the solar system are Mercury, Venus, Earth, Mars, Jupiter, Saturn, Uranus Neptune and Pluto.

Por último, una definición terminográfica partitiva enumera todas las partes que constituyen el todo. Si esta misma definición que hemos facilitado arriba fuera de <solar system>, en lugar de ser de <planet>, la definición ya no sería extensional, sino partitiva (Wright y Budin, 1997: 339).

Las definiciones extensionales y partitivas pueden resultar muy útiles para la comprensión del concepto; sin embargo, son menos utilizadas que las intensionales, ya que en ocasiones puede resultar muy difícil enumerar todos los ejemplos o todas las partes de un concepto. Asimismo, a veces no son recomendables, puesto que no ayudan a la comprensión del concepto definido. La norma ISO 704 (2000: 15) recomienda el uso de la definición intensional siempre que se pueda, ya que, según indica, «they most clearly reveal the essential characteristics of a concept within a concept system».

A pesar de que tradicionalmente se han preferido las definiciones intensionales debido principalmente a su sistematicidad, en los últimos años están apareciendo nuevas tendencias en Terminología que subrayan los inconvenientes de este tipo de definiciones:

Intensional definitions gives only the necessary and sufficient characteristics that would enable a specialist (i.e. someone who already “understand the term”) to identify the object in question as an example of a category (Meyer, 1992: 23).

O como indica Temmerman (2000: 76):

We want to show that for some units of understanding [...] such a definition [intensional definition] is neither feasible, nor desirable because the impossibility of clearly delineating a unit is functional in science — we see defining as a never ending process having progress of both understanding and having knowledge as its motive.

Esta autora (ibíd.: 94) demuestra las dificultades e imposibilidades que presentan las definiciones propuestas por la terminología tradicional (definición intensional y extensional) a la hora de definir tres unidades de comprensión (*units of understanding*)

(cf. 1.3.2.2.3) pertenecientes al ámbito de la biología. Según las conclusiones extraídas, estas definiciones clásicas no son adecuadas para la definición de todos los conceptos, puesto que a menudo pueden resultar demasiado restrictivas y no proporcionan los conocimientos necesarios para comprender la unidad objeto de análisis. A partir de estos resultados, propone que la definición de los conceptos se base en la posición que ocupa cada concepto en el sistema conceptual al que pertenece, pero que, al contrario de utilizar únicamente las relaciones jerárquicas (lógicas y ontológicas) para su definición, como propone la TGT, se consideren las estructuras intercategoriales (dentro de las categorías) e intracategoriales (entre categorías).

Defining a category involve describing the relative position this category has in the conceptual structure. The types of information which we consider to be relevant in trying to define categories within ICMs (Idealized Cognitive Models)⁵⁷² concern the internal structure of the category (the intracategorical structure) and those concerning the relationships the category has with other categories in the ICM (the intercategorical structure (ibíd: 96).

Así pues, la autora aboga por unos principios y métodos para la elaboración de definiciones más flexibles, a fin de recoger los aspectos más relevantes de cada concepto y conseguir la comprensión óptima en terminografía (ibíd.: 121). En este contexto, considera que la utilización de marcos semánticos (o ICM) da la posibilidad de definir los conceptos de forma completa, sistemática y adaptada a la situación comunicativa en la que se dé y, gracias a la especificación de las relaciones que mantiene un concepto con el resto, tanto dentro como fuera de su categoría, permite una comprensión mejor de dicho concepto.

Puede observarse, pues, cómo los marcos semánticos sirven de gran ayuda a la hora de elaborar definiciones de todos los conceptos propios de un campo de especialidad. Otros autores, como veremos a continuación, también han utilizado los marcos semánticos para definir conceptos.

⁵⁷² ICMs (*Idealized Cognitive Models*) hacen referencia a los modelos cognitivos idealizados de Lakoff (1987), que, según el autor, son los modelos con los que estructuramos el conocimiento en nuestra mente y los que dan lugar a las estructuras categoriales y a la creación de prototipos. Para profundizar en el tema, remitimos al primer capítulo (cf. 1.3.2.3.3.1).

Strehlow (1997: 77) compara la representación de la estructura de una definición con la representación de un sistema conceptual, aunque advierte que los métodos tradicionales empleados para la elaboración de sistemas conceptuales pueden resultar insuficientes e incluso confusos a la hora de representar definiciones de conceptos. Por eso, considera que el uso de marcos semánticos es una buena alternativa para ofrecer la representación de los conceptos y sus características y las relaciones que estos presentan con respecto a otros conceptos del sistema, a fin de a) ofrecer una representación adecuada y b) distinguir entre conceptos próximos.

Por su parte, Martin (1998) presenta su modelo de definiciones basadas en marcos (*frame-based definitions*, FBD) de acuerdo con la definición de «marco» de Fillmore (cf. 1.3.2.3.3.1.), con el que realiza las definiciones de los conceptos mediante *slots* (categorías conceptuales) y *fillers* (atributos y valores de los conceptos). Según este autor (ibíd: 206), establecer las definiciones con marcos favorece la uniformidad entre conceptos de la misma categoría, ya que se emplea el mismo modelo para su definición, y son más completas, puesto que permiten el empleo de otras relaciones además de las lógicas y ontológicas. Asimismo, las definiciones son mucho más flexibles, ya que no se trata de un listado rígido y almacenado de características, sino que es una estructura dinámica que se adapta a las necesidades del usuario y a la multidimensionalidad de cada concepto.

Por último, Fillmore (2003) también hace referencia al uso de los marcos para sus *double-decker definitions* y subraya la importancia de que la definición remita a la estructura conceptual que subyace a la palabra. Asimismo, destaca de nuevo la necesidad de que las definiciones sean flexibles y aporten información sobre las relaciones jerárquicas y no jerárquicas que se establecen entre conceptos.

The double-decker definitions [...] will obviously need to treat different kinds of words in very different ways, but in each case it is important to give the user access to the conceptual structure that underlies the word, information about the position or orientation of its meaning within that conceptual structure, and a generous supply of examples of how the word interacts with other words and phrases in the utterances of language users (ibíd.: 289).

En esta misma línea, se encuentra nuestra propuesta metodológica, en la que proponemos la elaboración de *patrones definitorios* concretos y sistemáticos basados en marcos semánticos elaborados para el campo de especialidad objeto de trabajo, con la intención de:

1. Elaborar definiciones que remitan al marco semántico del dominio, con objeto de facilitar la comprensión del concepto y de las relaciones establecidas con otros conceptos cercanos.
2. Elaborar definiciones que indiquen explícitamente la pertenencia del concepto a una categoría.
3. Elaborar definiciones completas de los conceptos, donde se indiquen tanto las relaciones jerárquicas (genéricas y partitivas) como las no jerárquicas o asociativas.
4. Elaborar definiciones flexibles que permitan ajustarse al grupo potencial de destinatarios, así como recoger la multidimensionalidad de los conceptos.
5. Elaborar definiciones coherentes y homogéneas de los conceptos miembros de una misma categoría.

Para establecer dichas plantillas o patrones definitorios es necesario partir de la estructura básica de cualquier definición terminológica, a saber: *Definiendum* = *genus* + *differentiae*, donde el *Definiendum* es el término que se debe definir; el *genus* se corresponde con su hiperónimo, que vendrá determinado por la relación jerárquica IS-A; y el *differentiae* hace referencia a las características propias que lo diferencian de otros conceptos, y que será introducido por las relaciones asociativas establecidas para cada dominio. Como ejemplo ilustrativo, podemos presentar una definición extraída de García de Quesada (2001) con respecto a un medicamento, la bleomicina, donde se observa claramente el patrón seguido, a saber: el *genus*, con la relación IS-A, que hace referencia a su hiperónimo, y el resto de características propias con las relaciones semánticas correspondientes incluidas entre corchetes:

antibiótico antineoplásico [IS-A], cuyo nombre comercial es Bleoxane [HAS-TRADENAME], utilizado en la quimioterapia [USED-IN] de

tumores germinales, osteosarcoma, enfermedad de Hodgkin, sarcoma de Kaposi y carcinoma epidermoide de cabeza y cuello, pene y cervix [USED-IN-THE-TREATMENT-OF].

Siguiendo esta metodología, será necesario, por un lado, utilizar una misma plantilla para los conceptos pertenecientes a una misma categoría, es decir, que compartan el mismo hiperónimo, pero que presentan diferentes características que los diferencian unos de otros (para el caso expuesto anteriormente, será necesario que todos los medicamentos definidos sigan el mismo patrón de definición); y, por otro lado, elaborar diferentes plantillas para definir conceptos pertenecientes a diferentes categorías.⁵⁷³

Una vez establecida la metodología que proponemos en nuestra investigación para la elaboración de definiciones basadas en marcos semánticos, es necesario tener en cuenta los principios básicos que se deben seguir a la hora de elaborar definiciones adecuadas y, por tanto, exentas de problemas terminológicos. Como vimos anteriormente (cf. 2.4.2.1), tanto la ISO como otros centros han establecido una serie de recomendaciones para elaborarlas, que empleamos para realizar una evaluación de las definiciones encontradas en recursos terminográficos. En este contexto, también las seguimos pero para un objetivo diferente: para elaborar definiciones adecuadas dentro de nuestra propuesta metodológica. De esta manera, nuestras definiciones deberán cumplir las siguientes condiciones:

- Evitar expresiones y formulaciones de tipo metalingüístico (por ejemplo, «Dícese de»);
- Evitar palabras más complicadas que la palabra definida;
- Adaptar la formulación de la definición a las necesidades específicas de los usuarios y utilizar palabras conocidas por los usuarios en general. En el caso de utilizar unidades más específicas, estas deberán figurar como términos definidos dentro del trabajo.
- No utilizar en la definición el término definido;

⁵⁷³En el capítulo cuarto (cf. 4.2.3.1.1.), se presentan las diferentes plantillas definitorias establecidas dentro de esta investigación para el dominio del turismo de aventura.

- Partir de un descriptor de la misma categoría gramatical que el término definido;
- Utilizar el mismo descriptor para todos los términos que pertenecen a una misma clase o subclase conceptual;
- Distinguir el concepto de otros conceptos similares del mismo o diferente campo de especialidad mediante las relaciones semánticas que ese concepto establece con el resto de los conceptos del mismo campo.
- Contener una sola oración, sin contener puntos en el interior;
- Evitar paráfrasis conceptualmente vacías, es decir, que reproducen la misma información que puede desprenderse del propio término.
- Evitar la circularidad, tanto en la definición como en el sistema de definiciones
- Evitar la definición de un concepto por la negación de otro.
- No traslucir ninguna ideología y ser veraz.
- Adecuarse a las finalidades del trabajo en que se muestran tales definiciones.

Llegados a este punto en el que hemos seleccionado el tipo de definición en la que vamos a trabajar y hemos revisado los principios básicos para su elaboración, ya tenemos establecidas las directrices para comenzar con la elaboración de definiciones. En definitiva, a pesar de haber fijado estos datos, el trabajo de elaboración es personal y dependerá de cada terminógrafo y del esquema categorial y los marcos que haya elaborado en las fases previas. No obstante, si se siguen estas indicaciones, las definiciones que se obtengan serán sistemáticas y adecuadas para los objetivos de nuestro trabajo.

A continuación, trataremos otro aspecto importante para la elaboración de la base de datos: la selección de los contextos.

2.6.1.4.3. Selección de ejemplos contextuales

Se entiende por «contexto» el entorno lingüístico de un término en un enunciado, es decir, la oración u oraciones donde aparece dicha unidad, que condicionan su existencia, su forma, su funcionamiento, su sentido, su valor y su uso (Gómez González-Jover, 2006: 596). Así pues, podemos afirmar que el contexto es la mejor manera de contextualizar los términos que se han extraído previamente para recopilar la nomenclatura de la base de datos, además de para comprender mejor parte del dominio de especialidad al que pertenece. Se trata de un trabajo cíclico, que parte primero del texto para extraer los términos de sus contextos reales para después volver al texto, de donde se extraerá esta vez únicamente una pequeña porción: el contexto de una unidad concreta. Según Dubuc y Lauriston (1997: 81), el contexto desempeña una doble funcionalidad:

The context [...] plays a double role: it first provides living proof that the term is used in the field of reference and then allows a conceptual content to be associated with the term being studied. The second role is the more important. It makes it possible to determine the specific relationship between a term and its subject field through concept identification.

De esta manera, se vuelve imprescindible la inclusión de contexto (o ejemplos contextuales) en el trabajo (onto)terminográfico, si tenemos en cuenta además la necesidad manifiesta del grupo de destinatarios al que se dirige cualquier trabajo (onto)terminográfico realizado con esta metodología, es decir, los traductores profesionales, que por lo general no cuentan con un conocimiento experto del dominio en el que están trabajando.

Las diferentes metodologías de trabajo analizadas proponen diversos criterios para llevar a cabo la selección de ejemplos contextuales, aunque todas coinciden en los siguientes:

- El contexto debe extraerse de una fuente fiable y ser textos originales.⁵⁷⁴

⁵⁷⁴ Por regla general, los contextos se extraen del propio corpus compilado en las fases previas. En aquellos casos en los que el corpus no contenga ejemplos de uso adecuados para el diccionario, el terminógrafo deberá acudir a fuentes externas (otros textos no incluidos, fuentes de referencia, etc.).

- El contexto deber ser conciso, pero completo, es decir, el fragmento seleccionado debe ser lo suficientemente explícito para aportar información relevante.
- El contexto debe contener la unidad terminológica que pretenda ilustrar en la misma categoría gramatical de la entrada, es decir, si la entrada hace referencia a un sustantivo, el término del contexto debe ser un sustantivo, aunque haya un adjetivo con la misma forma.
- El contexto debe ser representativo, es decir, debe representar el sistema conceptual al que pertenece el concepto al que hace referencia el término.

Asimismo, según Fauconnier y Turner (2002) y Evans y Greens (2006), los contextos proporcionan los medios para que el destinatario construya un espacio mental parcial del dominio sobre el que está consultando la información, por lo que los contextos que se seleccionen deberán también mostrar relaciones semánticas de la unidad en cuestión con algunos de sus conceptos relacionados, a fin de plasmar la realidad conceptual de dicha unidad y facilitar su comprensión al usuario. En este sentido, los contextos se asemejan a las definiciones, aunque se diferencian en el hecho de que los primeros se extraen de textos que plasman una situación comunicativa especializada real, mientras que las últimas se crean por los terminógrafos a partir del marco semántico establecido previamente.

En definitiva, la inclusión de los contextos en los trabajos (onto)terminográficos es imprescindible para transmitir a los usuarios toda la información necesaria con relación a la unidad consultada y su uso en el discurso propio del ámbito de estudio, por lo que su selección no debe ser arbitraria. Además, se vuelve especialmente importante en el caso de los traductores profesionales, que deben adquirir sus conocimientos expertos y discursivos para producir enunciados y utilizar correctamente los términos en las oraciones de la lengua meta. Como afirman Fuentes Morán y García Palacios (2002: 84), el contexto es algo más que «un mero valor añadido en la organización de la

Téngase en cuenta que no se recomienda la creación *ad hoc* de contextos, puesto que desvirtuaría completamente el trabajo basado en una comunicación real.

estructura informativa del artículo lexicográfico, y adquiere una polifuncionalidad que le capacita para proporcionar información de muy diversa índole».

2.6.1.5. Fase V: Revisión y validación

Esta fase incluye una revisión y validación final de la aplicación terminológica por parte del terminógrafo y de los colaboradores expertos del ámbito objeto de estudio. Estos últimos intervienen principalmente en cuestiones conceptuales y pragmáticas, puesto que advierten con cierta facilidad cualquier error conceptual, la ausencia de términos importantes para el ámbito, pueden sugerir la eliminación de algún término por considerarlo no pertinente o innecesario, el uso específico de algún término, entre otros. Por su parte, de los aspectos lingüísticos se encarga el terminógrafo, cuya formación en este campo le permite realizar valoraciones más precisas durante la revisión de las definiciones, la adecuación de los contextos o la validez de los equivalentes.

El objetivo primordial de esta fase consiste en asegurar que los datos terminológicos sean precisos y completos, al tiempo que satisfacen las necesidades de los usuarios finales. Para ello, es necesario que el terminógrafo realice dos tareas principalmente:

- a) Solucionar posibles casos problemáticos, que no se atendieron en algún momento, como pueden ser neologismos, la definición de algún término, la búsqueda de algún equivalente de traducción, o la gestión de algún tipo de variación.
- b) Revisar de forma detenida y exhaustiva toda la información contenida en cada una de las entradas con el fin de detectar y subsanar posibles errores (Vargas Sierra, 2006:
 1. Repetición de registros;
 2. Campos obligatorios completados;
 3. Funcionamiento correcto de referencias cruzadas;
 4. Pertinencia de los términos;
 5. Revisión ortográfica y gramatical;

7. Redacción de definiciones;
8. Sintaxis de contextos, contenidos, etc.;
9. Referencias bibliográficas;
10. Equivalente: misma categoría gramatical o misma función gramatical.

Una vez que se han realizado todas las modificaciones y correcciones pertinentes, se pasará a la siguiente fase para elaborar la versión definitiva del trabajo y proceder a su publicación.

2.6.1.6. Fase VI: Exportación y edición

La última fase de cualquier proyecto (onto)terminográfico consiste en la edición, ya sea en papel o en formato electrónico, de la aplicación que se pondrá a disposición del grupo de destinatarios final. Cabe destacar que, al igual que en algunas ocasiones anteriores, las tareas que se realizan en esta fase dependen directamente del soporte elegido en un principio y, según sus posibilidades la edición se realizará de una forma o de otra.

El aspecto más destacado de la edición es la exportación de la información terminológica que se encuentra en la base de datos. Esta tarea dependerá de los objetivos planteados inicialmente en el proyecto, es decir, se realizará en un formato u otro dependiendo de si la edición del recurso terminográfico se va a realizar en formato papel o electrónico (CD-ROM o Internet). En el caso de un recurso que va a ser publicado en papel, será necesario volcar los datos electrónicos contenidos en la base de datos a un formato de texto (.doc, .rtf, etc.) para que puedan manipularse en la etapa posterior de maquetación antes de su publicación. Por el contrario, si va a publicarse en formato electrónico, deberemos utilizar las posibilidades de exportación del sistema de gestión de la base de datos o, si no tuviera disponible esta posibilidad, utilizar algún tipo de programa que permitiera la edición correspondiente. Como vimos en los resultados de la encuesta realizada en esta investigación (cf. 2.5.3.3.), los traductores profesionales prefieren la consulta de recursos en formato electrónico debido a las ventajas que presenta: facilidad y rapidez de acceso, sin necesidad de almacenaje, acceso remoto,

etc., por lo que nuestra metodología estará dirigida a la publicación de recursos en línea. No obstante, los pasos que hemos descrito hasta el momento sirven tanto para la edición de recursos en papel como en formato electrónico.

En este contexto, se recomienda también haber utilizado un sistema de gestión de bases de datos que proporcione la posibilidad de editar los productos finales, es decir, que permita la exportación en formatos de edición de la base de datos terminológica, a fin de agilizar esta fase y evitar la necesidad de utilizar más herramientas. Esta posibilidad es una de las características que tendremos que tener en cuenta a la hora de seleccionar la herramienta con la que vamos a trabajar para que satisfaga las necesidades que nos hemos planteado en la fase inicial de nuestro proyecto. Para la edición de productos electrónicos, la aplicación nos deberá permitir la exportación en formato HTML, XML, TBX o cualquier otro formato que permita su publicación electrónica. En nuestro caso, defendemos la exportación en formato XML o TBX⁵⁷⁵ por ser los formatos estándares que permiten la reutilización en un futuro del material terminográfico elaborado por parte del terminógrafo o por otras aplicaciones informáticas, aunque estos no están siempre disponibles en las aplicaciones con posibilidad de acceso, ya sea este libre o comercial. Y, por tanto, la herramienta seleccionada para nuestro trabajo ontoterminográfico permitirá la posibilidad de exportar en dicho formato de intercambio de información.

Llegados a este punto, donde todo el trabajo ontoterminográfico está realizado y solo falta exportar la información elaborada según la macroestructura y microestructura determinadas en las fases previas y desarrollar la información acerca del recurso (características generales, objetivos del diccionario, instrucciones de uso, bibliografía consultada, etc.), para incluirla en un apartado específico del producto final,⁵⁷⁶ consideramos que es esencial prestar atención a las necesidades y expectativas que

⁵⁷⁵ El formato TBX es un tipo de .XML que sigue una estructura determinada y normalizada por la norma ISO 30042 (2008) (cf. 4.3.1.3.1.1.). Muchos editores de este tipo de productos no permiten aún la exportación en este formato, aunque sí exportar en formato .XML con una estructura especificada previamente mediante una DTD (*document type definition*). De esta manera, un formato .XML se puede convertir en un .TBX mediante la utilización de la DTD establecida por esta norma ISO.

⁵⁷⁶ Al tratarse de un recurso electrónico, podría dedicarse a esta información un apartado denominado *Acerca de* o cualquier otro título similar.

presentan los usuarios finales (en este caso, los traductores profesionales) para ofrecerles un recurso a su medida y útil para sus necesidades.

De esta manera, proponemos que la edición de la base de datos ontoterminográfica sea lo más flexible y dinámica posible, además de que ofrezca un acceso fácil y rápido. Para ello, consideramos imprescindible que los recursos finales destinados a los traductores profesionales ofrezcan una serie de aspectos básicos:

- remisiones mediante hiperenlaces a otros términos contenidos en la base de datos y que pertenecen al mismo marco semántico que el término buscado;
- varias formas de acceso a la información contenida, como por ejemplo la búsqueda mediante la introducción del término deseado, la búsqueda alfabética o, si es posible, la búsqueda textual.⁵⁷⁷
- flexibilidad a la hora de seleccionar la información que aparece en cada entrada según las necesidades del propio usuario.

Con respecto al primer punto, la remisión a otros términos del recurso no existe ninguna discusión al respecto, ya que, en general, se considera una forma óptima de ampliar la información contenida en cada entrada y proporcionar más conocimientos sobre el marco semántico al que pertenece el concepto en cuestión. En los recursos en formato papel, la posibilidad de remisiones más rápidas mediante hiperenlaces se pierde, aunque estas remisiones se realizan mediante la inclusión de indicaciones sobre los términos relacionados al final de la entrada, los cuales tienen que ser buscados explícitamente por el usuario si desea ampliar la información.

El segundo punto hace referencia a un aspecto más complejo, puesto que no se tiene conocimiento exacto sobre cuál es la mejor opción de búsqueda para los usuarios en general. Cada uno puede preferir un tipo de acceso u otro, por lo que no se debe generalizar. Por lo general, los recursos electrónicos suelen ofrecer más de un tipo de búsqueda, que son principalmente la búsqueda por término, es decir, se introduce el término que se desea consultar y se obtiene la entrada de dicho término, y la búsqueda

⁵⁷⁷ Otro tipo de búsqueda podría ser la onomasiológica, que consiste en la búsqueda de propiedades y relaciones que caracterizan a los términos (cf. Alcina Caudet, 2008).

alfabética, o sea, una lista de los términos incluidos en la base de datos ordenados de forma alfabética. Desde esta metodología, presentamos otra opción de búsqueda, que agilizaría la consulta y el acceso a la información por parte de los traductores profesionales, e incluso otros usuarios. La hemos denominado *búsqueda textual*, ya que la base de la búsqueda es el texto mismo donde aparecen los términos que se deben consultar.

Esta búsqueda textual consiste en destacar en el mismo texto de trabajo los términos contenidos en la base de datos, es decir, en el texto aparecen indicados con hiperenlaces los términos incluidos en la base de datos para que el traductor pueda, de un solo vistazo, ver cuáles son los términos que están disponibles y cuáles no. Una vez que el traductor obtiene el texto con los términos indicados, podrá acceder a la información terminológica de cada uno pinchando directamente en el término que desee consultar. De esta manera, se consigue un acceso rápido, ágil y fácil a la información contenida en las entradas de los términos y se evitan, por consiguiente, las búsquedas infructuosas de términos por no estar incluidos en los recursos utilizados. Este tipo de búsqueda ha sido ya incluido en algunos recursos conocidos como asistentes para la comprensión (*comprehension assistants*).⁵⁷⁸ No obstante, sugerimos que esta búsqueda, se acompañe además de otro tipo de búsqueda como puede ser la búsqueda por término, ya que en algunas ocasiones el traductor profesional desea realizar búsquedas de términos que no aparecen en el texto que posee.

Por último, nos parece relevante señalar dos aspectos más: por un lado, el hecho de que el usuario final (en nuestro caso, el traductor profesional) no siempre resulta satisfecho con la información que encuentra en un recurso y necesita la consulta de otros para completarla o confirmarla; y por otro, a veces los recursos disponibles no ofrecen toda la terminología necesaria para el traductor, ya sea porque se trata de un neologismo o porque no se incluyó en el momento de la elaboración de la base de datos. Por ello, destacamos la necesidad, en la medida de lo posible, de facilitar al usuario final

⁵⁷⁸ En el siguiente capítulo (cf. 4.3.1.1.) se realizará una revisión de algunos de los sistemas actuales que presentan estas características.

la búsqueda de la información que desea encontrar de manera externa, ya sea dándole acceso a otros recursos o incluso a buscadores generales como *Google*.

El tercer punto, que hace referencia a la flexibilidad a la hora de seleccionar la información que aparece en cada entrada según las necesidades del propio usuario, está relacionado con la teoría de la función (*function theory*), de Tarp (2008a), y con los diccionarios multifuncionales electrónicos (Spohr, 2009). En síntesis, estas dos propuestas indican que el contenido de las entradas se debe ajustar a las necesidades de cada grupo de destinatarios, a fin de que el producto satisfaga las funciones para las que se va a usar, por ejemplo la comprensión o la reproducción de mensajes. Asimismo, la segunda propone la confección de una base de datos terminográfica genérica compuesta por una completa información y que, a la hora de editar el recurso final, se realice una selección de la información que se desea mostrar según los usuarios finales. Estas dos propuestas son genéricas, es decir, su trabajo terminográfico no está centrado en ningún grupo de usuarios determinado, sino que a la hora de editar el recurso utilizan una u otra información según estos usuarios. En nuestro caso, esto no es así, ya que hemos dejado claro desde la primera fase del trabajo que nuestro grupo de destinatarios serían los traductores profesionales. No obstante, también se detectó mediante la encuesta *ad hoc* realizada en esta investigación (cf. 2.5.3.3.) que no todos los usuarios necesitan el mismo tipo de información, ya que esta depende de 1) la formación del traductor con respecto al dominio en el que está trabajando y 2) las preferencias de este a la hora de realizar cualquier consulta. Asimismo, un mismo traductor puede necesitar un tipo de información en un momento dado y otra en otro. Por ello, y aunque trabajemos con los campos terminológicos que han recibido mejores resultados en la encuesta, sugerimos que, antes de la consulta del producto por parte de los usuarios finales, se les proporcione la posibilidad de seleccionar la información que deseen visualizar en la entrada terminológica de cada término. Así, por ejemplo, un traductor podrá seleccionar que en la entrada aparezca simplemente el término que está buscando acompañado por su equivalente en una lengua determinada y la definición; y otro, podrá elegir el término, la categoría gramatical, el equivalente, la definición y el contexto, o simplemente el término y el equivalente. De esta manera, se conseguiría un recurso que ofrece la posibilidad de realizar consultas flexibles, dinámicas y rápidas, además de

ofrecer la posibilidad de reutilizar la información según las preferencias y necesidades de los usuarios finales.

En nuestra opinión, estas funcionalidades son básicas para la buena consulta del trabajo ontoterminográfico realizado en la fases previas y, por tanto, para la satisfacción de los usuarios del recurso, por lo que los proyectos terminográficos deberían tener en cuenta estos puntos a la hora de realizar la edición de sus productos. Asimismo, nos parece importante destacar que, a pesar de encontrar proyectos elaborados con una metodología sistemática, coherente y completa, carecen de métodos adecuados para la consulta de las bases de datos elaboradas. En otras palabras, dedican gran cantidad de esfuerzo a la gestión de la información y la elaboración de la base de datos, pero llevan a cabo una edición convencional y obviando los requisitos de los usuarios finales. De esta manera, a pesar de contar con numerosas herramientas para la gestión de corpus, búsqueda de concordancias, extracción de terminológica, etc., se siguen realizando ediciones de bases de datos en formato .rtf o .pdf para su posterior edición en papel o .html para su edición en formato electrónico, lo que dificulta el acceso a la información y, por tanto, la consulta por parte de los usuarios.

En definitiva, consideramos que se han realizado enormes avances en las herramientas de apoyo al trabajo (onto)terminográfico, pero no en lo que respecta a la edición de recursos terminológicos, concretamente en formato electrónico. En este sentido, detectamos carencias a este respecto que vamos a intentar solventar con una propuesta de una aplicación informática, *Trandix*, desarrollada en el marco de esta investigación y que tendrá como objetivo el ofrecer una herramienta novedosa que tenga en cuenta los aspectos básicos que comentábamos con anterioridad y que proporcione, así, la posibilidad de realizar consultas de forma flexible, cómoda y rápida.

Una vez que hemos tratado nuestra propuesta metodológica a lo largo de este capítulo, que se resume en la tabla siguiente y hemos explicado con profundidad en qué consisten las diferentes fases, vamos a continuar en el siguiente capítulo nuestra investigación centrándonos en un ámbito concreto de especialidad: el segmento de

turismo de aventura, con objeto de aplicar todas las fases y tareas concretadas en este capítulo según la metodología propuesta.

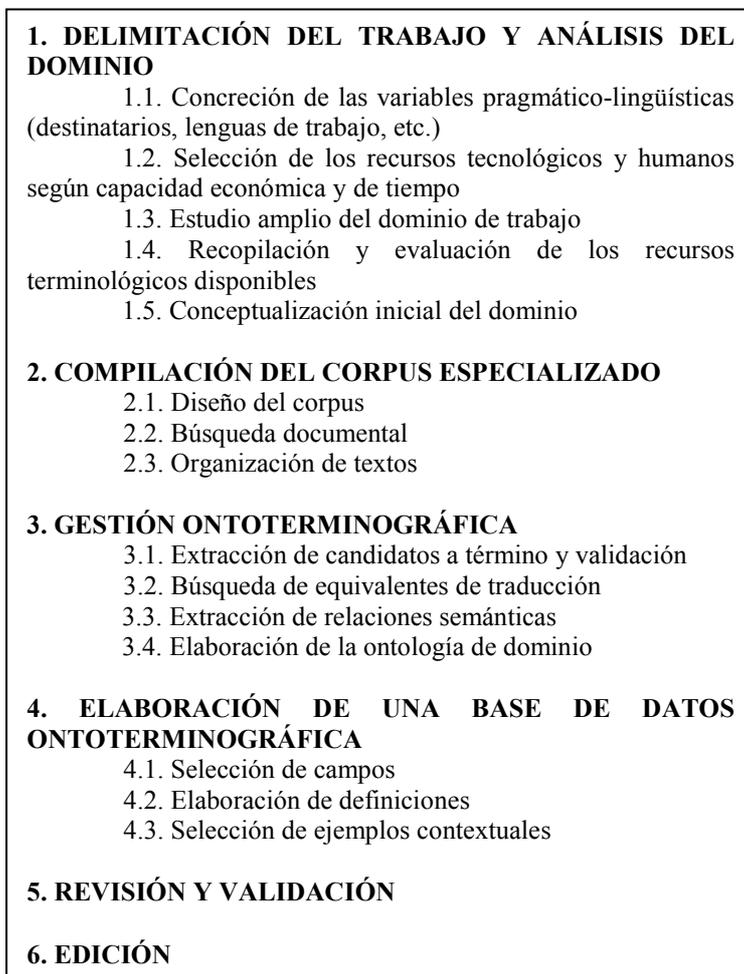


Figura 26. Fases de la propuesta metodológica para la elaboración de recursos ontoterminográficos para traductores.

CAPÍTULO 3.

PREPARACIÓN DEL TRABAJO ONTOTERMINOGRÁFICO

ÍNDICE DE CONTENIDOS

- 3.1. Introducción
- 3.2. Delimitación de trabajo
 - 3.2.1. Selección de las herramientas de trabajo
- 3.3. Análisis del dominio
 - 3.3.1. El concepto *turismo*
 - 3.3.2. El turismo como fenómeno económico en España
 - 3.3.3. Los segmentos turísticos en auge
 - 3.3.3.1. El turismo de aventura como turismo alternativo
 - 3.3.3.1.1. Vínculos entre turismo y deporte
 - 3.3.3.1.2. El turismo de aventura en España
 - 3.3.3.1.3. Actividades del turismo de aventura y propuestas para su clasificación
 - 3.3.4. El discurso turístico como lenguaje de especialidad
 - 3.3.4.1. Características del lenguaje del turismo como lenguaje de especialidad
 - 3.3.4.1.1. Plano lingüístico
 - 3.3.4.1.2. Plano pragmático
 - 3.3.4.1.3. Plano funcional
 - 3.3.4.2. Características formales del texto turístico
 - 3.3.5. La traducción de textos turísticos en España
 - 3.3.5.1. Características de la traducción de textos turísticos promocionales: problemas, dificultades y errores frecuentes
 - 3.3.6. Recursos terminológicos monolingües, bilingües y multilingües para las lenguas inglés, alemán y español en el segmento de turismo de aventura
 - 3.3.6.1. Búsqueda y selección de los recursos terminológicos
 - 3.3.6.2. Clasificación de los recursos terminológicos
 - 3.3.6.3. Evaluación de los recursos terminológicos
 - 3.3.7. Esquema categorial del dominio
- 3.4. Compilación del corpus AVENTURIS
 - 3.4.1. El subcorpus comparable AVENTURIS 1
 - 3.4.2. El subcorpus paralelo AVENTURIS 2

Confrontados con la realidad de treinta millones de términos es obligado reconocer que un terminógrafo, sea una persona o un grupo, no puede abarcar toda esa extensión. Por ello, es imprescindible trabajar sobre dominios o campos especiales (*subfields*) (Marcos Marín y Sánchez Lobato, 1991: 139).

3.1. INTRODUCCIÓN

Llegados a este punto en el que hemos realizado una revisión de las corrientes tanto teóricas como prácticas de la Terminología, así como una propuesta metodológica para la realización de un trabajo ontoterminográfico basado en corpus dirigido a un grupo de destinatarios concretos, tenemos establecido el marco de trabajo y los pasos que debemos seguir para llevar a cabo el trabajo que nos hemos propuesto para esta tesis doctoral. De esta manera, comenzaremos la parte práctica de nuestra investigación de acuerdo con la metodología propuesta, que se enmarca dentro de las teorías modernas de la Terminología y de la Semántica de Marcos.

Como vimos en el capítulo anterior, el trabajo ontoterminográfico y, por extensión, cualquier trabajo terminográfico parte de una fase de preparación, donde el terminólogo debe establecer las líneas de trabajo así como los materiales y los recursos que servirán de base para la elaboración de la base de datos ontoterminográfica y su posterior edición. Para ello, nuestra propuesta metodológica presentada en el capítulo

anterior divide esta fase inicial de preparación en varias fases diferentes, a saber: 1. La selección del dominio y delimitación del trabajo, donde se presentan las diferentes decisiones pragmático-lingüísticas tomadas por el terminógrafo para este proyecto de investigación en concreto, así como la presentación de las diferentes herramientas informáticas que van a servir de ayuda durante todo el proceso; 2. Análisis del dominio, donde se lleva a cabo un estudio sobre el dominio que se ha elegido en el punto anterior y se establecen los límites teóricos, lingüísticos y pragmáticos que concretan el estudio de la terminología y de sus características intrínsecas; 3. Estudio de los recursos disponibles según lenguas y dominio de trabajo, con objeto de establecer el estado de la cuestión en lo referente a los posibles recursos similares al que se desea elaborar y las ventajas y desventajas que presentan para el grupo de destinatarios objeto de estudio; 4. Compilación del corpus, que permite establecer la “materia prima” de la que parte el proceso de trabajo del terminólogo, tanto para el estudio más profundo de los rasgos del dominio como para la extracción de terminología, de relaciones semánticas, de equivalentes, etc., y 5. Elaboración de la conceptualización del dominio, donde se aúnan todos los conocimientos adquiridos hasta ese momento por parte del terminólogo en lo referente al dominio de trabajo a fin de proponer una representación conceptual inicial que sirva como base para continuar con el trabajo ontoterminográfico.

Siguiendo los pasos indicados previamente, ofreceremos como ejemplo un proceso de trabajo exhaustivo a lo largo de este capítulo, es decir, describiremos un modelo de trabajo enmarcado en la metodología ontoterminográfica donde se indicarán las decisiones previas relativas a las variables pragmático-lingüísticas que se deben tomar en lo que respecta a un proyecto de trabajo determinado, centrado en unos destinatarios concretos y con un dominio especializado, el turismo de aventura. Asimismo, veremos cómo se lleva a cabo un análisis del dominio seleccionado en todas sus vertientes, así como la compilación de un corpus textual centrado en el dominio objeto de estudio, para terminar con la propuesta de conceptualización del dominio.

En definitiva, en este capítulo trataremos en profundidad la fase de preparación del trabajo ontoterminográfico y ofreceremos una explicación exhaustiva de los diferentes pasos que se deben dar para continuar, en el capítulo siguiente, con la elaboración propia de la base de datos ontoterminográfica.

3.2. DELIMITACIÓN DEL TRABAJO

Como explicamos en la presentación de nuestra metodología, la primera etapa consiste en el establecimiento de las variables pragmático-lingüísticas que caracterizarán al recurso y que se mantendrán a lo largo de todo el proceso de trabajo según los objetivos de nuestra investigación. Las decisiones sobre estas variables hacen referencia al tema sobre el que se va a trabajar, la(s) lengua(s) de trabajo, el grupo de destinatarios al que va destinado, la finalidad que presenta, la aplicación que se desea dar al recurso y sus dimensiones.

Generalmente, las decisiones que se toman sobre estas variables suelen estar determinadas por la carencia de recursos sobre el campo o porque el material existente no cumple las necesidades o está desfasado (Arntz y Picht, 1995: 255), aunque también influye las decisiones administrativas de la empresa, organización o proyecto en el que se enmarque el trabajo. En nuestro caso, la razón de la elección de nuestro tema de trabajo, el turismo de aventura, ha sido debida a las dos causas principalmente, a saber: en primer lugar, como veíamos en la introducción de este trabajo, nuestra investigación se enmarca en el proyecto de investigación HUM-892, *La contratación turística electrónica multilingüe como mediación intercultural: aspectos legales, traductológicos y terminológicos*, cuyos objetivos contemplan el análisis de material promocional de los diferentes segmentos turísticos más importantes en España, la elaboración de un protocolo de evaluación electrónica de sitios web turísticos promocionales multilingües, la elaboración de repertorios lexicográficos electrónicos, multilingües y modulares, entre otros, y en el proyecto I+D *Ecosistema: Espacio único de sistemas de información ontológica y tesauros sobre el medio ambiente*, que tiene por objetivos la creación de ontologías y recursos ontoterminográficos sobre los segmentos turísticos relacionados con la naturaleza y el medio ambiente. La otra razón de la elección del tema se basa en el hecho de que el segmento del turismo de aventura es uno de los segmentos turísticos en auge actualmente, tanto por parte del turismo nacional como internacional, y que, sin embargo, carece actualmente de la atención merecida por parte de los terminógrafos y por parte de muchas empresas que ofrecen servicios de este sector. Por estas dos razones, decidimos embarcarnos en el estudio de este segmento turístico y trabajar en un recurso que cubriera las necesidades que se presentan en la actualidad, especialmente a

los traductores profesionales que deben trabajar con textos promocionales de este sector, y que fuera descriptivo, es decir, no perseguimos la normalización de la terminología de este sector, aunque sí poner un poco de orden en las variantes denominativas y en los errores conceptuales existentes.

Derivado de las dos razones anteriores, seleccionamos las lenguas de trabajo, que serían, además del español, el inglés y el alemán, puesto que son los dos mercados emisores de turistas más importantes para nuestro país según los informes oficiales aportados por el Instituto Nacional de Estadística.⁸⁸¹ De esta manera, determinamos que nuestro recurso sería trilingüe, aunque unidireccional, es decir, el español sería la lengua de trabajo para la elaboración de definiciones, marcas de uso, etc. y se ofrecerían equivalencias en inglés y alemán.

Con respecto al grupo de destinatarios, ya hemos avanzado que serían los traductores profesionales, uno de los grupos de usuarios que más utilizan los recursos terminológicos y que menos se consideran a la hora de elaborar cualquier producto. De este modo, la aplicación de la obra queda también clara: se tratará de un recurso destinado tanto a la descodificación, es decir, a la comprensión del término original, como a la codificación, es decir, a la reproducción del término original en la lengua meta.

Por último, se deben decidir las dimensiones de la obra, no tanto para determinar el número exacto de entradas que va a incluir, sino más bien para determinar el equipo y los recursos materiales necesarios para llevarla a cabo. En nuestro caso, la obra contará con unas dimensiones muy limitadas, debido a que el dominio seleccionado no es muy extenso, por lo que bastará con un terminógrafo que esté asesorado por expertos en la materia (monitores de las actividades, expertos en las diferentes actividades y en turismo, etc.).

Asimismo, después de haber concretado las variables anteriores, es necesario establecer las siguientes cuestiones (IULA, 2008):

⁸⁸¹ <<http://www.ine.es/>>. En los siguientes apartados profundizaremos en la situación del mercado turístico en España, tanto con respecto a los mercados emisores más importantes como con relación a los diferentes segmentos turísticos.

- Tipos de unidades terminológicas seleccionadas. En nuestro caso, seleccionaremos los términos propios del segmento del turismo de aventura y, concretamente, los recogidos en los textos promocionales de dicho dominio. De esta manera, encontraremos términos especializados, propios de las diferentes actividades, y divulgativos, puesto que el público al que va dirigido es general lego en la materia. Además, se incluirán principalmente nombres, que harán referencia a los principales conceptos del dominio, y los verbos que hagan referencia a las actividades que se desarrollan. Por último, se indicará información fraseológica de dichas unidades y sinónimos, cuando sea pertinente.
- Orden de presentación. En nuestro caso, la presentación se realizará principalmente mediante una aplicación desarrollada para este proyecto que permite la búsqueda textual (cf. 2.6.1.6. y 4.3.1.) además de una búsqueda por término, por lo que la decisión ya viene determinada por la herramienta que vamos a utilizar.
- Inclusión de anexos. En todo recurso, es muy importante la descripción del recurso, especialmente su modo de uso, a fin de que los usuarios puedan sacarle el máximo partido. En nuestro caso, la aplicación de búsqueda incluirá un manual de usuario, donde se explicará brevemente los pasos que el usuario debe seguir para utilizar la herramienta de forma adecuada y todas sus funcionalidades, así como una serie de preguntas frecuentes.⁸⁸²
- Microestructura de cada entrada terminológica. Nuestro recurso incluirá los tres tipos de informaciones requeridas por los traductores profesionales, a saber: a) información lingüística (término, categoría gramatical, género, abreviaturas y acrónimos, variantes ortográficas o terminológicas y colocaciones); b) información conceptual (área temática, definición, equivalente de traducción, sinónimos, remisiones a conceptos relacionados y material gráfico), y c) información pragmática (ejemplos contextuales, marcas de uso, marcas de variación lingüística y observaciones). De esta manera, satisfará las necesidades

⁸⁸² Esta información se ha incluido en la herramienta bajo el título de *Manual de usuario* y se puede consultar en el Anexo III.

de nuestro grupo de destinatarios y posibilitará una mejor codificación del mensaje meta.

- Soporte y tipología de su edición. Como hemos indicado, nuestro recurso se consultará a través de una aplicación en línea desarrollada para esta investigación, por lo que será un recurso electrónico con acceso libre a través de Internet.

Finalmente, es imprescindible determinar el software informático con el que se cuenta. Debido a la revisión que hace falta realizar para seleccionar el adecuado, dedicaremos la siguiente sección a este tema.

3.2.1. SELECCIÓN DE LAS HERRAMIENTAS DE TRABAJO

Como indicábamos en el capítulo anterior, cada proyecto terminológico determinará las herramientas de trabajo según sus posibilidades económicas y según las necesidades a la hora de alcanzar sus resultados, por lo que no se puede determinar *a priori* de forma estricta y general las aplicaciones necesarias. En unos casos, se tendrá disponibilidad para trabajar con un conjunto de herramientas que cubran diferentes pasos del proceso terminológico; en otros casos, se podrán utilizar herramientas integradoras, es decir, una aplicación que cuente con varias herramientas integradas en la misma plataforma sin la necesidad de utilizar diferentes aplicaciones; o incluso se requerirá de un trabajo manual en algunas fases. Salvando las posibles limitaciones económicas que pueda presentar un proyecto terminológico, es importante que se elija la herramienta de trabajo que responda como mínimo al 80% de sus exigencias, como nos indican Pavel y Nolet (2001). Asimismo, si no se encontrara ninguna lo suficientemente adecuada, se podría recurrir a la creación de una propia.

Antes de seleccionar la herramienta con la que vamos a trabajar, necesitaremos plantearnos una serie de cuestiones a la hora de escoger la que más se ajusta a nuestras necesidades. Básicamente, estas cuestiones están relacionadas con los siguientes puntos:

- Capacidad. Este punto hace referencia al volumen de datos que se prevé gestionar con la herramienta, es decir, seleccionar una herramienta que

permita gestionar un volumen adecuado de información de acuerdo con el proyecto que se pretende llevar a cabo, ya que podría ocurrir que se viera limitado el trabajo debido a la falta de capacidad de la herramienta para almacenar o gestionar los datos requeridos. O, incluso, que se necesitara seleccionar otra herramienta diferente por esta limitación una vez que hubiera empezado el proyecto.

- **Acceso.** Es necesario determinar el número de personas que van a trabajar con la herramienta, el uso que van a realizar y desde dónde se va a producir. Así pues, si el trabajo se va a realizar simultáneamente por varias personas, la herramienta deberá permitir el acceso de múltiples usuarios; si se va a trabajar desde diferentes puntos, la herramienta deberá permitir un acceso remoto, etc.
- **Flexibilidad.** La herramienta debe ofrecer la flexibilidad adecuada para ajustarse a las necesidades del proyecto con respecto al número de campos terminológicos que se desean incluir en la base de datos, la selección de dichos campos, las lenguas de trabajo, etc. Es decir, la herramienta seleccionada debe ajustarse a las variables pragmático-lingüísticas que se establecen al inicio de cualquier proyecto y, o bien ofrece directamente los campos, las lenguas, etc. que se han seleccionado para el proyecto, o bien permite la creación o modificación de dicha información de forma flexible y fácil.
- **Especificaciones.** En este punto es necesario determinar si se necesitan funcionalidades específicas para el proyecto terminológico que se desea llevar a cabo, como por ejemplo la necesidad de crear ontologías, de gestionar corpus textuales, de incluir material de audio, etc. En todo caso, la selección de la herramienta deberá tener en cuenta las especificaciones concretas de cada proyecto.
- **Intercambio.** Dependiendo de los fines que se establezcan para el proyecto terminológico, este punto será relevante o no. En el caso de que se desee que la información introducida en la base de datos de la herramienta sea utilizada

con posterioridad por otras personas, proyectos o aplicaciones informáticas, se deberá tener en cuenta el formato que ofrece la herramienta para facilitar el intercambio de dicha información.

- Exportación. Relacionado con el punto anterior, la selección de la herramienta vendrá determinada por los formatos que ofrece para la fase de exportación. De esta manera, dependiendo de los objetivos marcados inicialmente se deberá seleccionar una herramienta que permita la exportación en formato .rtf, .pdf, .html, .xml, u otro.

Estos puntos indicados arriba no pretenden ser exhaustivos, aunque se pueden considerar como los puntos básicos sobre los que se deberían sustentar las decisiones a la hora de seleccionar una u otra herramienta de gestión terminológica para el trabajo que se desea llevar a cabo.

En nuestra investigación, hemos realizado una revisión de las diferentes herramientas disponibles para evaluarlas y determinar su adecuación a las necesidades de nuestro proyecto, que, según las cuestiones antes mencionadas, serían las siguientes:

- Capacidad. En un primer momento, la capacidad de la herramienta no debería ser muy grande, puesto que el trabajo se limita únicamente al segmento del turismo de aventura. No obstante, la intención de nuestra investigación es ampliar los segmentos de trabajo, por lo que sí que se requerirá una herramienta que pueda gestionar y almacenar un volumen elevado de información terminológica.
- Acceso. El acceso requerido para este trabajo debe ser remoto, puesto que se necesita disponer de la herramienta en diferentes momentos y ubicaciones, por lo que deberá disponer de un acceso a través de Internet. Asimismo, se necesita un acceso múltiple de usuarios, puesto que se planea la intervención de diferentes usuarios en un trabajo futuro.
- Flexibilidad. La herramienta debe ofrecer flexibilidad a la hora de seleccionar los campos terminológicos con los que se van a trabajar, así como permitir la selección de diferentes lenguas de trabajo. En un primer

momento, serán únicamente las lenguas español, inglés y alemán, aunque se planea ampliar el trabajo a otras lenguas como el francés y el italiano.

- Especificaciones. Las funcionalidades específicas para nuestro proyecto están basadas principalmente en la posibilidad de elaborar ontologías y de gestionar corpus textuales de gran tamaño. Asimismo, se considerarán positivamente las herramientas que permitan la inclusión de material gráfico, como imágenes, en la base de datos.
- Intercambio. La herramienta seleccionada deberá permitir el intercambio de información con otras aplicaciones, por lo que el formato de exportación básico que debe ofrecer debe ser el formato TBX (cf. 4.3.1.3.1.1.).
- Exportación. Como hemos indicado en el punto anterior, la herramienta seleccionada deberá permitir una exportación en formato TBX. Asimismo, se considerará de forma positiva la posibilidad de exportar en formatos para una edición en papel, como pueden ser formatos con extensión .rtf o .pdf.

Una vez establecidas nuestras necesidades básicas, hemos llevado a cabo un estudio de los diferentes sistemas de gestión terminológica que permiten alcanzar los objetivos de nuestro proyecto. Para ello, hemos realizado un análisis de las funcionalidades que ofrecen dichos sistemas y los hemos comparado con nuestras necesidades. Finalmente, hemos realizado la selección de la herramienta que vamos a utilizar.

Antes de empezar con la presentación de estos sistemas de gestión, consideramos necesario establecer una distinción entre los sistemas de gestión terminológica (SGT) y los sistemas de gestión ontoterminológica (SGOT) y determinar el tipo de sistema de gestión que nos interesa en el marco de nuestra investigación y, por tanto, el que vamos a utilizar en nuestro trabajo.

Por un lado, los SGT son unos gestores terminológicos muy utilizados por los traductores y por proyectos terminológicos de diferente naturaleza, ya que presentan gran cantidad de funcionalidades y ventajas de utilización. Estos gestores son programas informáticos compuestos por una base de datos que posibilita la gestión de los datos, es decir, permiten la creación, explotación y administración de bases de datos

terminológicas, así como la definición y manipulación de modelos de datos y el almacenamiento y acceso controlado de estos. Los SGT más destacados que existen actualmente en el mercado son los siguientes: *SDL Multiterm* de SDL-Trados, *SDL TermBase* de SDL y *TermStar XV* de Star, que están vinculados a memorias de traducción (aunque algunos también pueden ser utilizados de forma independiente), y *TshwaneLex Suite*, entre otros. Ninguno de estos programas son de libre distribución, pero todos permiten una versión de prueba para poder comprobar sus características y determinar si se ajustan a las necesidades del proyecto en cuestión antes de su adquisición. Asimismo, todos estos programas ofrecen gran capacidad de gestión y almacenamiento de datos, diferentes tipos de acceso (remoto/local, usuario único o múltiple), gran flexibilidad a la hora de seleccionar los campos de las entradas terminológicas y lenguas de trabajo, así como diferentes formatos de exportación y posibilidad de intercambio. Sin embargo, a pesar de todas estas ventajas que presentan, no se ajustan a las necesidades iniciales planteadas en nuestro proyecto terminológico, ya que no permiten la elaboración de ontologías ni tampoco la gestión de corpus de forma integrada.⁸⁸³ Asimismo, el intercambio de datos que proporcionan estos sistemas de gestión se limitan a usuarios que cuentan con los mismos sistemas, es decir, es posible intercambiar datos creados en los mismos sistemas de trabajo (por ejemplo, la base de datos creada por un usuario A en *SDL Multiterm* al sistema de otro usuario B), pero no entre aplicaciones diferentes, lo que limita razonablemente el intercambio de datos tal y como se defiende desde organizaciones internacionales como la ISO o LISA.

Estos SGT nos obligan a recurrir a lo que se ha denominado «bricolaje informático» (en inglés *do-it-yourself IT*) (Vargas Sierra, 2009: 45), que hace referencia al empleo de varias aplicaciones informáticas de forma simultánea y a la adaptación de estas para satisfacer las necesidades concretas del usuario, a fin de obtener los resultados deseados. Por este motivo, se recomienda el uso de otros sistemas de gestión que funcionan como estaciones de trabajo integradas a la hora de llevar a cabo cualquier trabajo terminológico de forma sistemática y completa.

⁸⁸³ El gestor *TshwaneLex Suite* permite cierta gestión de corpus mediante la búsqueda de contextos y la selección de estos para su inclusión en las diferentes entradas, aunque no permite la búsqueda de concordancias ni otras formas de gestión.

Estas estaciones de trabajo son sistemas de gestión terminológica integrales que ofrecen todos los módulos necesarios para llevar a cabo un trabajo terminográfico completo, desde la compilación del corpus de trabajo hasta la exportación de la base de datos para su posterior edición, pasando por la gestión de corpus (concordancias, colocaciones). Como ejemplo de estos sistemas integradores, podemos citar *TERMINUS*,⁸⁸⁴ del grupo IULA de la Universidad Pompeu Fabra, y *System Quirk*,⁸⁸⁵ de la Universidad de Surrey.

Los propios autores del programa *TERMINUS* (Mesa Lao, 2009: 45) lo definen como una estación de trabajo para la terminología, puesto que integra en un único entorno de trabajo todas las fases para la elaboración completa, individual o en equipo, de proyectos terminológicos monolingües con equivalencias. En otras palabras, esta herramienta permite llevar a cabo desde la compilación de un corpus textual de trabajo hasta la edición final del recurso, pasando por la gestión de corpus y el modelado de los campos que integran las entradas terminológicas. Asimismo, permite crear y gestionar diferentes perfiles de usuarios, crear y gestionar más de un proyecto terminológico, así como el trabajo simultáneo desde diferentes puntos de acceso.

Esta estación de trabajo ofrece una estructura modular, que consta de los diferentes apartados que permiten llevar a cabo el trabajo terminográfico, a saber:

- Proyectos: permite crear uno o varios proyectos terminológicos según las necesidades de los usuarios.
- Fuentes: permite gestionar las fuentes utilizadas en un proyecto terminológico.
- Estructuración conceptual: permite crear un árbol de campo para estructurar los términos de un dominio.
- Documentos: permite añadir archivos de texto en diferentes formatos que constituirán con posterioridad el corpus de trabajo.
- Corpus: permite formar corpus con los documentos incluidos en el programa o realizar búsquedas en Internet para compilarlo.

⁸⁸⁴ <<http://igraine.upf.edu/Terminus2/index.html>>

⁸⁸⁵ <<http://www.computing.surrey.ac.uk/SystemQ/>>

- Análisis: permite analizar corpus mediante frecuencias, concordancias, n-gramas y cálculo de asociación para buscar colocaciones.
- Glosarios: permite crear los glosarios que forman parte de un proyecto creado en la fase inicial por un usuario.
- Términos: permite introducir los datos terminológicos en el glosario, consultarlos y modificarlos (en su caso).
- Exportación: permite la edición de diccionarios una vez finalizado el trabajo terminológico mediante la exportación en los siguientes formatos: .html, .pdf, .txt y .xml.

En general, se observa que esta estación de trabajo constituye un recurso útil para la realización de trabajos terminológicos, puesto que presta servicios a todas las fases en las que se divide el trabajo (según la metodología propuesta por Cabré Castellví, 1993) y permite su adaptación a diferentes perfiles de usuarios.

Por su parte, *System Quirk* es también un paquete flexible e integrador de herramientas para la elaboración y gestión de bases de datos terminológicas, pero incluye algunas funcionalidades más que *TERMINUS*. Al igual que el anterior, este sistema asiste al terminólogo durante las fases del trabajo terminológico, desde la compilación del corpus textual de trabajo hasta la edición de diccionarios, y permite una configuración flexible a la hora de establecer el número de usuarios (usuario único o múltiple), el número de proyectos, los campos de la entrada terminológica, entre otros aspectos. Sin embargo, ofrece algunas diferencias que lo distinguen del anterior y lo hacen un poco más completo. Entre las principales diferencias, cabe destacar que este sistema permite la extracción automática de términos a partir de los corpus textuales cargados en la aplicación con la mínima interacción por parte del usuario.

La arquitectura de este sistema también es modular, con objeto de organizar fácilmente el trabajo y permitir mayor flexibilidad. De esta manera, encontramos diferentes programas integrados en el sistema que nos permiten llevar a cabo las tareas de cualquier trabajo terminológico:

- Virtual Corpus: un programa que permite la gestión de textos para crear corpus.
- KonText: un programa que permite analizar textos mediante la generación de listas de palabras, búsqueda de concordancias y de colocaciones, etc.
- Ferret: un programa que permite realizar análisis estadísticos de textos para localizar términos, especialmente términos compuestos, mediante el uso de stoplists.
- Browser/Refiner: un programa que permite crear y modificar bases de datos terminológicas.

Asimismo, ofrece otros módulos adicionales, como son la elaboración de resúmenes automáticos, un alineador de textos, entre otros programas, con el objetivo de completar las necesidades de los terminólogos. Por último, nos parece relevante destacar que, a pesar de ser un sistema muy completo, flexible y fácil de manejar, no permite la exportación de la base de datos terminológica, es decir, el sistema está concebido para elaborar y consultar los datos dentro del mismo programa, por lo que no permite actualmente la exportación a ningún formato.

En el marco de un trabajo terminográfico, cualquiera de estos SGT integrados serían adecuados para llevar a cabo todas las fases, desde la compilación del corpus hasta la edición de los recursos terminológicos. Sin embargo, teniendo en cuenta que nuestro trabajo se enmarca dentro de una metodología ontoterminográfica, se requeriría el empleo simultáneo de editores de ontologías independientes a la estación de trabajo (cf .1.4.3.4.) en caso de que se optara por utilizar algunos de estos SGT presentados con anterioridad, lo que daría lugar a unas constantes limitaciones por la falta de integración entre las dos metodologías.⁸⁸⁶

Por este motivo, pasamos a tratar el otro grupo de sistemas de gestión, los sistemas de gestión ontoterminológicos (SGOT), que son las aplicaciones más adecuadas para los terminógrafos que trabajan en el marco ontoterminográfico. Estas aplicaciones son las más escasas en este ámbito debido principalmente a tres razones:

⁸⁸⁶ Algunos de los SGT presentados anteriormente permiten la creación de referencias cruzadas, como TshwaneLex, o de mapas conceptuales, como TERMINUS, pero esto no sería suficiente para una metodología enmarcada en la ontoterminografía como es nuestro caso.

primero, la ontoterminografía es una línea de investigación relativamente reciente, por lo que aún quedan mucho que hacer y que investigar en este ámbito, especialmente en lo relacionado con las herramientas adecuadas para desarrollar recursos ontoterminográficos; segundo, existe una preferencia de la mayoría de los terminógrafos de trabajar fuera de un marco ontoterminográfico, en particular debido a las dificultades que se encuentran a la hora de llevar a cabo proyectos para la elaboración de estos recursos; y tercero, las herramientas que se elaboran dentro de esta metodología suelen desarrollarse dentro de proyectos de investigación y, por tanto, su empleo está restringido a un uso interno. No obstante, se pueden encontrar actualmente algunos, aunque escasos, SGOT que, sin llegar a cumplir todas las necesidades de los terminógrafos en el marco de una metodología ontoterminográfica, son herramientas que permiten la creación de este tipo de recursos aunque con ciertas limitaciones y carencias (Durán Muñoz, 2010b).

Como ejemplo de herramientas que tienen acceso libre y se encuentran en la web, presentaremos brevemente las dos siguientes: *Ontoterm*, *Terminography Tools* y *Corpógrafo*.

En primer lugar, *Ontoterm* (<http://www.ontoterm.com/>) (Moreno Ortiz, 2000, 2004) es un sistema de gestión terminológico basado en ontologías de libre acceso y dirigido a un uso académico y no comercial que permite elaborar, de forma sencilla, ontologías de nueva creación sobre cualquier dominio. Gracias a su fácil manejo y flexibilidad a la hora de permitir el trabajo en equipo con acceso remoto, esta herramienta se ha utilizado en varios proyectos terminológicos recientes en España, como en el proyecto GENOMA-KB (<http://www.iula.upf.edu/>) y el proyecto ONCOTERM (<http://www.ugr.es/~oncoterm/>).⁸⁸⁷

Esta aplicación, que se enmarca dentro de la Teoría Comunicativa de la Terminología, de Cabré (1993), está dividida en dos módulos principales: en primer lugar, un editor de ontologías, que permite la creación de ontologías desde cero o basados en ontologías previas y, por tanto, la representación del conocimiento de un dominio de especialidad en cuestión; y, en segundo lugar, un gestor de bases de datos

⁸⁸⁷ Estos proyectos se describirán al final de este capítulo.

terminológicas, que posibilita la creación y edición de información terminológica de los conceptos introducidos en la ontología, así como la navegación por la base de datos creada. Como la mayoría de los sistemas de gestión terminológica basados en ontologías, Ontoterm es un sistema orientado al concepto y, por tanto, no permite la introducción de términos o de información en la base de datos sin antes haber creado ese concepto en el gestor de ontologías, es decir, parte de un enfoque onomasiológico (del concepto al término). De esta manera, se mantiene la coherencia del sistema al ser imprescindible la creación previa de conceptos para llevar a cabo su posterior labor terminológica y, así, evitar la posible introducción de términos que no están incluidos en la ontología y que, por tanto, no se ajustan a la representación conceptual del dominio. Además de estos dos módulos principales, Ontoterm consta también de un navegador de ontologías, que permite la consulta y visualización de la ontología mediante una estructura de notaciones, y de un generador de informes en HTML, que permite al usuario seleccionar aquellos conceptos para los que desea generar páginas web, en las que aparecerá la información terminológica introducida anteriormente.

Por su parte, las categorías de datos incluidas en el gestor de bases de datos terminológicas están basadas en la norma ISO 1620 (1999), lo que le aporta estandarización y homogeneización en su contenido terminológico. Asimismo, permite que el usuario sea quien elija las categorías de datos que desea que aparezcan en las entradas terminológicas de la base de datos basadas en la norma anterior, aunque no permite que los usuarios creen categorías nuevas. Por otra parte, permite completar la información de la base de datos con material gráfico, lo que favorece la comprensión del término en cuestión y otorga mayor complejidad al trabajo final, y da la posibilidad al usuario de seleccionar las categorías de datos que necesita a la hora de utilizar la herramienta. A pesar de las ventajas y posibilidades que presenta esta herramienta, también tiene ciertas desventajas que impiden la completa satisfacción de los proyectos que la utilizan, como son la falta de apoyo técnico y de mantenimiento de la herramienta, la imposibilidad de exportar la base de datos creada con la herramienta a formatos como .rtf o .xml, así como la complejidad con la que se muestra la información contenida en la base de datos una vez que se selecciona su publicación en formato HTML (Figura 27) (cf. Durán Muñoz, 2010b).

METASTATIC-CANCER
METHOTREXATE
METHYLSALLEN
METHYLPREDNISOLONE
MICROCALCIFICATION
MICROGRAPHIC-SURGERY
MICROMETASTASIS
MICROSTAGING
MITOSIS
MITOTIC-INHIBITOR
MITOKANTRONE
MIXED-CELLULARITY
MIXED-TUMOR
MODIFIED-RADICAL-HYSTERECTOMY
MODIFIED-RADICAL-MASTECTOMY
MOHS-MICROGRAPHIC-SURGERY
MOLE
MONOCLONAL-ANTIBODY
MONOCYTE
MONOCYTIC-LEUKEMIA
MONOKINE
MONODITY
MOUTHWASH
MOVING-BEAM-THERAPY
MR-SCAN
MRI
MUCINOUS-CARCINOMA-OF-THE-BREAST
MUCOUS-MEMBRANE
MULLERIAN-TUMOR
MULTICENTRIC-DISEASE
MULTIDRUG-RESISTANCE
MULTIFOCAL-OSTEOSARCOMA
MULTIPLE-BASAL-CELL-CARCINOMA-SYNDROME
MULTIPLE-ENDOCRINE-NEOPLASIA
MULTIPLE-MYELOMA
MUSCULOSKELETAL-SYSTEM
MUSTARD-GAS
MYASTHENIA-GRAVIS
MYASTHENIC-SYNDROME
MYCOSIS-FUNGOIDES
MYELOBLASTIC-LEUKEMIA
MYELODYSPLASTIC-SYNDROME
MYELOFIBROSIS
MYELOGRAM
MYELOGRAPHY
MYELOID-LEUKEMIA

MYELOID-LEUKEMIA

originator: Chantal
origination date: 08/12/1999
check date: 27/08/2001
checker: Pamela
URL: <http://www.epi.bris.ac.uk/rd/publicat/dec/dec52.htm>

originator: Ariane
origination date: 11/01/2001

| Conceptual Structures | |
|-----------------------|---|
| ISA | MALIGNANT-NEOPLASM-OF-LYMPHATIC-AND-HEMATOPOIETIC-TISSUE |
| SUBCLASSES | ACUTE-MYELOID-LEUKEMIA CHLOROMA CHRONIC-MYELOID-LEUKEMIA MYELOID-SARCOMA SUBACUTE-MYELOID-LEUKEMIA |
| DESCENDANTS | ACUTE-MYELOID-LEUKEMIA ACUTE-PROMYELOCYTIC-LEUKEMIA ACUTE-PROMYELOMONOCYTIC-LEUKEMIA CHLOROMA CHRONIC-MYELOID-LEUKEMIA EOSINOPHILIC-LEUKEMIA MYELOBLASTIC-LEUKEMIA MYELOID-SARCOMA NEUTROPHILIC-LEUKEMIA SUBACUTE-MYELOID-LEUKEMIA |

| | |
|------------------|---|
| ANCESTORS | CHANGE-STATE-ANIMAL DISEASE-EVENT EVENT HAVE-NEOPLASM LIVING-EVENT MALIGNANT-NEOPLASM MALIGNANT-NEOPLASM-BY-BODY-PART MALIGNANT-NEOPLASM-OF-LYMPHATIC-AND-HEMATOPOIETIC-TISSUE PHYSICAL-EVENT |
| RELATIONS | AFFECTS-PHYSIOLOGICAL-SYSTEM : LYMPHATIC-SYSTEM SPLEEN DIAGNOSED-WITH : BONE-MARROW-ASPIRATION HAS-SYMPOM : FLU BECOME-TIRED ANEMIA |

English

| | |
|----------------------------|--|
| myeloid leukemia | <p>term type: main entry term</p> <p>note: Hemos asignado al término "myeloid leukemia" el termtpe "main entry term" puesto que de todas las variantes posibles es el que se usa en el ICD9-CM (International Classification of Diseases, Clinical Modification de 1998). (es)</p> <p>part of speech: noun number: singular hyphenation: my-e-loid leu-ke-mi-a usage note: usually further subclassified as acute myeloid leukemia and chronic myeloid leukaer (en)</p> <p>context: By calorie restriction, the incidence of myeloid leukemia was significantly decreased. In addition, the latent period of the myeloid leukemia in the groups for calorie restriction was significantly extended at a greater extent as compared with the control diet groups. (en)</p> <p>source identifier: Harrison's Plus</p> <p>definition: leukemia characterized by an uncontrolled proliferation of the myeloid lineage and precursors in the bone marrow and other sites (UMLS). (en)</p> <p>context (x): i myeleu.txt</p> <p>origination date: 12/06/2002 originator: Ariane</p> |
| myelocytic leukemia | <p>term type: synonym part of speech: noun number: singular usage note: usually further subclassified as acute myelocytic leukemia and chronic myelocytic leukemia (en)</p> <p>frequency: commonly used context: Alkylating agents have been shown to cause an increased incidence of myelocytic leu and probably other malignancies. (en)</p> <p>source identifier: Harrison's Plus</p> <p>definition: see myeloid leukemia. (en)</p> |
| myelosis leukemia | <p>term type: synonym part of speech: noun number: singular geographical usage: en US (American English) (en)</p> <p>definition: see myeloid leukemia. (en)</p> |
| myelogenous | <p>term type: synonym part of speech: noun</p> |

MICROMETASTASIS
MICROSTAGING
MITOSIS
MITOTIC-INHIBITOR
MITOKANTRONE
MIXED-CELLULARITY
MIXED-TUMOR
MODIFIED-RADICAL-HYSTERECTOMY
MODIFIED-RADICAL-MASTECTOMY
MOHS-MICROGRAPHIC-SURGERY
MOLE
MONOCLONAL-ANTIBODY
MONOCYTE
MONOCYTIC-LEUKEMIA
MONOKINE
MONODITY
MOUTHWASH
MOVING-BEAM-THERAPY
MR-SCAN
MRI
MUCINOUS-CARCINOMA-OF-THE-BREAST
MUCOUS-MEMBRANE
MULLERIAN-TUMOR
MULTICENTRIC-DISEASE
MULTIDRUG-RESISTANCE
MULTIFOCAL-OSTEOSARCOMA
MULTIPLE-BASAL-CELL-CARCINOMA-SYNDROME
MULTIPLE-ENDOCRINE-NEOPLASIA
MULTIPLE-MYELOMA
MUSCULOSKELETAL-SYSTEM
MUSTARD-GAS
MYASTHENIA-GRAVIS
MYASTHENIC-SYNDROME
MYCOSIS-FUNGOIDES
MYELOBLASTIC-LEUKEMIA
MYELODYSPLASTIC-SYNDROME
MYELOFIBROSIS
MYELOGRAM
MYELOGRAPHY
MYELOID-LEUKEMIA
MYELOMA
MYELOPROLIFERATIVE-DISORDER
MYELOSUPPRESSION-CHEMOTHERAPY
MYELOSUPPRESSION-SIDE-EFFECT
MYELOSUPPRESSION
MYOMA
MYXOMA

N

N-VALUE
NASOGASTRIC-TUBE
NASOPHARYNGOSCOPE
NASOPHARYNGOSCOPY
NASOPHARYNX
NATURAL-KILLER-CELL
NEAR-TOTAL-THYROIDECTOMY
NECRODISSECTION
NEEDLE-ASPIRATION-BIOPSY
NEEDLE-LOCALIZATION
NEOADJUVANT-THERAPY
NEPHRECTOMY
NEPHROSTOMY
NEPHROTOMOGRAPHY
NEPHROTOMOGRAPHY
NEPHROURETERECTOMY
NERVE-BLOCK
NERVE-ROOT-CLIPPING

Figura 27. Resultado para la búsqueda del término «Myeloyd-Leukemia» en la base de datos Oncoterm.
Fuente: <<http://www.ugr.es/~oncoterm/alpha-index.html>>.

La segunda herramienta que vamos a comentar en este apartado es *Termonotography Tools* (<http://taalkunde.ehb.be/cvc/software>), que se enmarca dentro de la Teoría Sociocognitiva de la Terminología, de Temmerman (2000). Esta aplicación se trata de un conjunto de tres herramientas vinculadas entre sí cuya finalidad es la creación de recursos ontoterminográficos. Estos programas están desarrollados en JAVA y fueron creados en el seno del proyecto europeo FFPOIROT (de Baer et ál. 2006),⁸⁸⁸ en el Centrum voor Vaktaal en Communicatie (Erasmushogeschool, de Bruselas), bajo la dirección de la profesora R. Temmerman. En las primeras fases de su desarrollo, estas herramientas estaban basadas en el editor de ontologías *Protégé*, pero posteriormente fueron modificadas y ampliadas, y adquirieron su propia estructura independiente.

Estas tres herramientas dividen la creación de los recursos ontoterminográficos en tres fases, una por cada herramienta: en primer lugar, se realiza la conceptualización del dominio, es decir, se establece la ontología de dominio; en segundo lugar, se completa la información terminológica en la base de datos a partir de textos y de la representación conceptual elaborada previamente; y, por último, se consultan los términos incluidos en la base de datos elaborada previamente. Como se puede observar con esta metodología, esta herramienta se trata también de una aplicación orientada al concepto, es decir, sigue un enfoque onomasiológico.⁸⁸⁹

En la primera fase se utiliza el programa *Categorisation Framework Editor*, cuyo objetivo es elaborar un marco predefinido de conceptos (o categorías, según Temmerman, 2000) y de relaciones, es decir, esta fase consiste en la creación de una conceptualización (o categorización, según Temmerman, 2000) del dominio de especialidad en cuestión. De esta forma, esta fase es independiente del lenguaje objeto

⁸⁸⁸ El proyecto FFPOIROT (Financial Fraud Prevention Oriented Information Resources using Ontology Technology) se enmarcó en el V Programa Marco de la Unión Europea y tuvo una duración de tres años (2002-2005). Tenía como principal objetivo la elaboración de una base de conocimiento en varios idiomas (holandés, italiano, francés e inglés) para el dominio del fraude fiscal en Internet.

⁸⁸⁹ No obstante, la metodología propuesta por Kerreman y Temmerman (2003) se trata de una metodología mixta, es decir, va desde el concepto al término pero también desde el término al concepto, por lo que el enfoque no sería íntegramente conceptual como se concluye de las herramientas que utilizan. Como se comentó en el capítulo anterior (cf. 2.6.), esta metodología defiende una constante alimentación de la información, tanto a partir de textos (de términos) como a partir de los conceptos propios del dominio.

de estudio y trabaja solo en el nivel conceptual, organizando y representando el conocimiento del campo de especialidad, según los conocimientos del propio terminólogo y a través de las consultas a expertos en la materia. Este editor de ontologías es muy sencillo de manejar y permite gran flexibilidad a la hora de crear, mover, eliminar o modificar conceptos y relaciones. Esto se debe principalmente a que se trata de una herramienta que se ha diseñado por y para los terminógrafos, por lo que se parte de la premisa de que un terminógrafo o lingüista en general cuenta con un conocimiento muy limitado de todo lo relacionado con la ingeniería del conocimiento y la informática y, por tanto, cuanto más fácil y manejable sea la herramienta que deba utilizar, mejor realizará su trabajo. De esta manera, es una herramienta bastante intuitiva y fácil de usar, basada como se puede ver en el editor de ontologías *Protégé*, a pesar de contar con algunas dificultades y limitaciones. Entre los principales problemas que encontramos en esta herramienta es el hecho de establecer, por defecto, relaciones ontológicas, es decir, relaciones parte-todo, siempre que se crean conceptos vinculados entre sí. De esta manera, obliga al terminólogo a modificar todas las relaciones que se establecen de forma automática en la aplicación, lo que provoca una pérdida grande de tiempo. Por otro lado, establece una separación entre meta-categorías y categorías que no es adecuada en este nivel, puesto que, al encontramos en el nivel conceptual, todas las entradas de las ontologías deberían considerarse como categorías. Y, de nuevo, encontramos aquí otra dificultad para el uso de la herramienta.

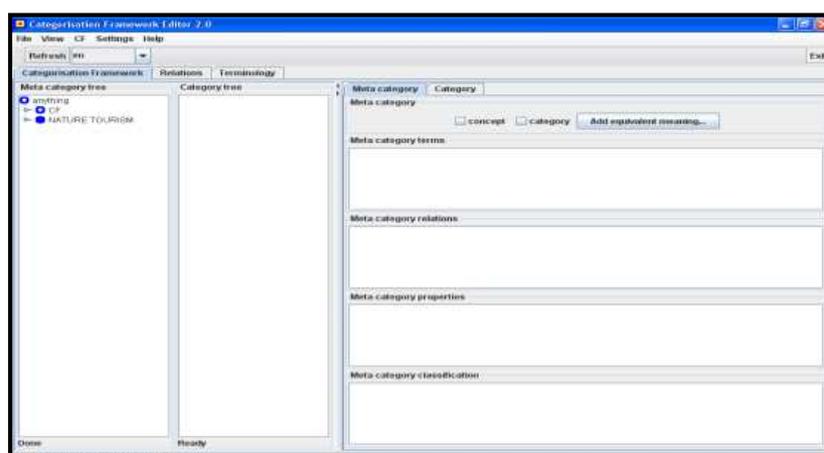


Figura 28. Pantalla inicial del programa *Categorisation Framework Editor*.

Una vez que se ha representado el dominio mediante sus conceptos, es decir, se ha creado la ontología del dominio, se pasa al nivel terminológico con la siguiente herramienta, *Termonography Workbench*, para empezar a trabajar en la base de datos terminológica. En esta herramienta, se cargan los textos que comprenden el corpus textual del dominio de trabajo que se ha compilado previamente, estableciendo durante la subida una organización de los textos por dominio o por criterios establecidos *ad hoc*, y se comienza a extraer la terminología de los textos y a incluirlas en la categorización realizada con la primera herramienta. Esta segunda herramienta también es muy fácil de manejar e intuitiva y, además, tiene una peculiaridad muy importante: el trabajo terminográfico se basa directamente en textos; es decir, la parte básica del trabajo terminográfico, que es la compilación de un corpus textual adecuado, se integra en la herramienta para que la terminología se extraiga directamente a partir de la visualización de los textos y, por tanto, de su contexto real. Esta funcionalidad de la herramienta presenta cierto interés e innovación con respecto a otras herramientas, ya que permite el tratamiento del texto directamente por parte del terminólogo en su contexto real. Sin embargo, este proceso obliga a desarrollar un trabajo laborioso y extenso, puesto que se deben consultar todos los textos compilados, que se han tenido que subir a la aplicación de uno en uno, para extraer posteriormente la terminología de forma manual, en vez de utilizar extractores terminológicos para extraerla de forma (semi)automática.

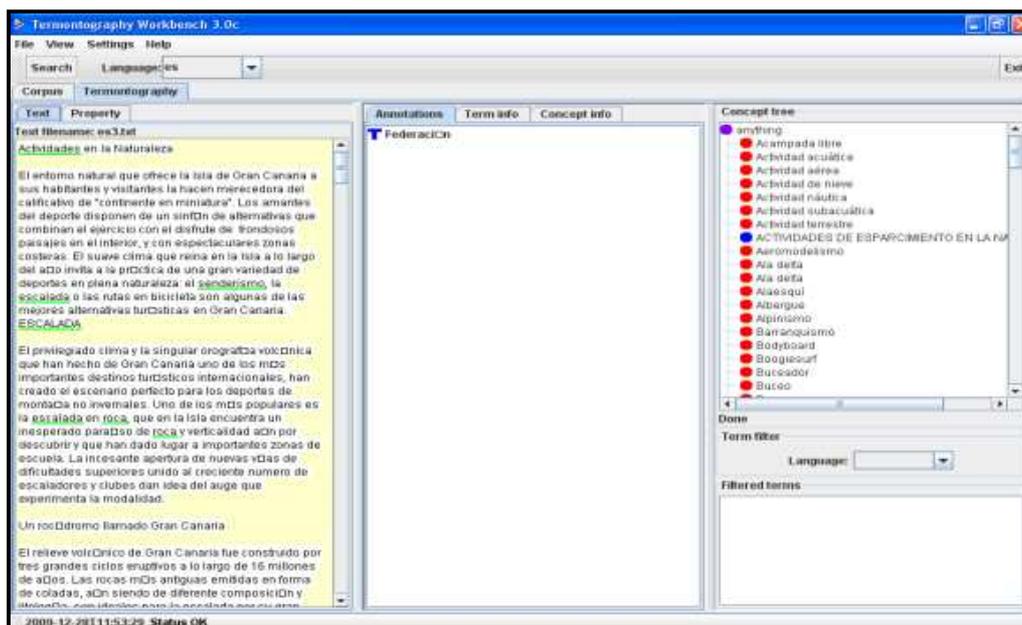


Figura 29. Pantalla del programa *Termtography Workbench* con texto (izquierda) y categorización (derecha) cargados.

Por último, encontramos la herramienta *Dictionary*, que carga automáticamente los datos incluidos en las herramientas anteriores y presenta como resultado un diccionario con diferentes opciones de búsqueda (por orden alfabético, por término, por término con comodines, es decir, lo que se conoce como búsqueda aproximada *-fuzzy search-*). Esta herramienta es bastante limitada, puesto que no incluye toda la información que se ha ido añadiendo en las herramientas anteriores, sino simplemente la definición, los términos relacionados con la entrada que se ha consultado y los datos de la fuente y el idioma. A pesar de contar con bastantes fortalezas, como son los diferentes tipos de búsqueda que dispone, la carga automática de datos introducidos anteriormente, así como el fácil manejo y la correcta visualización de los datos, su uso se ve limitado al no permitir la selección de información que debe aparecer en cada entrada, las relaciones establecidas anteriormente, los conceptos relacionados u otro tipo de información de interés por parte del usuario.

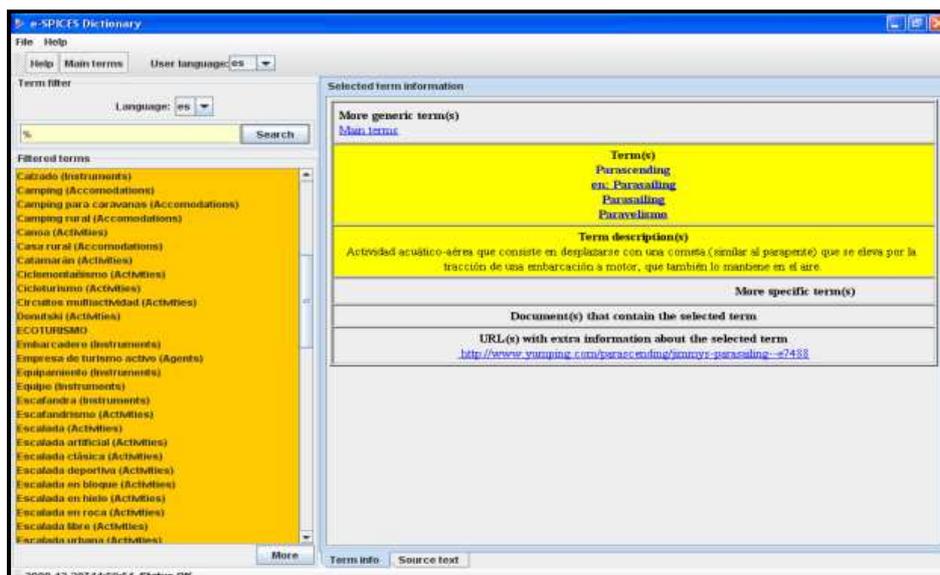


Figura 30. Pantalla del programa *Dictionary*.

En general, podemos decir que los dos programas comentados con anterioridad, ambos de libre acceso para uso académico presentan grandes fortalezas en su diseño y manejo; sin embargo, sus debilidades hacen que su uso y extensión sean limitados, especialmente con respecto a *Termtography Workbench*, que necesitaría la revisión de algunas de sus funcionalidades para conseguir una herramienta más útil en el campo de la ontoterminografía. En nuestra opinión, si estas debilidades comentadas aquí se subsanaran, ambas herramientas serían óptimas para llevar a cabo un trabajo ontoterminográfico completo, pero, actualmente, consideramos que ninguna de las dos son adecuadas para nuestros propósitos.

A continuación, trataremos una aplicación denominada *Corpógrafo* (<http://www.linguateca.pt/corpografo/>), que, al igual que las herramientas presentadas con anterioridad, es una aplicación de libre distribución desarrollada en el ámbito académico, en este caso la Universidad de Oporto, con el objetivo de apoyar el trabajo de los investigadores lingüísticos, especialmente terminólogos. A diferencia de las dos anteriores, que presentaban una orientación conceptual, esta herramienta está orientada al término y, por tanto, no requiere como punto de partida la representación conceptual de un dominio mediante ontologías. Además, consiste en un conjunto de herramientas integradas en una misma aplicación que asisten al terminólogo durante el proceso de

elaboración de una base de datos partiendo de un corpus textual que contenga documentos relativos a un dominio concreto, lo que presenta cierta similitud con el programa *TERMINUS*, que comentábamos con anterioridad, o con el prototipo *Spaterm*, que comentaremos más adelante. No obstante, debemos destacar que se trata de un sistema que permite el trabajo semántico, por lo que no seguiría la misma línea que *TERMINUS*, que era una estación integrada para el trabajo terminográfico sin ontologías.

Corpógrafo incluye gran variedad de herramientas que van desde conversores de formato a buscadores de concordancias y extractores semiautomáticos de relaciones semánticas, pasando por herramientas que permiten a los usuarios crear sus propias bases de datos, lo que la convierte en una aplicación muy completa. Su estructura está dividida en cuatro grandes áreas de trabajo, a saber (Figura 31):

- El gestor (*Gestor*), donde se encuentran las herramientas de edición y preprocesamiento de los archivos, como son las opciones para etiquetar, segmentar, limpiar o convertir el formato de textos, así como para realizar comparaciones entre corpus;
- El analizador (*Pesquisa*), que permite la extracción automática de candidatos a término mediante n-gramas, extracción de colocaciones y búsqueda de concordancias;
- El centro de conocimiento (*Centro de Conhecimento*), donde encuentran las herramientas que permiten generar y organizar el conocimiento, ya sea para llevar a cabo análisis específicamente relacionados con información conceptual o para detectar relaciones semánticas entre términos;
- El centro de comunicación (*Centro de Comunicação*), donde se puede encontrar toda la documentación existente sobre la aplicación, enviar mensajes al administrador del sistema y comprobar los mensajes recibidos.

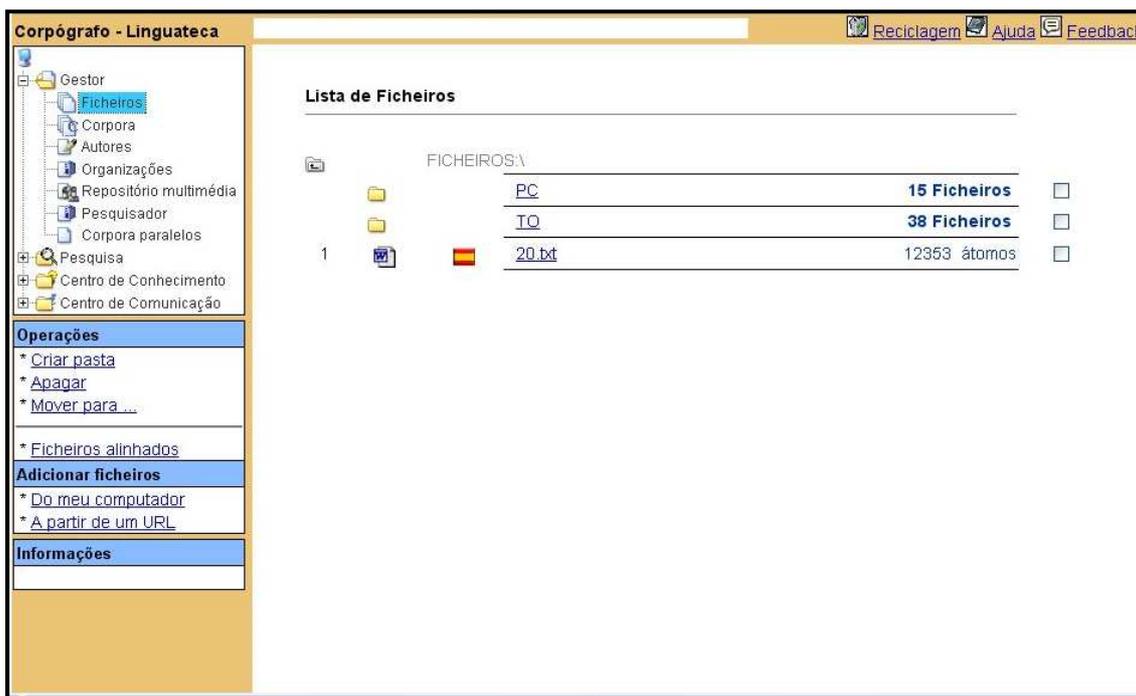


Figura 31. Pantalla del programa *Corpógrafo*.

Estos cuatro módulos que comentábamos se pueden observar en la columna izquierda de la pantalla, donde al seleccionar en cada apartado se abre un menú que permite al usuario utilizar las diferentes herramientas disponibles de cada apartado, además de aparecer en la pantalla central la información relativa a dicha opción una vez que se pincha en ella. Junto a estas opciones, el programa también dispone de otras funcionalidades, como son la de alinear textos, añadir documentos de forma individual, mover archivos, etc., que se pueden acceder en la parte baja de la misma columna.

Asimismo, como decíamos, permite la creación de recursos terminológicos basados en el conocimiento, como glosarios especializados, tesauros, bases de datos terminológicas y ontologías, así como listas de palabras categorizadas. Por último, esta aplicación también permite la exportación de los resultados a varios formatos y aplicaciones, siguiendo los estándares terminológicos de bases de datos (XML) y de memorias de traducción.

A pesar de presentar gran número de ventajas para el trabajo ontoterminográfico que pretendemos llevar a cabo en este proyecto gracias a las funcionalidades de las herramientas que integran el sistema, su uso se ve limitado por completo al trabajo

terminológico en lengua portuguesa. En otras palabras, el sistema se ha desarrollado por y para investigadores portugueses, lo que hace que todo el software esté en portugués, los campos de la entrada terminológica estén en esa lengua, así como todos los apartados automáticos que permite la herramienta (extracción de relaciones semánticas, de colocaciones, etc.) están basados en la lengua portuguesa. Todo esto nos permite utilizar la herramienta para observar su funcionamiento, pero lamentablemente no utilizarla para nuestro proyecto.

Llegados a este punto, y después de haber presentado tres de los programas de libre acceso para la creación de recursos ontoterminográficos y de haber comentado sus ventajas e inconvenientes, trataremos brevemente otras herramientas basadas en la misma concepción pero con acceso restringido al grupo de investigación en cuestión: por un lado, el conjunto de herramientas desarrollado para los investigadores del grupo Condillac, de la Universidad de Saboya (<http://ontology.univ-savoie.fr>); y por otro, el prototipo *Spaterm*, desarrollado en el marco del proyecto *La contratación turística electrónica multilingüe como mediación intercultural: aspectos legales, traductológicos y terminológicos* (N.º de ref. HUM-892, 2006-2009. Junta de Andalucía), dirigido por la Dra. Gloria Corpas Pastor, y en el que hemos tenido la oportunidad de trabajar como investigadora.

El grupo de investigación Condillac, dirigido por el profesor C. Roche, utiliza tres herramientas diferentes para la creación de recursos ontoterminográficos, de la misma manera que lo hacía *Terminography Tools*. La primera, llamada *iMap* (o SNCW de Semantic Network Craft Workbench), permite la creación de conceptualizaciones semiformales, es decir, redes semánticas o mapas conceptuales. Esta fase utiliza una aplicación orientada al concepto y se lleva a cabo conjuntamente con expertos del dominio de especialidad en cuestión, que son los que se encargan de representar los conceptos y las relaciones que existen entre ellos. Esta herramienta muestra la concepción básica de su metodología, la ontoterminología (Roche, 2007), que consiste en establecer que el trabajo ontoterminológico debe partir de los conceptos

del dominio y no de los corpus textuales, donde la información puede aparecer distorsionada.⁸⁹⁰

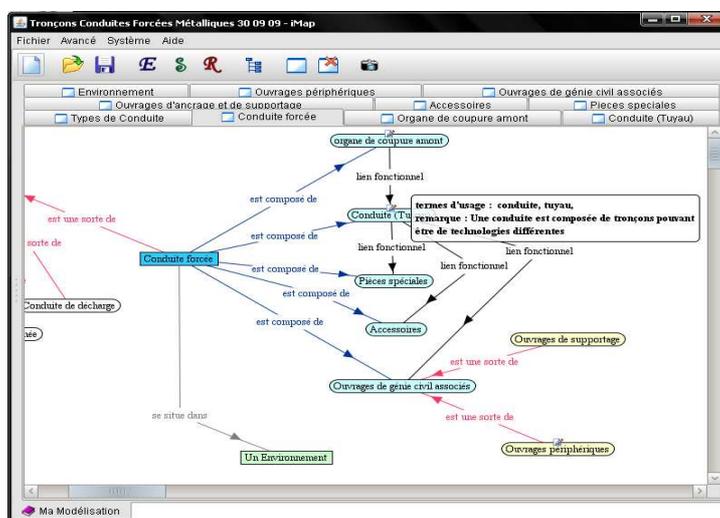


Figura 32. Pantalla del programa *iMap*.

Después de esta primera herramienta, que presenta ciertas limitaciones a la hora de introducir la información semántica, debido a la creación de relaciones y de tamaño de redes, se continúa con la segunda, *iConcept* (o OCW de Ontology Craft Workbench), que está destinada a la creación de ontologías propiamente dichas. Además, en esta fase se introduce la información lingüística y se establecen las relaciones entre términos y conceptos (establecidas con anterioridad).

⁸⁹⁰ Las demás herramientas expuestas en este apartado también consideran el concepto como punto de partida, aunque defienden una metodología mixta (*middle-up*), que utiliza los corpus textuales como fuente de información además del trabajo conjunto con expertos en la materia en cuestión para la elaboración de las ontologías.

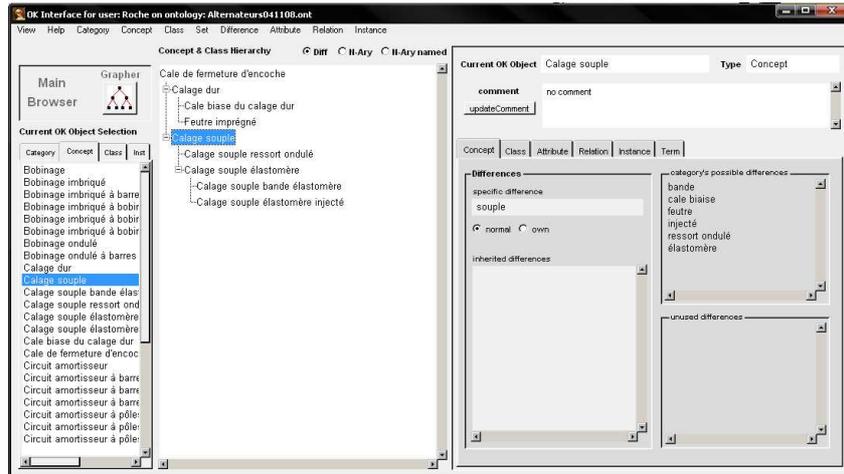


Figura 33. Pantalla del programa *iConcept*.

La última herramienta permite la visualización y la consulta de toda la información que se ha ido incluyendo en las fases anteriores en forma de diccionario especializado. De esta manera, se puede acceder a la información semántica y lingüística que se ha ido incluyendo en las fases previas, así como a las imágenes y esquemas que se han ido creando.

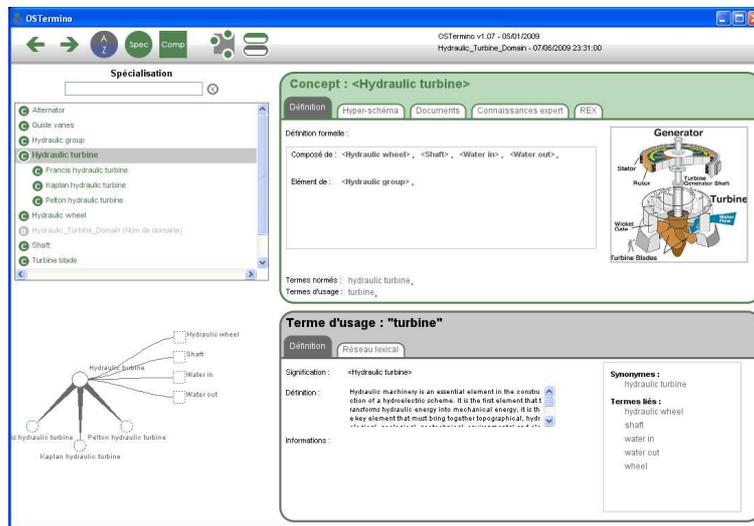


Figura 34. Pantalla del diccionario especializado como resultado de las fases anteriores.

Estas tres herramientas del grupo Condillac comparten la misma estructura y esencia que las de *Termtography Tools*, aunque el resultado en forma de diccionario

especializado es mucho más completo y pertinente, y presenta un menor número de limitaciones. En general, salvando las concepciones en lo referente a la metodología defendida por este grupo de investigación, nos parece que estas herramientas son las más completas que hemos examinado hasta el momento, y su empleo sería muy adecuado para nuestro trabajo de investigación. Sin embargo, se trata de unas herramientas de acceso restringido para la Universidad de Saboya, lo que nos impide su utilización en nuestro proyecto.

Por último, comentaremos la herramienta *Spaterm*,⁸⁹¹ que como el sistema de gestión anterior también se trata de una aplicación de acceso restringido, ya que ha sido elaborada dentro del proyecto de investigación HUM-892, de la Universidad de Málaga. Esta aplicación es una herramienta integrada de gestión y edición de recursos ontoterminográficos que permite la ejecución de todas las fases de trabajo en el marco de una metodología ontoterminográfica (cf. 2.6.1.), a saber: gestión de corpus, búsqueda de concordancias, extracción de candidatos a término y a equivalentes de traducción (mediante cognados), edición de ontologías, elaboración de entradas terminológicas y edición en formato .rtf, .pdf y .html. Asimismo, permite el acceso tanto de usuario único como múltiple, la elaboración simultánea de diferentes proyectos, así como el acceso remoto.

Esta aplicación está diseñada para trabajar con las lenguas inglés, español, francés e italiano y está orientada al término, es decir, trabaja con una metodología semasiológica. Presenta una interfaz sencilla e intuitiva y una estructura modular, formada por los diferentes apartados que permiten al terminólogo llevar a cabo las fases del trabajo ontoterminográfico. Estos apartados están organizados en dos niveles: por un lado, nos encontramos el primer nivel, donde están los tres módulos principales en los que se divide la herramienta: *corpus* (gestor de corpus), *term bank* (gestor de base de datos) y *knowledge patterns* (gestor de patrones semánticos). Y, por otro lado, observamos el segundo nivel, que se alcanza seleccionando uno de los tres gestores de forma independiente. De esta manera, pinchando en cada módulo del primer nivel

⁸⁹¹ En Castillo Rodríguez (2010) se realiza una descripción de un trabajo ontoterminográfico con el programa *Spaterm* para la terminología del sector del turismo de salud y belleza. Se puede acceder a esta herramienta a través de la siguiente URL: < <http://clg.wlv.ac.uk/MCQDemo/login.php>>.

(*corpus*, *term bank* y *knowledge patterns*), accederemos a las diferentes herramientas que nos permitirán llevar a cabo nuestro trabajo que comentaremos a continuación.

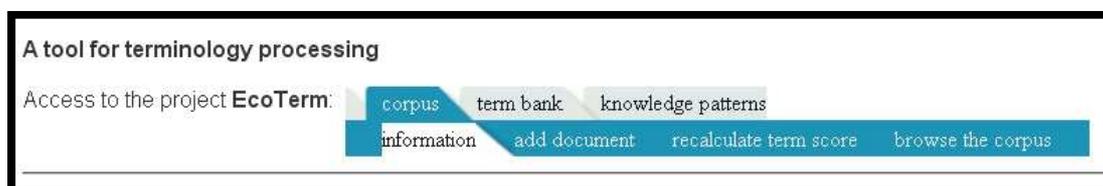


Figura 35. Interfaz de *Spaterm* con los tres módulos del primer nivel y los módulos del segundo nivel del gestor de corpus.

Dentro del gestor de corpus, encontramos cuatro secciones diferentes que nos dan la posibilidad de manipular, organizar y analizar el corpus que se vaya a utilizar para el trabajo terminológico. En primer lugar, este módulo nos da la posibilidad de subir un corpus textual formado por varios archivos en formato de texto plano (.txt) en un archivo comprimido, así como la subida individual de documentos en el mismo formato de texto plano, ya sea para completar el corpus anterior o para crear un corpus en la misma aplicación. Una vez que los documentos de trabajo se encuentran en la aplicación, este módulo nos permite visualizar los documentos cargados, eliminar documentos irrelevantes para el proyecto y consultar la información general referente al corpus, como por ejemplo el número de documentos subidos, el número de *tokens* total, el número de documentos por lengua, etc. Por último, este módulo de la herramienta permite extraer de forma automática los candidatos a términos de los textos disponibles mediante una técnica híbrida, es decir, una técnica que aúna las características de la extracción estadística y lingüística (cf. 1.4.2.4.; Ha ét al, 2008a y 2008b), de dos formas diferentes: una, el usuario puede seleccionar el número mínimo de candidatos a término que desea visualizar por lengua; y dos, el usuario realiza una selección de ocurrencias mínimas del candidato en el texto.

Una vez que se ha cargado el corpus de trabajo en la aplicación y se ha realizado la extracción de los candidatos a término, se pasa al siguiente módulo del primer nivel: el gestor de base de datos. Este módulo, al igual que el anterior, está formado por una serie de submódulos que contienen las herramientas de trabajo del terminólogo y que se pueden observar en la Figura 36.

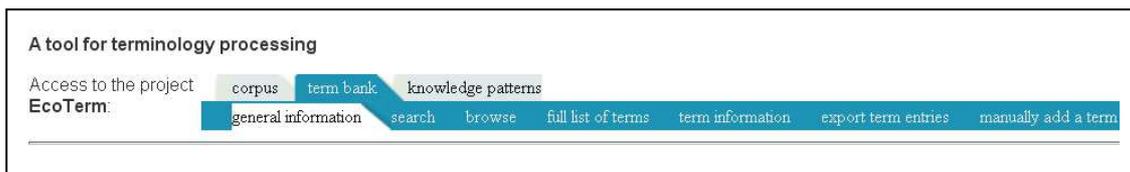


Figura 36. Módulo y submódulos del editor de base de datos de *Spaterm*.

Así pues, encontramos un submódulo que permite consultar la información general contenida en la base de datos (número de candidatos a términos extraídos, el número de términos validados, el número de términos eliminados por el usuario, etc.), así como realizar búsquedas de términos incluidos en la base de datos, consultar la lista de todos los candidatos a término extraídos organizados por lengua, validar los términos y eliminar los que no se consideran términos, entre otras funcionalidades generales. En este mismo módulo, el terminógrafo puede, además, comenzar a introducir la información terminológica pertinente para cada término validado a partir de una plantilla que contiene las categorías establecidas por la norma ISO 12620 (1999) más utilizadas en los recursos ontoterminológicos, como son 1) unidad terminológica; 2) categoría gramatical; 3) género, 4) número; 5) estado del término; 6) entidad estandarizadora; 7) uso geográfico; 8) dominio; 9) contexto; 10) definición; 11) material gráfico; 12) equivalente(s); 13) tipo de término; y 14) observaciones. Asimismo, en caso de que el usuario detectara carencias en la extracción realizada de forma automática por la aplicación, tendría la posibilidad de introducir términos de forma manual en otro de los submódulos disponibles (*Manually add a term*). Por último, dentro de este nivel, encontramos también un submódulo para la exportación de entradas terminológicas (Figura 37), donde se procede a seleccionar los campos que se desean exportar y el formato (HTML, PDF y RTF) en el que se desea realizar dicha exportación.

| select fields | |
|---------------------|---|
| terminological unit | <input checked="" type="checkbox"/> |
| term type | <input checked="" type="checkbox"/> Abbreviation <input type="checkbox"/> <input type="checkbox"/> Acronym <input type="checkbox"/> <input checked="" type="checkbox"/> Full form <input checked="" type="checkbox"/> <input type="checkbox"/> Scientific term <input type="checkbox"/> <input type="checkbox"/> Symbol <input type="checkbox"/> <input checked="" type="checkbox"/> Synonym <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Orthographical variant <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Terminological variant <input checked="" type="checkbox"/> |
| part of Speech | <input checked="" type="checkbox"/> Noun <input checked="" type="checkbox"/> <input type="checkbox"/> Verb <input type="checkbox"/> <input type="checkbox"/> Adjective <input type="checkbox"/> <input type="checkbox"/> Adverb <input type="checkbox"/> <input type="checkbox"/> Other <input type="checkbox"/> |
| gender | <input checked="" type="checkbox"/> Masculine <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Feminine <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Neuter <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Other <input checked="" type="checkbox"/> |

Figura 37. Ejemplo de selección de campos para la exportación.

Por último, encontramos el módulo de gestor de patrones semánticos, que se encuentra también en el primer nivel de la aplicación. Este módulo permite al usuario cargar en la aplicación patrones semánticos pertinentes para las diferentes relaciones que se hayan determinado relevantes para un dominio en cuestión y, a partir de ellos, detectar de forma automática las relaciones semánticas entre términos presentes en el corpus de trabajo. Asimismo, partiendo de los resultados obtenidos, la aplicación también ofrece una representación gráfica de forma automática (Figura 38), además de la posibilidad de realizar modificaciones, inclusiones y eliminaciones de dichos resultados (Figura 39).



Figura 38. Ejemplo de representación gráfica de relaciones semánticas (Castillo Rodríguez, 2010: 579).

| suspected relations | |
|--|---|
| suspected relations | <input type="button" value="delete"/> |
| aromaterapia HOWIT WORK consiste en el uso de aceites esenciales | <input type="checkbox"/> |
| aromaterapia ISA masaje corporal | <input type="checkbox"/> |
| aromaterapia <input type="text" value="relation type (ISA, WOR) relation with"/> | <input type="button" value="add more.."/> |

Figura 39. Ventana para la creación, eliminación o modificación de relaciones semánticas ofrecidas de forma automática por *Spaterm* (Castillo Rodríguez, 2010: 579).

Además de los módulos que conforman el primer y segundo nivel de esta aplicación, encontramos otras funcionalidades incluidas dentro de la plantilla para la creación de entradas terminológicas que son de gran ayuda al terminólogo durante el proceso de trabajo. Por un lado, encontramos el buscador de concordancias, que nos permite analizar los contextos en los que aparece la entrada terminológica sobre la que estamos trabajando para observar sus concordancias y poder seleccionar el contexto más adecuado para la entrada. Y, por otro, encontramos un módulo de análisis por el que se sugieren candidatos a variantes ortográficas y terminológicas del término con el que se está trabajando, así como candidatos a equivalentes de traducción a partir de cognados (Mitkov et ál., 2008). En ambos casos, el usuario puede validar o eliminar las

propuestas de la aplicación, así como añadir sus propios equivalentes en caso de no estar de acuerdo con los arrojados de forma automática por la aplicación.

Llegados a este punto, y después de haber realizado una revisión de las herramientas ontoterminográficas a las que hemos tenido acceso y después de haber evaluado sus ventajas e inconvenientes, podemos confirmar que el prototipo *Spaterm* es una de las herramientas más completas disponibles en la actualidad para llevar a cabo un trabajo ontoterminográfico (cf. 2.6.1.) de forma fácil, sistemática e integral. Y, por ello, la seleccionaremos como la aplicación que utilizaremos para llevar a cabo nuestra elaboración de base de datos ontoterminográfica, gracias a la posibilidad de acceso por nuestra pertenencia como investigadora al proyecto *La contratación turística electrónica multilingüe como mediación intercultural: aspectos legales, traductológicos y terminólogos* (N.º de ref. HUM-892, 2006-2009. Junta de Andalucía), dirigido por la Dra. Gloria Corpas Pastor, en el marco del cual se ha desarrollado dicho prototipo. No obstante, cabe destacar que este prototipo cumple la mayoría de los requisitos establecidos al inicio de nuestro proyecto, a saber: 1. la capacidad en la gestión de grandes volúmenes de datos y de documentos, 2. las posibilidad de acceso por usuario múltiple y remoto, 3. la flexibilidad a la hora de seleccionar campos terminológicos que se desean exportar y 4. la creación de ontologías. Sin embargo, existen otros requisitos que no se ven cumplidos, como son la flexibilidad a la hora de seleccionar las lenguas (en *Spaterm* solo están disponibles las lenguas inglés, francés, español e italiano) y el intercambio de datos mediante la exportación de bases de datos en un formato estándar.⁸⁹²

En nuestro proyecto, como decíamos, utilizaremos como base el prototipo *Spaterm*, pero realizaremos una serie de modificaciones para subsanar los inconvenientes de la versión actual que no se ajustan a nuestros objetivos iniciales, así como para ampliar sus funcionalidades e independencia del equipo de programadores que la ha desarrollado. Todo ello se llevará a cabo dentro del proyecto I+D *Ecosistema: Espacio único de sistemas de información ontológica y tesauros sobre el medio ambiente* (N.º de ref. FFI2008-06080-C03-01/FILO, 2009-2011. Ministerio de Ciencia e

⁸⁹² Recordemos que el formato estándar recomendado por la ISO es TBX y, en un marco de empresas de traducción y recursos terminológicos, el formato TBX-Basic.

Innovación), dirigido por la Dra. Gloria Corpas Pastor, y dará lugar a la aplicación informática *ProTermino*.

En primer lugar, se añadirá el idioma alemán como otra lengua de trabajo junto con las ya disponibles en *Spaterm* (español, inglés, francés e italiano). De esta manera, se podrá llevar a cabo el trabajo ontoterminográfico completo, incluido tanto las tareas automatizadas de la aplicación (extracción de candidatos a términos y equivalentes, la creación de ontologías, etc.) como las tareas manuales (introducción de datos terminológicos, inclusión manual de términos, etc.).

En segundo lugar, se incorporará la opción de exportar la base de datos elaborada con la aplicación en formato TBX (concretamente TBX-Basic), para posibilitar el intercambio de datos entre aplicaciones, usuarios o empresas, lo que proporciona un carácter internacional o más extenso a la herramienta.

En tercer lugar, se realizarán una serie de cambios menores para mejorar el contenido de la aplicación, concretamente en la sección de introducción de datos en la entrada terminológica. A continuación, se indican brevemente algunos de ellos:

- Se ha incorporado el campo de colocaciones, que no estaba disponible con anterioridad y que puede ser de gran utilidad dependiendo del grupo de destinatarios al que vaya dirigido el recurso que se está elaborando;
- Se ha automatizado la inclusión del contexto que aparece en la entrada terminológica de un término a partir del buscador de concordancias. En la versión anterior, el usuario debía escribir manualmente el contexto que había seleccionado en el módulo de concordancias, lo que provocaba falta de agilidad al trabajar con la herramienta. Actualmente, es posible que el usuario seleccione el contexto que le parezca más oportuno y que, directamente seleccionándolo en la ventana del buscador, aparezca en el campo *Contexto* de la entrada terminológica.
- Se ha establecido la posibilidad de incluir varios equivalentes de un mismo término, ya que en la versión anterior solo se permitía añadir un único equivalente por lengua. Además, se puede indicar el género de cada equivalente, que puede resultar muy útil en lenguas como el alemán.

Por último, también se ha introducido una funcionalidad para agilizar el trabajo del terminólogo y que le permite indicar al usuario el estado en el que se encuentra una entrada en la que ha estado trabajando. De esta manera, le da la opción de señalar si una entrada está ya finalizada, si le falta la revisión, si está sin terminar, etc. para agilizar la continuación de su trabajo y evitar, así, que tenga que llevar un recuento externo de las entradas en las que ha trabajado o que tenga que acceder a la entrada para consultar si está lista o no.

Como conclusión, y después de haber examinado las diferentes herramientas disponibles y de haber realizado los cambios pertinentes en el prototipo *Spaterm*, podemos confirmar que tenemos a nuestra disposición una herramienta muy completa, fácil de usar y que satisface todas las necesidades presentadas durante el proceso de trabajo en un marco ontoterminográfico, desde la gestión del corpus de trabajo hasta la exportación en formato de intercambio TBX. Todo lo cual nos permitirá llevar a cabo un trabajo adecuado, sistemático y coherente en el dominio del turismo de aventura.

Una vez terminado el trabajo ontoterminográfico con esta herramienta, nuestra tarea no finaliza con la edición de dicha información en formato de diccionario (ya sea en formato PDF o RTF) o con la simple consulta de términos en la misma base de datos, como ocurre con la mayoría de los trabajos de este tipo, sino que, desde nuestra investigación, defendemos una edición más compleja y acorde con las necesidades de consulta de los traductores profesionales, como exponíamos en el capítulo anterior (cf. 2.6.1.6.). Para ello, la edición de nuestra base de datos elaborada con la aplicación *ProTermino*, expuesta con anterioridad, trabajaremos con una aplicación informática independiente, desarrollada dentro de esta investigación, que nos permitirá realizar búsquedas textuales además de búsquedas por términos y búsquedas externas (en buscadores y recursos terminológicos externos). Esta aplicación se presentará en la última fase del capítulo siguiente por considerarse fuera del proceso de elaboración de la base de datos ontoterminográfica, ya que pertenece al proceso de edición y consulta de la base de datos.

3.3. ANÁLISIS DEL DOMINIO

Una vez que hemos tomado las decisiones iniciales del proyecto, es decir, hemos seleccionado el dominio y las herramientas de trabajo y hemos establecido las variables pragmático-lingüísticas del proyecto, debemos comenzar con un análisis del dominio en cuestión, a fin de adquirir los conocimientos necesarios para poder trabajar de forma eficiente en el dominio y evitar cualquier error conceptual posible. No obstante, a pesar de realizar un análisis tan exhaustivo como sea posible, por lo general los terminógrafos no podrán considerarse expertos en la materia, por lo que es imprescindible que se encuentren en constante contacto con expertos que puedan asesorarles ante cualquier duda y que puedan revisar el trabajo realizado.

En nuestra opinión, este análisis debe ser lo más completo posible y tratar los puntos más relevantes del dominio de trabajo, como son el mercado del dominio, las características del lenguaje de especialidad, la legislación pertinente y los recursos terminológicos disponibles en el momento de empezar a trabajar. Asimismo, considerar las necesidades del grupo de destinatarios al que va dirigido y, si fuera posible, realizar un estudio de los posibles problemas que presenta la traducción en ese dominio y para las lenguas de trabajo elegidas.

A continuación, presentamos el análisis llevado a cabo como parte de nuestro trabajo ontoterminográfico para el segmento del turismo de aventura, que consta de las diferentes partes que se han especificado con anterioridad.

3.3.1. EL CONCEPTO TURISMO

Antes de comenzar a tratar la actividad turística en España, se nos hace imprescindible tratar el concepto de <turismo>, a fin de comprenderlo de forma adecuada y poder distinguirlo de otras actividades que se podrían confundir con él pero que no lo son.⁸⁹³ A lo largo del tiempo, se han ido sucediendo diferentes definiciones, las cuales suelen incidir en una serie de aspectos tales como la idea de desplazamiento o

⁸⁹³ Este punto es especialmente importante en el segmento del turismo de aventura, nuestro objeto de estudio, puesto que se debe hacer una separación entre actividades como el senderismo o el montañismo, cuando no se pueden considerar actividades turísticas sino actividades deportivas.

el motivo de la estancia. No obstante, la tarea de definir qué es el «turismo» no es sencilla, debido a las siguientes razones:

- Razones derivadas por su carácter multifacético: económico, sociológico, político, psicológico, geográfico, etc., lo que provoca que las definiciones referentes a «turismo» adquieran un punto de vista desde la disciplina que la elabore;
- Razones derivadas de la multitud de segmentos turísticos que lo compone: turismo de sol y playa, turismo de aventura, turismo cultural, etc.⁸⁹⁴ y que presentan sus propias características intrínsecas, tanto en lo que respecta a las actividades que lo componen como a los turistas.
- Razones derivadas de la diversidad de subsectores relacionados con el turismo: alojamiento, agentes de viajes, transporte, comercio y restauración, medio ambiente, etc., que están regulados por normativas específicas y a menudo de fuera del ámbito turístico.

La Organización Mundial del Turismo (OMT),⁸⁹⁵ la principal institución internacional en el campo del turismo, elaboró la siguiente definición de «turismo» con fines estadísticos: «[el turismo se define como un] conjunto de actividades que realizan las personas durante sus viajes y estancias en lugares distintos al de su entorno habitual, por un periodo de tiempo consecutivo inferior a un año, con fines de ocio, por negocio o por otros motivos, y no por motivos lucrativos».⁸⁹⁶ Asimismo, esta organización consideró pertinente la diferenciación entre términos como «visitante», «turista» y «excursionista», con el fin de clarificar estos conceptos durante la elaboración de

⁸⁹⁴ Cf. 3.3.1.

⁸⁹⁵ La Organización Mundial del Turismo es un organismo especializado de las Naciones Unidas y desempeña un papel central y decisivo en la promoción del desarrollo de un turismo responsable, sostenible y accesible para todos. La Organización promueve el cumplimiento del Código Ético Mundial para el Turismo, con miras a garantizar que los países miembros, los destinos turísticos y las empresas maximicen los efectos económicos, sociales y culturales positivos del turismo y maximicen sus beneficios, reduciendo a la vez al mínimo los impactos ambientales y sociales negativos. URL: http://www.unwto.org/index_s.php.

⁸⁹⁶ Esta definición se aprobó en la *Conferencia sobre Estadísticas de Viajes y Turismo* celebrada en Ottawa en junio de 1991, donde se acordaron una serie de recomendaciones sobre estadísticas de turismo de la OMT.

estadísticas. Estas definiciones, acordadas también en la Conferencia de Ottawa (1991), son las siguientes:

[Un visitante es una] persona que visita por no más de un año un país diferente de aquél en el cual tiene de ordinario su residencia y cuyo motivo principal para la visita es distinto del de ejercer una ocupación remunerada en el mismo país que visita.

En esta definición, se contemplan a su vez dos conceptos diferentes: «turista» y «excursionista»:

(1) Turistas, visitantes como se definen anteriormente, que permanecen al menos 24 horas pero no más de un año en el país que visitan y cuyos motivos de viaje pueden clasificarse del modo siguiente:

a) placer, distracción, vacaciones, deporte;

b) negocios, visita a amigos o parientes, misiones, reuniones, conferencias, salud, estudios, religión.

(2) Excursionistas, visitantes como se definen anteriormente, que permanecen menos de 24 horas en el país que visitan (incluidos los pasajeros en crucero).

En este caso, distinguiríamos entre un turista y un “no turista”, dependiendo de si pernocta o no en un lugar diferente de su residencia habitual, donde se encuentra por motivos no lucrativos. Esta distinción, no obstante, no es compartida por todas las definiciones de turismo y, en algunos casos, no se establece dicha diferenciación y, por tanto, se consideran como «turista» ambos conceptos.⁸⁹⁷

La mayoría de las definiciones que existen hacen referencia a la demanda de los turistas y a sus aspectos principales: lugar, duración y motivación. Sin embargo, también se encuentran definiciones que parten de otros puntos de vista, como pueden ser el sociológico o el económico. A continuación, indicamos tres definiciones que muestran diferentes aproximaciones: la primera, desde un punto de vista profesional del turismo; la segunda, con una perspectiva sociológica; y la tercera, con una aproximación económica.

⁸⁹⁷ En nuestro caso, también consideraremos «excursionista» como «turista», debido a la naturaleza *sui generis* de nuestro segmento turístico, el turismo de aventura, donde las actividades que se practican son complementarias a otros tipos de turismo, como el turismo rural.

- «Conjunto de técnicas y actividades encaminadas a atraer y satisfacer las necesidades y motivaciones de los turistas en sus desplazamientos vacacionales. Por extensión, organización de los medios conducentes a facilitar estos viajes» (Montaner y Antich, 1998: 369);
- «El turismo es un fenómeno social que consiste en el desplazamiento voluntario y temporal de individuos o grupos de personas que fundamentalmente por motivos de recreación, descanso, cultura o salud, se trasladan desde su lugar de residencia habitual a otro, en el que no ejercen ninguna actividad lucrativa ni remunerada, generando múltiples interrelaciones de importancia social, económica y cultural» (de la Torre Padilla, 1989:33).
- «The sum of processes, activities and outcomes arising from the interactions among tourists, tourism suppliers, host governments, host communities, origin governments, universities, community colleges and non-governmental organisations in the process of attracting, transporting, hosting and managing tourists and other visitors» (Weaver y Lawton, 2006).

A pesar de existir un reconocido número de definiciones, la más aceptada y utilizada por todos es la propuesta por la OMT, que ya indicábamos anteriormente, la cual será la definición que tendremos en cuenta en nuestro trabajo de investigación. No obstante, en nuestro trabajo consideraremos «turista» a todo aquél que realice una actividad turística, ya sea con pernoctación o no, al contrario de lo que considera la OMT para sus fines estadísticos.

3.3.2. EL TURISMO COMO FENÓMENO ECONÓMICO EN ESPAÑA

El turismo, tal y como se conoce actualmente, es un fenómeno relativamente reciente, ya que se empezó a desarrollar en las sociedades industriales de finales del siglo XIX, motivado por la mejora de las condiciones de vida, el incremento del tiempo libre y al abaratamiento de las comunicaciones y el transporte. A lo largo del siglo XX, la experiencia turística se va aumentando y se difunde entre los diferentes grupos de la sociedad hasta la explosión del turismo de masas de los años cincuenta y sesenta. Actualmente, el turismo se ha convertido en una necesidad para los habitantes de los

países desarrollados, lo que provoca millones de desplazamientos anuales por motivos turísticos varios.

En España, la presencia del turismo ha sido constante desde el *boom* turístico que tuvo lugar en Europa en los años cincuenta y continúa siéndolo en la actualidad. El sector servicios representa hoy día unos de los pilares más importantes de la economía de España, y el turismo es considerado uno de los factores principales de este sector. Esta actividad no es solo una anécdota en nuestro país, sino que se trata de uno de los pilares que sustentan la economía española, ya que la mayor parte de los ingresos de España se obtienen gracias al turismo y un número muy elevado de empresas se dedican a este sector.⁸⁹⁸ Como indica la OCDE (2010: 248), el turismo ocupa el 11% del PIB total y el 13% del empleo en nuestro país, unos datos nada desdeñables.

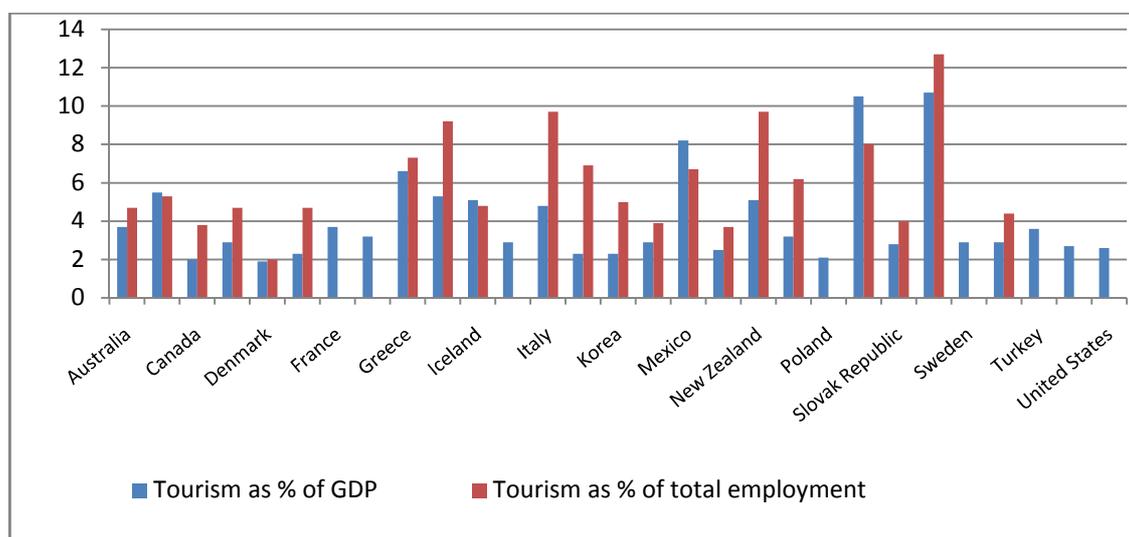


Gráfico 3. El sector turístico en las economías de los países de la OCDE. Fuente: OECD (2010: 8).

En el año 2006, España ocupaba el segundo puesto en el mundo, tanto como país más visitado (después de Francia) como en lo referente al gasto generado por los turistas (después de los Estados Unidos), según el informe del año 2006 del Instituto de Estudios Estadísticos (IET), dependiente del Ministerio de Industria, Turismo y

⁸⁹⁸ Según el *Boletín informativo del Instituto Nacional de Estadística* (vol. 3, 2006), el 78,4% de las empresas activas a 1 de enero de 2005 se dedican a este sector y han generado un volumen de negocio que supera el billón de euros. Asimismo, la actividad turística para España representa alrededor del 11% del PIB (IET. *Balance del turismo. Año 2009, 2010*).

Comercio.⁸⁹⁹ Sin embargo, como veremos más adelante, esta posición ha variado en la actualidad debido a los cambios producidos por la crisis económica de 2008. Esta posición del año 2006 se mantuvo durante el año siguiente cuando se llegó a registrar un total de 59,2 millones de visitas de turistas no residentes según la encuesta de Movimientos Turísticos en Fronteras (Frontur),⁹⁰⁰ que elabora el mismo instituto, lo que significaba un incremento del 4,5% con respecto al año 2006. En aquel momento, según los pronósticos de estas fuentes oficiales consultadas, los datos de crecimiento positivo recogidos en los años 2006 y 2007 se repetirían en los años sucesivos.

A pesar de estos datos positivos, el crecimiento en la recepción de turistas y del gasto comenzó a disminuir en el año 2008 a raíz de la crisis económica y financiera que ha sacudido la economía mundial y que continúa presente actualmente. Según el informe anual sobre el año 2009 publicado por el IET (IET, 2010b: 14), España sufrió en 2008 un retroceso del 2,5% con relación a la cifra alcanzada de turistas internacionales, un dato importante si tenemos en cuenta que se trata del primer retroceso desde que se empezó a realizar esta encuesta en 2002. En total, visitaron nuestro país 1,3 millones de turistas menos que el año anterior, lo que supuso un revés para nuestra economía. Por su parte, en 2009, la llegada de turistas internacionales volvió a disminuir y, en esta ocasión, con una cifra mucho mayor: un 8,7% respecto del año anterior (ibíd.: 14), lo que provocó un segundo retroceso del mercado turístico español y aumentó las dificultades económicas de nuestro país.

Gracias a la leve recuperación de la crisis económica y financiera mundial, el año 2010 vio un aumento moderado de la recepción de turistas y del gasto en el mercado turístico español y consiguió cerrarse en positivo, lo que cambia la tendencia al romper dos años consecutivos de descensos y abre una etapa que se presume de sostenida recuperación (IET, 2010c: 1). Concretamente, en el año 2010 España recibió 52,7 millones de turistas, cifra un 1% superior a la alcanzada en 2009 (ibíd.). No obstante, a pesar de los datos positivos registrados, el año 2010 se puede calificar de convulso según la *Nota coyuntural* publicada por el IET (ibíd.) debido a una serie de

⁸⁹⁹ < <http://www.iet.tourspain.es/informes/documentacion/publicaciones/TurisCifras2006Esp.pdf>>

⁹⁰⁰ <<http://www.iet.tourspain.es/informes/documentacion/FronturFamiliar/Balance%20del%20turismo%20en%20Espana%20en%202006.pdf>>

factores externos a los que se han tenido que hacer frente durante el transcurso del año y que dieron lugar a descensos en la recepción de turistas, a saber: en primer lugar, durante la primera parte del año se produjo un fuerte retroceso (13,3%) con respecto al año anterior debido a la nube de cenizas volcánicas que invadió el espacio aéreo europeo y la consiguiente cancelación de numerosos vuelos; y, en segundo lugar, al cierre del año tuvo lugar la huelga salvaje de controladores, que provocó el cierre del espacio aéreo español y la cancelación de numerosos vuelos, junto con el cierre de varios aeropuertos europeos provocado por las adversas condiciones meteorológicas. Estos factores externos provocaron, por tanto, que los resultados del balance del año 2010 fueran inferiores a las que podrían haberse obtenido.

Otro aspecto importante que cabe destacar de este año 2010 consiste en el retroceso de recepción de turistas procedentes de los principales mercados emisores, a saber: Reino Unido y Alemania. Estos, a pesar de seguir siendo los principales mercados para nuestro país, han registrado una reducción en el número de turistas durante el año 2010, que no ha provocado números negativos de crecimiento gracias al aumento de turistas procedentes de Francia, Italia y del conjunto de los países nórdicos.



Gráfico 4. Llegada de turistas internacionales a España durante los años 2005-2010 (elaboración propia a partir de los datos extraídos del IET).

Con respecto a 2011, los últimos datos disponibles nos indican que el comienzo del año para el mercado turístico español ha sido muy positivo, ya que en el mes de enero se registró un total de 2,7 millones de turistas internacionales, lo que ha supuesto un crecimiento del 4,7% y una cifra de 119 mil turistas más que en enero del año anterior. Este dato se considera relevante, ya que no hay que olvidar que fue en enero de 2010 cuando se produjo el primer avance desde julio de 2008, lo cual refuerza aún más el carácter positivo de este dato y las expectativas de crecimiento para este año.

El retroceso en la llegada de turistas extranjeros a España sufrido durante los años 2008 y 2009, causados por la crisis económica y financiera, provocó un cambio en la posición que ocupaba nuestro país con respecto al año 2006, cuando España se encontraba en el segundo puesto tanto en lo referente al número de visitas como de gasto generado por los turistas. Según datos publicados por la OMT en su informe *Panorama del turismo internacional* sobre el año 2009 (OMT, 2009: 5), los tres primeros puestos, tanto en lo referente a llegadas de turistas internacionales como en ingresos por turismo internacional, siguen correspondiendo a los Estados Unidos, España y Francia, aunque en un orden diferente, a saber: Francia sigue ocupando el primer puesto como destino del turismo internacional en términos de llegadas, aunque se sitúa en tercer lugar en lo referente a ingresos. Los Estados Unidos se sitúan en primer lugar en ingresos y el segundo lugar en llegadas, que habían perdido frente a España después del 11 de septiembre. Por su parte, España descendió al tercer lugar en términos de llegadas, aunque mantuvo su posición de segundo país con más ingresos del mundo y primero de Europa.

Según el informe sobre la situación del turismo mundial en el año 2010 presentado por la OMT en el Foro Mundial del Turismo en marzo de 2011 en Andorra, España sigue siendo el segundo destino mundial por ingresos turísticos, por detrás de los Estados Unidos, por lo que ha mantenido su posición desde antes del comienzo de la crisis mundial. Sin embargo, esta organización ha confirmado que China la ha relevado de su tercera posición con referencia al número de visitas de turistas internacionales, lo que ha venido provocado por la escalada china en el mercado turístico. De esta manera, nuestro país se sitúa ahora en la cuarta posición con respecto a recepción de turistas de una lista encabezada por Francia y los Estados Unidos. Finalmente, la OMT confirma

también en este informe la recuperación de España durante el año 2011, aunque aún no volverá a las cifras previas a la crisis.

Con este panorama, observamos la relevancia que tiene el turismo internacional en nuestro país y que solo el hecho de bajar de un tercer a un cuarto puesto puede ocasionarnos grandes dificultades en la economía española. En este contexto, consideramos necesario una continua atención hacia el motor de nuestra economía en todos los niveles: promoción exterior, calidad de servicios, facilidad de transporte, atención multilingüe, etc. así como mantener el contacto y el atractivo para nuestros principales mercados emisores de turistas, que son Reino Unido y Alemania, Francia e Italia.

Con respecto a este tema, los datos del Instituto Nacional de Estadística⁹⁰¹ y del IET nos indican que Reino Unido es el principal emisor de turistas hacia España, seguido por Alemania y Francia (en este orden), los cuales concentran entre los tres el 61% de los turistas no residentes de nuestro país.

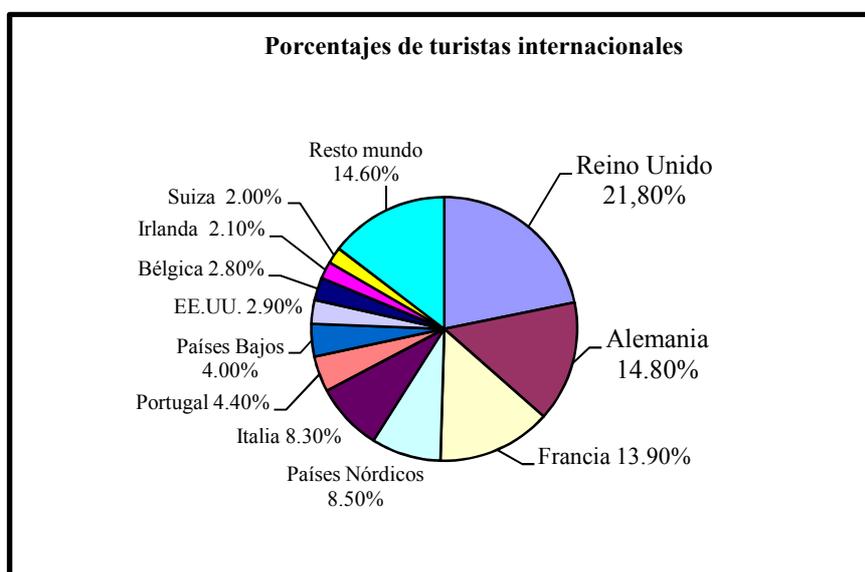


Gráfico 5. Porcentaje de turistas internacionales según el IET. Movimiento de estudios en Fronteras (Frontur).

⁹⁰¹ <<http://www.ine.es/>>

A pesar de la crisis y de la consecuente disminución de turistas internacionales procedentes de estos mercados emisores hacia España, el orden de estos países se ha mantenido durante los años 2008, 2009 y 2010. Durante estos tres años, se han producido algunas fluctuaciones mensuales de turistas en las que a veces el mercado francés ha adelantado al mercado alemán en número de turistas, como por ejemplo ocurrió en abril de 2010 cuando el mercado emisor alemán se vio reducido un 16,2% y el mercado francés aumentó un 18,8%. No obstante, en cálculos generales siguen siendo Reino Unido y Alemania los mercados emisores de turistas hacia España. Con la recuperación económica de estos principales países emisores, junto con el resto de países importantes que emiten turistas hacia nuestro país (Francia, Italia, países nórdicos, etc.) se ha presenciado un aumento importante del flujo de turistas hacia España, con la excepción en el momento actual de Reino Unido, lo que nos indica una tendencia de crecimiento positivo durante el año 2011 y siguientes, especialmente si tenemos en cuenta el número de turistas que elegirá España como destino debido a las revueltas que se están sucediendo en algunos países árabes.

La mayoría de estos turistas internacionales se concentran en seis comunidades autónomas, que son consideradas como las más receptoras en nuestro país: Cataluña, Baleares, Canarias, Andalucía, Comunidad Valenciana y Madrid, por este orden. Estas comunidades, aunque se han mantenido inamovibles en sus puestos, han sufrido descensos tanto en la recepción de turistas (por ejemplo, Andalucía ha recibido un 7,7% menos de turistas internacionales) como en el gasto (por ejemplo, en Cataluña y Canarias) debido a la bajada de turistas por la crisis económica (IET, 2010b: 7-8).⁹⁰²

A causa de esta crisis económica, sumado a la aparición de otros posibles mercados competitivos para España, como son Turquía o Croacia, que ofrecen precios más bajos y cuentan con unos recursos naturales casi intactos,⁹⁰³ se están llevando a cabo desde hace unos años políticas turísticas y programas y proyectos, tanto en el

⁹⁰² Al no contar en este momento con los datos definitivos relativos al año 2010, nos basamos en el informe elaborado sobre el año 2009.

⁹⁰³ Datos extraídos del Plan de marketing del turismo en Andalucía 2005. Consejería de Turismo, Comercio y Deporte. URL: http://www.juntadeandalucia.es/turismocomercioydeporte/ctcd-docs/custom_doc/PLAN_de_MARKETING_2005.pdf.

ámbito nacional como regional,⁹⁰⁴ con objeto de aumentar la competitividad turística de España en todos sus aspectos: calidad de la oferta turística, marketing exterior, nivel de precios, satisfacción del turista, infraestructuras, etc.⁹⁰⁵ Con estas medidas, los diferentes gobiernos pretenden mejorar los criterios de calidad y fomentar la sostenibilidad, diferenciación y especialización del sector turístico español y de sus diferentes segmentos (sol y playa, activo, rural, cultural, etc.), a fin de poder competir con los demás destinos turísticos emergentes y mantener su posición tanto en Europa como en el mundo.

3.3.3. LOS SEGMENTOS TURÍSTICOS EN AUGE

Como indicábamos con anterioridad, una de las dificultades encontradas en la definición de «turismo» se debe a la heterogeneidad de actividades que se engloban dentro de este concepto, ya que existe tal cantidad de oferta turística diversa que es complicado considerarlas todas dentro de una definición. No obstante, comparten unos objetivos comunes, que principalmente son:

- Satisfacer al turista según sus necesidades y motivos de viaje (relajación, aventura, descanso, visitas culturales, etc.); y
- Ofrecer servicios de calidad.

Desde hace algunos años se está fomentando la actuación hacia un turismo sostenible, especializado y flexible desde los diferentes niveles gubernamentales (regional, nacional y europeo), con objeto de mejorar la calidad de los servicios turísticos y de competir con otros mercados emergentes, además de adaptarse a las nuevas necesidades y valores que

⁹⁰⁴ Destacan el Programa de Caminos Naturales del Ministerio de Medio Ambiente, el programa de evaluación de calidad del Instituto para la Calidad Turística Española, el Plan de Calidad Turística de Andalucía 2006-2008 (Consejería de Turismo, Comercio y Deporte, 2006), el Plan General del Turismo Sostenible de Andalucía 2008-2011 (Consejería de Turismo, Comercio y Deporte, 2008), Plan Turismo Español. Horizonte 2020 (Ministerio de Industria, Comercio y Deporte) y el Plan Turismo Español 2008-2012.

⁹⁰⁵ En el informe Balance del turismo. Año 2009. (IET, 2010a: 15 y ss.) se puede consultar un estudio realizado sobre la competitividad turística de España con otros mercados competidores (Turquía, Túnez, Egipto, Croacia, Grecia, entre otros).

presentan los “nuevos” turistas. Estos turistas exigen un turismo ético⁹⁰⁶ y de calidad, son cada vez más independientes, están concienciados con la protección del medio ambiente y la sostenibilidad y utilizan preferiblemente la Red para buscar, seleccionar y contratar sus vacaciones. Según Goytia Prat (2000), las nuevas tendencias en la demanda turística se podrían resumir de la siguiente manera:

i) Nuevos hábitos vacacionales:

- Reducción de la duración de las vacaciones principales.
- Incremento de viajes de corta duración.
- Incremento de viajes de largo recorrido.
- Reducción del nivel de fidelidad al destino.

ii) Cambio en los hábitos de compra.

- Incremento de la preferencia por los viajes organizados individualmente.
- Búsqueda de canales de compra alternativos (Internet).

iii) Cambios en los hábitos de consumo.

- “Cama y comida” no es suficiente.
- Mayor nivel de actividad (de “pasar” a “vivir” las vacaciones).
- Preferencia por establecimientos con elevado nivel de equipamiento.
- Preocupación por el entorno medio ambiental.
- Combinación de sol y playa con actividades lúdico-culturales.
- Deseo de un servicio personalizado y flexible.
- Deseo de garantía de servicio y servicios de atención y quejas.

⁹⁰⁶ El objetivo de este turismo debe ser promover y desarrollar nuevas formas de turismo que proporcionen a todos los participantes los mayores beneficios posibles sin causar cambios ecológicos o sociales intolerables” (en Goytia Prat, 1995: 128).

En este contexto, se está produciendo una tendencia a la segmentación del mercado turístico⁹⁰⁷ como medio para dar respuesta a estas necesidades y motivaciones nuevas de los turistas y, por tanto, para acercar la oferta turística a la demanda y no viceversa, como ocurría anteriormente. Así, por ejemplo, encontramos el tradicional turismo de sol y playa, el turismo rural, el turismo cultural, el turismo gastronómico, entre otros tipos, todos ellos con unas características propias y unos turistas potenciales con determinadas necesidades y motivaciones.

Asimismo, en los últimos años se están fomentando nuevas formas de *turismos* denominadas como «turismo alternativo», un turismo de baja densidad en áreas naturales o rurales, que ha surgido en contraposición al turismo de masas, característico de las formas de turismo tradicionales como el turismo de sol y playa. Según Fennell (1999: 6):

Alternative Tourism is a generic term that encompasses a whole range of tourism strategies (e.g. “appropriate”, “eco-”, “soft”, “responsible”, “people to people”, “controlled”, “small scale”, “cottage”, and “green tourism”) all of which purport to offer a more benign alternative to conventional mass tourism in certain types of destinations.

Las formas de turismo alternativo que podemos encontrar actualmente son muy variadas y, de hecho, a menudo sus límites son difusos y se entremezclan, llegando a menudo a coincidir en las actividades que integran cada forma de turismo. Esto es debido principalmente a la juventud de este tipo de turismo,⁹⁰⁸ al rápido surgimiento de las diferentes ofertas turísticas y a la libertad de las empresas turísticas de denominar sus propias actividades.

⁹⁰⁷ La segmentación de mercado, en este caso el turístico, es la división del mercado en cuestión en grupos homogéneos y más pequeños a los que se les aplicará una estrategia comercial diferenciada con el fin de satisfacer necesidades y deseos de los individuos que forman los segmentos y con la finalidad de que la empresa alcance sus objetivos (Llamas Arjona, 2009: 30).

⁹⁰⁸ Se empezó a tener consciencia de este turismo alternativo a finales de la década de los ochenta, pero fue a partir de la década de los noventa cuando se empezaron a tomar medidas para fomentar este tipo de turismo y la gestión sostenible de todas las formas de turismo. Ejemplos de ello son la Carta Mundial del Turismo Sostenible, redactada en el marco de la Conferencia Mundial de Turismo Sostenible que se celebró en Lanzarote en abril de 1995; o la Carta Europea de Turismo Sostenible en los espacios protegidos, elaborada en 1995, una iniciativa de la Federación EUROPARC (<http://www.european-charter.org/home/>) que tiene como objetivo global promover el desarrollo del turismo en clave de sostenibilidad en los espacios naturales protegidos de Europa y está fundamentada en las recomendaciones de la Agenda 21, adoptadas en la Cumbre de la Tierra en Río de Janeiro en 1992 y por el sexto Programa de acciones comunitarias para el desarrollo sostenible (Fundación EUROPARC, 2007).

Entre los diferentes tipos de turismo que se conocen como «turismo alternativo», destacamos los siguientes: el ecoturismo (*ecological tourism*), definido por *The International Ecotourism Society* (TIES)⁹⁰⁹ como el «responsible travel to natural areas that conserves the environment and sustains the well being of local people», es decir, un turismo dirigido al enriquecimiento de la experiencia turística por medio de una mayor integración en el medio visitado y de una participación más importante en la vida social y cultural del lugar de vacaciones basado en el respeto por el medio ambiente y las costumbres locales;⁹¹⁰ el turismo activo (en contraposición al turismo pasivo tradicional), un tipo de turismo determinado por unas «vacaciones activas» (Goytia Prat, 2000), caracterizado por el deseo de realizar actividades de mayor o menor intensidad en la naturaleza durante el periodo vacacional, las cuales pueden ir desde la práctica del montañismo hasta un paseo en catamarán; el agroturismo, un tipo de turismo rural en el que el componente principal es la participación en tareas relacionadas con la explotación agraria y ganadera; el turismo de aventura, una forma de turismo que consiste en la práctica de actividades de aventura de mayor a menor intensidad en el medio natural; y el turismo de naturaleza o turismo verde,⁹¹¹ un tipo de turismo que se desarrolla en el medio natural, especialmente en los espacios naturales protegidos, y que

⁹⁰⁹ <<http://www.ecotourism.org/>>. Según esta asociación, el ecoturista debe seguir los siguientes principios: a) Minimizar el impacto turístico de la zona; b) Respetar la cultura local; c) Proporcionar experiencias positivas tanto para los visitantes como para las zonas receptoras; d) Proporcionar beneficios económicos para la conservación y para el desarrollo local; e) Fomentar la sensibilidad en cuanto a la política, medio ambiente y sociedad de los países receptores.

⁹¹⁰ No existe actualmente una definición universal para describir este tipo de turismo alternativo, por lo que a menudo se considera como una macrocategoría que engloba al resto de formas de turismo relacionadas con la naturaleza (turismo de aventura, turismo de naturaleza, agroturismo, etc.) (Wight, 2001: 48) y en otras ocasiones, se considera como una categoría diferente, donde impera el respeto y aprendizaje de la cultura local y el contacto con la naturaleza (Hawkins y Lamoureux, 2001: 64). Asimismo, recibe diversas denominaciones como, por ejemplo, «turismo ambiental», «turismo verde», «turismo ecológico», entre otras. Para profundizar en esta forma de turismo, véanse los diferentes capítulos que integran Weaver (2001a).

⁹¹¹ Estrictamente, tanto el agroturismo como el turismo de naturaleza y el turismo de aventura serían considerados como turismo activo, puesto que se trata de formas de turismo en el que el turista realiza diferentes actividades y su actitud es activa en todo momento. Sin embargo, el turismo activo se ha venido considerando en la legislación española (cf. 3.3.3.1.2.) como un turismo vinculado más a la práctica de deportes que a otro tipo de actividad recreativa, por lo que está más relacionado con el turismo activo de deporte. No obstante, cuando el turismo de naturaleza implica la realización de actividades recreativo-deportivas también sería considerado turismo activo.

se caracteriza por el deseo de disfrutar la naturaleza de forma activa pero relajada (avistamiento de pájaros, visita a espacios naturales protegidos, etc.).⁹¹²

En general, como se ha podido observar hasta ahora, no existe unanimidad para la segmentación de las formas de turismo que integran la categoría de turismo alternativo y, de hecho, muchas de ellas podrían incluirse dentro de algunas de las opciones indicadas, como es el caso del agroturismo, que se podría incluir dentro del ecoturismo, del turismo activo o del turismo rural, según la aproximación que se eligiese. A pesar de esta variedad de denominaciones aún poco claras, se advierte la tendencia actual de las actividades turísticas hacia un acercamiento al medio rural y natural, así como hacia turismo respetuoso con el medio ambiente. De esta manera, cabe destacar que estas formas de turismo no se consideran mejores que las formas más tradicionales, como el turismo de sol y playa o el turismo de salud y belleza, sino que se defiende una gestión sostenible del turismo en todos los segmentos turísticos, ya que sin ella cualquier tipo de turismo sería dañino para la zona donde se reciba.

A continuación, nos centraremos en el segmento de turismo alternativo elegido para esta tesis doctoral, el turismo de aventura, a fin de exponer su marco conceptual y sus principales características.

3.3.3.1. El turismo de aventura como turismo alternativo

La historia más reciente de las actividades practicadas en el turismo de aventura nos indica que aparecieron en Europa y España en la década de los setenta, siendo en la costa californiana donde se iniciaron la mayoría de las actividades; en los años 80 se desarrollaron y en los 90 se consolidaron, llegando a convertirse en uno de los segmentos turísticos que más ha crecido en la última década. En su comienzo, se trataban de actividades propias de aventureros, pero hoy en día las puede practicar cualquier tipo de persona, gracias al desarrollo del equipamiento que utilizan y de los

⁹¹² Si entendemos turismo rural como el conjunto de actividades recreativas realizadas en zonas rurales y basadas en elementos de cultura rural (Sayadi y Calatrava, 2001: 133). El turismo de naturaleza podría ser considerado sinónimo del ecoturismo, si consideramos este en su acepción más general, o como una categoría superordinada que incluiría el ecoturismo, así como cualquier otra forma de turismo que se desarrollara en la naturaleza (turismo de aventura, turismo de sol y playa, el turismo cinegético) (Weaver, 2001b: 74).

niveles de seguridad que presentan. Tal y como afirman Fuster y Elizalde (1995: 95), en los últimos años, la práctica de estas actividades físicas en el medio natural ha evolucionado desde concepciones tradicionales o minoritarias a formas innovadoras al alcance de una amplia masa social.

Debido al surgimiento relativamente reciente y al rápido desarrollo que ha tenido este segmento, no cuenta aún con una denominación específica y acordada por los expertos, ya que, como afirma Luque Gil (2003: 130), se usan indistintamente diferentes términos, a saber: «turismo de aventura», «turismo deportivo en la naturaleza», «turismo activo-deportivo en la naturaleza», «turismo activo y de aventura», «turismo activo en la naturaleza» «turismo deportivo y de aventura», etc., lo que provoca cierto desconcierto a la hora de utilizar y comprender el concepto. No obstante, a pesar de que estos términos se suelen utilizar de modo indiferente en numerosas ocasiones, la mayoría de los autores parecen decantarse por la utilización del término de «turismo de aventura».⁹¹³ En este contexto, seguiremos la línea general y emplearemos también este término para referirnos al segmento objeto de nuestra investigación.

Como veíamos anteriormente, este segmento turístico se trata de una forma de turismo alternativo que consiste en la práctica de actividades de aventura de mayor a menor intensidad al aire libre y en contacto con el medio natural. De esta manera, podemos determinar las características principales de este tipo de turismo:

- Se trata de un turismo sostenible, que cuida el medio ambiente y, por tanto, es contrario a la masificación que presentan otros segmentos.
- Consiste en la realización de actividades de aventura, es decir, actividades que incluyen cierto elemento de riesgo, desconocimiento y desafío, tanto mental como psicológico, como son el rafting, el barranquismo, el puenting, etc. Estas actividades se consideran como deportes de aventura, puesto que muchas de ellas son deportes federados y cuentan con numerosos profesionales que las practican.

⁹¹³ No obstante, encontramos otros muchos que prefieren la denominación de «turismo activo» como sinónimo de «turismo de aventura», ya que consideran que la idea de «aventura» implica que algo se sabe como comienza pero no como acaba, y el turismo debe acabar siempre con la satisfacción del cliente (cf. Aspas Aspas, 2000: 28).

- Estas actividades ofrecen diferentes niveles de riesgo, con objeto de llegar a un público más general, que pueda elegir una actividad de mayor o menor intensidad según su forma física y sus intereses.
- Se realiza en el medio natural, es decir, ofrece una interacción directa con la naturaleza.

A partir de estas características, el turismo de aventura se suele considerar como parte del macrosegmento del turismo en la naturaleza,⁹¹⁴ puesto que se practica en el medio natural, donde se incluirían además el ecoturismo, el turismo rural, etc., es decir, todas las actividades que se realizan de forma activa o pasiva en la naturaleza. A su vez, se suele englobar dentro de lo que se conoce como el turismo activo, especialmente en el ámbito institucional español y en gran parte del sector empresarial de los diferentes países, con objeto de enfatizar su carácter activo al realizar actividades de aventura de diferente naturaleza e intensidad. En este segmento, el turismo de aventura quedaría al mismo nivel que el agroturismo, el turismo cinegético, el turismo de pesca deportiva, etc., es decir, con el resto de segmentos que implican el comportamiento activo del turista durante el periodo vacacional. A su vez, dentro del turismo activo, se diferencian diferentes segmentos, y el turismo de aventura se englobaría dentro de lo que se conoce como turismo deportivo o turismo de deportes, ya que, como decíamos, las actividades que se realizan en este segmento se conocen como deportes de aventura. En este sentido, el turismo de aventura se encontraría al mismo nivel que el segmento del turismo náutico, del turismo de golf y del turismo de nieve. En la Figura 40, se puede observar el marco conceptual del turismo de aventura que proponemos en esta investigación, es decir, la relación de inclusión que presenta este segmento, en primer lugar, con el turismo en la naturaleza, que englobaría todos los tipos de turismos siguientes; en segundo lugar, con el turismo activo, que incluiría los más internos; y por

⁹¹⁴ Téngase en cuenta que el turismo en la naturaleza es diferente del turismo de naturaleza (o turismo verde), ya que el primero se trata de un macrosegmento que incluye todas las actividades turísticas que se realizan en el medio natural, lo que incluiría segmentos como el turismo de aventura, el turismo rural, el turismo cinegético, etc. Por el contrario, el segundo hace referencia a las actividades que se basan en el disfrute y observación de la propia naturaleza, como es el avistamiento de pájaros o cetáceos, visitas a parques naturales, etc. De esta manera, el primero incluye al segundo, pero no son equiparables.

último, con el turismo deportivo, que incluiría el turismo de aventura junto con otros tipos de turismo deportivo activo.⁹¹⁵

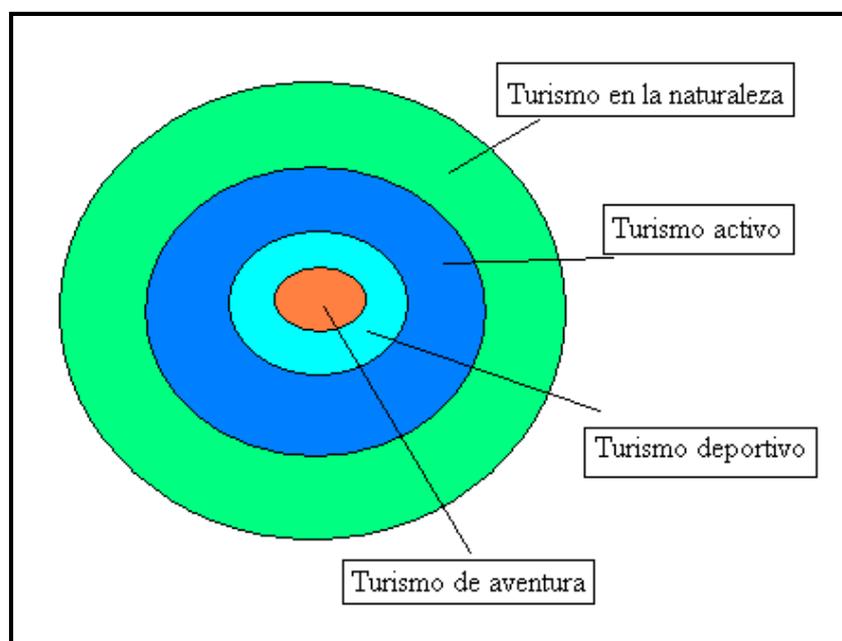


Figura 40. Marco conceptual del turismo de aventura.

En definitiva, en nuestra opinión, el turismo de aventura es a) un turismo alternativo, puesto que se trata de un turismo de baja densidad, opuesto al turismo de masas o convencional, se realiza en el medio natural o rural, etc.; b) es un turismo en la naturaleza, ya que el medio natural se utiliza como lugar de desarrollo de sus actividades; c) es un turismo activo, puesto que requiere de un comportamiento activo de los turistas a la hora de realizar las actividades; y d) es un turismo deportivo, cuya motivación principal es la práctica de una actividad deportiva por ocio y recreo.

Para continuar con el análisis de este segmento turístico, con respecto a su evolución y estado actual, su legislación, sus peculiaridades y sus actividades turísticas,

⁹¹⁵ El turismo deportivo aparece representado como parte del turismo activo, aunque también turismo no activo. Esto es debido al hecho de que este segmento se clasifica en dos: por un lado, encontramos el turismo de práctica deportiva, en la que se incluiría el turismo de aventura, y que consiste en la práctica de deportes durante las vacaciones, ya sea esta el principal objetivo vacacional o complementario; y por otro, encontramos el turismo de espectáculo deportivo, que consiste en la asistencia a acontecimientos deportivos durante la estancia vacacional, ya sea como principal objetivo o secundario. Este sería el caso de los turistas que se desplazan para asistir a las Olimpiadas, campeonatos, etc. Dentro de este segundo tipo, podemos incluir también las visitas a instalaciones deportivas, como estadios, circuitos, etc. (de Carvalho y Lourenço, 2009: 126).

presentaremos a continuación una serie de apartados donde abordaremos cada uno de estos aspectos.

3.3.3.1.1. Vínculos entre turismo y deporte

El primer aspecto básico que se debe considerar en un estudio sobre las actividades del segmento del turismo de aventura es la dificultad existente a la hora de establecer una frontera entre lo turístico y lo deportivo, ya que, como veíamos antes, este segmento se engloba dentro de lo que se conoce como turismo deportivo y sus actividades son consideradas como deportes de aventura.⁹¹⁶

Tradicionalmente, el término «deporte» se ha considerado como una actividad reglamentada, institucionalizada y competitiva, con unos valores intrínsecos determinados que lo diferencian de cualquier actividad física con un objetivo de recreación y de ocio y exentas de reglas y competición (Olivera y Olivera, 1995: 111). Según los autores que defienden esta concepción, se rechazaría la idea de considerar como deportes las actividades que se realizan dentro del segmento del turismo de aventura, ya que, según su postura, se trataría de actividades físicas de aventura que tienen lugar en el medio natural. Sin embargo, en la actualidad se está cambiando esta consideración y se tiende a una utilización del término «deporte» más amplia, donde se incluirían los «deportes de aventura» del segmento turístico (Hudson, 2003). Entre los cambios que se han producido para cambiar la concepción tradicional sobre el deporte, destaca la generalización de la práctica deportiva a todos los estratos de la población,

⁹¹⁶ Además de esta denominación, que se trata de la más extendida, las actividades practicadas en el turismo de aventura reciben las siguientes (Olivera, 1995, 5; León Guzmán y Parra Boyero, 2001): *nuevos deportes*, por su carácter innovador y alternativo; *deportes tecno/ecológicos*, por la unión entre naturaleza y tecnología; *deportes en libertad*, por la carencia de normas reglamentarias y federaciones reguladoras; *deportes californianos*, por su origen geográfico; *deportes salvajes*, *deportes extremos* o *deportes de riesgo*, por su búsqueda de sensaciones límites y difíciles; *actividades de reto en la naturaleza*, por su capacidad de luchar con uno mismo o contra la naturaleza; *actividades deslizantes de aventura* o *deportes glisse*, por considerar el deslizamiento como la actividad básica de estos deportes; *deportes Fun*, por su carácter divertido; *outdoor adventure recreation*, *actividades de reto en la naturaleza*, *actividades de emoción y aventura al aire libre*, por sus posibilidades de unir emoción y sensación y realizarse preferiblemente al aire libre; *actividades deportivas de esparcimiento y turísticas de aventura*, *actividades físicas de aventura en la naturaleza*, *actividades deportivas en el medio natural*, *actividades físico-deportivas en el medio natural*, por considerar su aspecto de esfuerzo físico en el medio natural; o *deportes adaptados*, por su carácter flexible en cuanto a la intensidad de la actividad para atraer al público en general.

asistiéndose a un crecimiento de la influencia del deporte sobre toda la sociedad y adquiriendo gran relevancia la concepción del «Deporte para Todos» entendido por el Comité Olímpico Internacional como:

El movimiento donde el objetivo es alcanzar el ideal olímpico, que proclama que el deporte es un derecho que pertenece a todos los hombres, sin distinción de raza ni clase. Este movimiento tiene por objetivo el promover la salud, la condición física y el bienestar mediante actividades deportivas que pueden ser practicadas por personas de sexo, edad, condiciones sociales y económicas diferentes, cual sean las culturas regionales o locales (Sánchez, 2001: 4).

Otro de los aspectos importantes que ha alterado la concepción del deporte en la actualidad es el interés surgido por las prácticas deportivas al aire libre y relacionadas con la naturaleza, que empiezan a adquirir gran importancia para las poblaciones urbanas acosadas por los problemas de las ciudades (tráfico, ruido, estrés, etc.) (Luque Gil, 2003: 122) y que, por tanto, fomenta la práctica de actividades deportiva en el medio natural.

Según Parra Boyero (2002: 182), algunos de los factores que pueden explicar la eclosión de estas actividades deportivas en el sector turístico y que, por tanto, han despertado el interés de los turistas por su práctica, se pueden resumir en los siguientes puntos:

- Factores de carácter organizativo:

- Se pueden realizar en un fin de semana.
- Se adapta a los periodos vacacionales.
- Son deportes de temporada sin necesidad de regularidad.
- Están mínimamente institucionalizados (no es necesario federarse, no hay entes oficiales, etc.).
- Se practican fuera del ámbito urbano.
- Pueden ser multidisciplinarios.
- Existe una amplia gama de actividades a realizar, casi a la carta.
- Son susceptibles de adaptarse fácilmente en complejidad y dificultad a cualquier nivel de destreza o de aprendizaje (de ahí por ejemplo el interés que adquieren para el campo turístico).

- El desarrollo empresarial ha supuesto que las empresas den un servicio total para realizar las prácticas en plenas condiciones sin preocuparse de nada (ha supuesto ampliar enormemente el número de demandantes).

- Factores de carácter emocional:

- Propician un sentimiento de libertad.
- Proporcionan mucho placer y sensaciones fuertes en poco tiempo.
- Se asocia a actividades placenteras, educativas y saludables. Suponen una manifestación cultural a la vez que un instrumento educativo/recreativo, que posibilitan el reencuentro con la naturaleza y la creación de hábitos de vida saludable.
- Se asocian a actividades de carácter físico-lúdico, a otra forma de vivir las vacaciones y la eclosión de otro tipo de turismo.

- Factores relacionados con la práctica:

- Pueden ser practicadas inmediatamente sin la preparación que exigen otros deportes.
- Mejora de los materiales, equipamientos y artefactos que son necesarios para su ejecución.
- No poseen un elevado coste económico.
- Se ha dado una gran respuesta empresarial que estimula constantemente el mercado, haciéndolo de más calidad y diversificando cada vez más la práctica.
- Aparecen continuamente nuevas modalidades.

En este contexto, se puede afirmar, por tanto, que hoy en día los conceptos «deporte» y «turismo» poseen una relación simbiótica,⁹¹⁷ percibiéndose un incremento en el número de personas que practican algún deporte y en el deseo de disfrutar de unas vacaciones activas o en las que estén presentes estas actividades deportivas (Luque Gil, 2003: 115). De ahí que el deporte se haya convertido en un factor de dinamización de la actividad turística y que las relaciones entre deporte y turismo estén generando nuevos horizontes

⁹¹⁷ Según Latiesa et ál. (2002: 22), para que una actividad sea considerada al mismo tiempo deportiva y turística se deben cumplir una serie de condiciones: el desplazamiento a un lugar distinto del habitual, debe tener una finalidad de ocio (no se incluye negocio para el caso del turismo o competición para el caso del deporte) y debe darse una actividad física participativa e intencionada (activa) o espectáculo (pasiva).

para el desarrollo del turismo, un apoyo para el desarrollo local y un nuevo nicho económico (Pigeassou, 1997: 16).

3.3.3.1.2. El turismo de aventura en España

El escenario turístico planteado en el informe de *Horizonte 2020*, elaborado por la OMT (2002), detecta un comportamiento muy positivo del turismo deportivo en todo el mundo y particularmente como componente importante del turismo en la naturaleza, coincidiendo con el interés cada vez mayor por el contacto directo con el medio natural y los destinos bien preservados desde el punto de vista medioambiental, así como con las sinergias crecientes entre el turismo y el deporte como dos actividades de gran repercusión social y económica en el mundo actual.

En esta misma línea, el último estudio realizado por el Ministerio de Industria, Turismo y Comercio sobre el turismo en la naturaleza (2005), confirma el crecimiento constante de este segmento en los últimos cinco años y su pronóstico favorable para los años siguientes. Según este estudio, las motivaciones principales para el viaje de este tipo son el descanso y la diversión (38%), seguidos de cerca por el deseo de realizar actividades deportivas (32%), lo que deja patente la importancia que está adquiriendo el turismo deportivo en la naturaleza en los últimos años. En definitiva, se puede constatar que tanto en el ámbito internacional como nacional, el turismo deportivo en la naturaleza es un segmento emergente y con futuro.

En España, los orígenes de la práctica deportiva en la naturaleza se remontan a los años ochenta y, desde entonces, su demanda ha estado en constante crecimiento. Este interés por las actividades de este segmento turístico se deben principalmente a los cambios experimentados por los consumidores del sector turístico y deportivo, que comentábamos más arriba, y al impulso dado a nuevos productos turísticos desde los planes y programas oficiales realizados por distintas administraciones, cuyos objetivos principales son la desmasificación del espacio litoral, un turismo más sostenible, la desestacionalización del mercado turístico, la diferenciación y recuperación de algunos territorios, etc. Entre los planes que destacan en el ámbito estatal, podemos citar el *Plan de Turismo Español. Horizonte 2020*, el *Plan de Turismo Español 2008-2012*, el *Plan*

Integral de Calidad del Turismo Español. PICTE 2000-2006. En este último, se especifica, además, que «el turismo de aventura tiene un gran potencial».

Debido a su reciente aparición, el turismo de aventura se considera un segmento turismo aún joven y poco consolidado, pero que poco a poco se está haciendo un hueco en el mercado turístico de España, gracias al aumento de empresas que se dedican al turismo activo,⁹¹⁸ a la presencia de estas empresas en Internet, a las subvenciones de las administraciones para mejorar infraestructuras, etc. En apenas un par de décadas, los deportes de aventura han dejado de considerarse especializados y orientados a grupos minoritarios de población para convertirse en una oferta cada vez más diversificada, adaptada y conectada con amplios segmentos de usuarios tanto nacionales como extranjeros (Rivera Mateos, 2007: 23), aunque muchas de estas prácticas quedan difuminadas hoy en día como oferta complementaria de ocio de otros segmentos turísticos, como son el segmento de sol y playa o el turismo rural.⁹¹⁹ Por su parte, el crecimiento de la oferta y la demanda en España de este tipo de actividades no impide reconocer la existencia aún en nuestro país de una cierta barrera cultural que ha originado hasta fechas recientes una menor tradición en su práctica y desarrollo, a diferencia de otros países cercanos como Francia, Italia, Reino Unido y, en general, todos los países alpinos centroeuropeos que cuentan al menos con treinta años de trayectoria (ibíd.: 24). Esto da lugar a un fuerte déficit de estadísticas y datos oficiales sobre la oferta y la demanda de este segmento, lo que hace muy complicado también el acceso a referencias sólidas acerca de este sector que permitan cuantificarlo de forma razonable. No obstante, se están llevando a cabo poco a poco estudios oficiales sobre los productos ofertados en este segmento turístico para contrarrestar esta situación, y un ejemplo de ello es un estudio reciente realizado por el Instituto de Turismo de España

⁹¹⁸ Desde el punto de vista de la oferta, son significativas las cifras de empresas y entidades turísticas ofertantes de turismo activo distribuidas por el conjunto de la geografía española y recogidas parcialmente en la guía oficial de *Fitur Activo*, que han alcanzado un total de 1.670 en el año 2005 con un incremento de un 40% respecto a la primera guía de 1997, registrándose un mayor número precisamente en las comunidades autónomas de Andalucía, Canarias, Cataluña y Madrid (Secretaría de de Estado de Turismo y Comercio, 2005b).

⁹¹⁹ En algunos lugares de España, estas prácticas dejan de ser complementarias para convertirse en objetivos principales del viaje, como ocurre en ciertos destinos especializados como las comarcas catalanas del Pallars y Valle de Arán, la Sierra de Guara o, en el caso de Andalucía, zonas como Tarifa (*surf*), Cabo de Gata (buceo y submarinismo) o Doñana (senderismo y rutas en 4x4).

(TURESPAÑA: 2008)⁹²⁰ sobre el turismo de montaña, donde se deja patente la importancia de este sector para la diversificación y especialización de la oferta, así como las diferentes actuaciones que se pretenden llevar a cabo para fomentar este tipo de turismo, entre las que destacan la promoción de los deportes de aventura, como el barranquismo, el rafting, la escalada, etc. especialmente para los mercados emisores de turistas más importantes (Reino Unido, Alemania y Francia), el desarrollo de planes de especialización de destinos turísticos españoles y de apoyo a la creación de empresas de este sector, la creación de un sistema de información, entre otras.

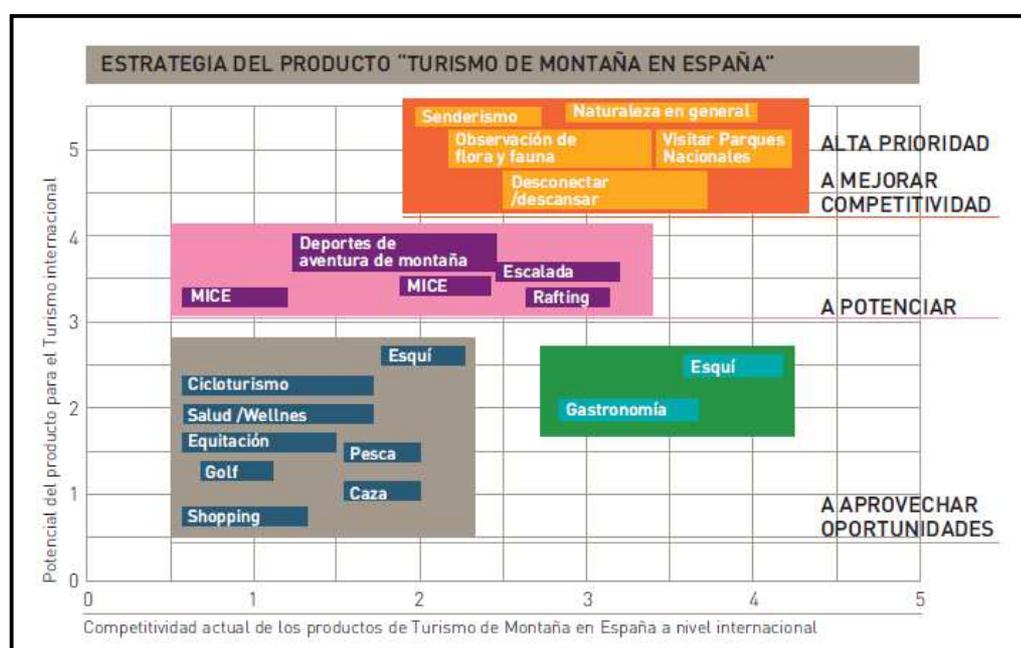


Gráfico 6. Situación actual y estrategia de actuación para fomentar el turismo de montaña en España. Fuente: TURESPAÑA (2008: 148).

A continuación, presentamos los principales rasgos que caracterizan al segmento del turismo de aventura en general según varios autores (Costa et ál., 1999: 779; Garau et ál., 2000, 600):

- La oferta de este tipo de turismo en España es muy reciente y no cuenta todavía con una promoción fuerte por parte de agencias especializadas, touroperadores o

⁹²⁰ <<http://www.tourspain.es/>>. El informe sobre el estado actual del turismo de montaña se puede consultar en la siguiente URL: <<http://www.tourspain.es/NR/rdonlyres/BE4C3474-E960-4150-9316-60FD56181D82/18739/turismomontana200994.pdf>>.

empresas fuertes de intermediación, tanto en el exterior como en el turismo nacional.

- Las actividades de aventura suele ser complementarias al objetivo principal del viaje y, por tanto, de más corta duración.
- Los turistas que realizan actividades de este segmento suelen ser nacionales (por tratarse de viajes de corta duración y no constituir el objetivo principal del viaje). A pesar de que la práctica de deportes no figura entre los principales motivos de los turistas extranjeros (en un ranking de 18 puntos ocupa el 15), el 20% de los turistas de temporada media y más de un 25% en temporada alta considera como importante o muy importante la práctica de deportes.
- En las instalaciones creadas la inversión requerida es habitualmente pequeña, poco más que la compra del material necesario.
- Suelen ser empresas muy jóvenes (menos de 10 años), de pequeño tamaño y con personal titulado por la federación correspondiente.
- Por lo que respecta al perfil de los practicantes de los deportes de aventura ofertados por este segmento, podemos observar la siguiente tabla:

| VARIABLES | CARACTERÍSTICAS PRINCIPALES |
|-------------------------------------|---|
| Demográficas | |
| Edad | Entre 18-40 años |
| Sexo | Masculino, sobre todo, con tendencia creciente género femenino |
| Situación familiar | Jóvenes solteros, jóvenes casados sin hijos y adultos de edad media casados y con hijos |
| Ingresos | Nivel económico medio-alto |
| Nivel educativo | Nivel medio-alto. |
| Nivel social | Bajo superior, medio y alto inferior. Estudiantes y profesionales medios |
| Geográficas | |
| Zona de procedencia | Núcleos de más de 100.000 habitantes |
| Mercados emisores | Principalmente turismo nacional (más del 70%). Alemanes, franceses y británicos acaparan el mercado extranjero. |
| Distancia | Destinos próximos a la zona de influencia el destino, normalmente dentro de la misma Comunidad |
| Psicográficas | |
| Estilo de vida | Activo, busca posición, autorrealización y protagonismo en la práctica viajera. El ocio es tan importante como el trabajo. |
| Personalidad | Inquieta, decidida, con ambición |
| Beneficios buscados | Práctica de deportes al aire libre con una cierta mitología de retorno y contacto con la naturaleza, evasión, descubrimiento, prestigio y experiencia singular |
| Posición del turista | Predomina el turista como sujeto activo de la experiencia vivida, al que no le importa la incertidumbre que puede generar la práctica deportiva al aire libre. |
| Grado de repetición | Entre 5-10 % |
| Características vacacionales | |
| Destinos preferidos | Prefiere visitar nuevos destinos y experimentar nuevas sensaciones, pero de manera un tanto selectiva. Espacios naturales protegidos y parajes de alto valor paisajístico o buenas condiciones medioambientales. |
| Frecuencia estacional | Escalona sus vacaciones, evita la temporada alta. Realiza al menos un viaje al año al medio rural y dispone de mayor número de vacaciones que la media |
| Tipo organización del viaje | Autoorganizado, con predominio de prácticas individuales y de carácter individualista, en pequeños grupos de amigos y compañeros, familias o parejas, sin recurrir generalmente a viajes organizados, de paquete turístico e intermediarios comerciales |
| Actividades más desarrolladas | Senderismo, cicloturismo, montañismo. Actividades hedonistas sin reglamentación o normas fijas que no requieran un esfuerzo intenso o una destreza especial y con riesgo controlado o previsible. |
| Mercados emisores | Nacional (en torno a un 70-80%). |
| Tipo de desplazamiento | Por carretera y en vehículo propio o compartido. |

Tabla 13. Perfil aproximativo de los practicantes de deportes de aventura en España. Fuente: Rivera Mateos (2007: 24).

En definitiva, el turismo de aventura se considera como un nicho de mercado cada vez más importante en España y con un pronóstico de crecimiento alto para los próximos años, aunque aún es necesario resolver ciertas debilidades que presenta y legislar para proteger las zonas naturales donde se realizan estas actividades. En el año 2000, se presentó un análisis DAFO del sector (Garau et ál., 2000: 602), donde se pueden apreciar de forma resumida las Debilidades, Amenazas, Fortalezas y Oportunidades (DAFO) de este segmento y que hoy en día puede seguir aplicándose.⁹²¹

| Fortalezas | Debilidades |
|--|---|
| <ul style="list-style-type: none"> - Climatología favorable, especialmente en comparación con otros países europeos. - Espacios naturales terrestres y marítimos de alta calidad. - Facilidad de comunicaciones para los turistas, especialmente los que utilizan chárter. - Infraestructuras y oferta complementaria existente. - Variedad de ofertas posibles. - Creciente oferta de restauración y alojamiento en zonas rurales. - Importante oferta de senderos señalizados e incremento constante con la recuperación de vías verdes. | <ul style="list-style-type: none"> - Actividades estacionales (las actividades náuticas se concentran en los meses más calurosos y las de tierra, en los meses de primavera y otoño). - Sector reciente y poco estructurado. - Imagen de España todavía poco asociada al turismo activo. - Dificultades para que las empresas especializadas actúen en el exterior de manera conjunta. - La oferta de alojamiento y de servicios complementarios es poco competitiva y poco preparada para el turismo internacional (idiomas, flexibilidad horaria, etc.). - Falta de personal cualificado en muchas actividades. |
| Oportunidades | Amenazas |
| <ul style="list-style-type: none"> - Posicionar España como espacio turístico ideal para la práctica del turismo activo (en sus diferentes modalidades) durante todo el año. - Consolidar una oferta complementaria de turismo activo de alta calidad. - Mejorar la información a los turistas existentes de las posibilidades de actividades disponibles. - Fomentar la interrelación entre las empresas existentes para comercializar o promocionar sus productos de manera conjunta. - Aprovechar el auge de la actividad como motivo de viaje como elemento desestacionalizador global. | <ul style="list-style-type: none"> - Saturación de los espacios donde se desarrollan las actividades de turismo activo. - Degradación del medio ambiente. - Crecimiento desordenado del sector con riesgos de deterioro de la calidad ofertada, intrusismo, etc. |

Tabla 14. Análisis DAFO del segmento del turismo de aventura. Fuente: Garau et ál. (2000: 602).

⁹²¹ Hemos actualizado e incluido algunos puntos de este análisis DAFO, a fin de adaptarlo a la situación actual de este sector.

A fin de contrarrestar los aspectos más negativos que presenta este sector, es decir, las debilidades y amenazas incluidas en este análisis DAFO, se ha comenzado a regular este sector desde las administraciones autonómicas, que son las que ostentan la competencia del turismo en España. Una de las principales dificultades que se encuentran los legisladores a la hora de elaborar la regulación pertinente se debe a la imbricación de tres sectores diferentes: uno, el turismo; dos, el deporte, y tres, el medio ambiente, lo que convierte su regulación en un tema muy complicado, influido por diversos aspectos legales, tales como la protección del medio natural donde se realizan las actividades de aventura,⁹²² la regulación de la práctica deportiva, la titulación requerida por los monitores, el registro de empresas turísticas, etc. A pesar de la influencia de estos tres aspectos, hasta el día de hoy la legislación disponible ha venido impulsada solo por consejerías de turismo, aunque incluyendo aspectos relevantes de las otras dos materias.

Desde el año 1991, que tan solo Cataluña contaba con un decreto, ha ido apareciendo normativa en otras comunidades cuyo objetivo era la regulación de este sector turístico, especialmente en lo referente a la normalización de los establecimientos de alojamiento rural⁹²³ y a las actividades de aventura que se pueden practicar, hasta llegar a los actuales doce decretos en todas España:⁹²⁴

- Cataluña (Decreto 81/1991, derogado por el Decreto 56/2003)
- Cantabria (Decreto 31/1997)
- Extremadura (Decreto 120/1998)
- Aragón (Decreto 146/2000 y Decreto 92/2001, unificados por el Decreto 55/2008)
- Galicia (Decreto 42/2001)

⁹²² Este aspecto es extremadamente importante, puesto que la mayoría de las actividades se llevan a cabo en espacios naturales protegidos, como son los parques nacionales, parques naturales, reservas de la Biosfera, etc.

⁹²³ En España, el turismo rural y el turismo de aventura están muy vinculados y la mayoría de los decretos los tratan conjuntamente, como por ejemplo el Decreto de Andalucía (2002).

⁹²⁴ Véase la URL: <http://www.feaav.org/common/mt/compendio/turismo_activo.shtm>, para acceder a los diferentes decretos citados.

- Andalucía (Decreto 20/2002 y Orden de 20 de marzo de 2003)
- Asturias (Decreto 92/2002)
- La Rioja (Decreto 111/2003)
- Navarra (Decreto 288/2004)
- Castilla-La Mancha (Decreto 77/2005)
- Castilla-León (Decreto 96/2007)
- Murcia (Decreto 320/2007)

Los doce decretos tratan el tema desde aproximaciones diferentes: en unos casos, se hace más hincapié en la relación con el turismo rural y con el alojamiento, como es el caso del decreto andaluz; y en otros casos, se enfatiza más el aspecto de proteccionismo del medio natural donde se realizan las actividades turísticas, como es el caso del decreto aragonés. A pesar de presentar ciertas diferencias entre sí, como es el caso de la denominación de las actividades que están regulando («actividades físico-deportivas en el medio natural», en el decreto catalán; «actividades propias de las empresas de turismo activo», en el decreto gallego; «actividades de turismo activo y de aventura», en el decreto aragonés; «actividades de turismo activo», en el decreto andaluz y en el decreto de Castilla-La Mancha, etc.), comparten algunos aspectos entre ellos, como es el hecho de partir de una consideración del producto ofertado muy parecida, fundamentada en dos aspectos: a) prácticas que se realizan sirviéndose de los recursos que ofrece la naturaleza y b) presencia de cierto riesgo en la práctica de la misma (a esto unen el decreto aragonés, gallego y andaluz la demanda de cierto grado de esfuerzo físico o destreza para llevar a cabo la práctica de las actividades) (Gil Luque, 2003: 140). Asimismo, todos comparten la denominación que le otorgan al segmento, a saber: «turismo activo», ya que en él incluyen tanto los deportes de aventura como los deportes de invierno (como el esquí), deportes náuticos (como la vela) o actividades cinegéticas (como la caza), es decir, todas las actividades activas que se pueden llevar a cabo en un entorno natural.

Otro aspecto que comparten todos los decretos autonómicos es el objetivo por regular la ordenación y actuación de las empresas turísticas que se dedican a este

segmento, así como las normas mínimas de seguridad que deben cumplir (homologación de material, titulación de los monitores, seguros de responsabilidad civil, etc.). De esta manera, pretenden establecer los pasos y requisitos necesarios para constituir una empresa de este tipo y para que pueda llevar a cabo sus actividades dentro del marco legal.

Por último, la mayoría de estos decretos incluyen una lista con las actividades turísticas propias de este segmento, aunque, debido a la rapidez con la que aparecen nuevas actividades y evolucionan las existentes, no se pueden considerar completas ni definitivas. A continuación, presentamos una tabla con las listas de las actividades de turismo activo contenidas en algunos de los decretos en vigor.

| ANDALUCÍA (2002) | ARAGÓN (2008) | GALICIA (2001) | CATALUÑA (2003) | MURCIA (2007) |
|---|---|---|---|--|
| 1. Bicicleta de montaña. 2. Buceo o actividades subacuáticas. 3. Descenso de barrancos. 4. Descenso en bote. 5. Escalada. 6. Esquí de río. 7. Esquí acuático. 8. Esquí alpino. 9. Espeleología. 10. Globo aerostático. 11. Heliesquí. 12. Heliexcursión. 13. Hidrobob. 14. Hidrotrineo. 15. Hidropedales. 16. Mushing. 17. Montañismo. 18. Motos de nieve. 19. Motos acuáticas. 20. Navegación a vela. 21. Paracaidismo. 22. Piragüismo. 23. Quads. 24. Turismo ecuestre. 25. Salto desde el puente. 26. Salto con elástico. 27. Senderismo. 28. Surf y windsurf. 29. Todoterreno con motor. 30. Travesía. 31. Vuelo libre. 32. Vuelo con ultraligero. 33. Vuelo sin motor. | 1. Actividades Subacuáticas: 1.1. Escafandrismo. 2. Actividades náuticas: 2.1. Navegación a vela. _ Vela ligera. _ Tabla de desplazamiento a vela o windsurf. 2.2. Piragüismo. 2.3. Rafting. 2.4. Hidrospeed. 2.5. Navegación a motor: - Esquí náutico. - Paseos y excursiones en embarcaciones con motor. - Motos de agua. 3. Turismo ecuestre y similares: 3.1. Hípica. 3.2. Paseos en otros animales. 4. Ciclismo: 4.1. Bicicleta de montaña. 4.2. Ciclismo de paseo o por carretera. 5. Actividades aéreas: 5.1. Ala Delta. 5.2. Parapente. 5.3. Vuelo a vela. 5.4. Vuelo con ultraligero. 5.5. Vuelo en globo. 6. Actividades de montaña y escalada: 6.1. Montañismo. 6.2. Barranquismo. 6.3. Escalada en hielo. 6.4. Escalada en roca y rocódromo. 6.5. Senderismo. | - Parapente - Ala delta - Bicicleta todo terreno - Descenso de barrancos - Descenso en bote - Hidrotrineo - Marcha a caballo - Piragüismo - Salto desde el puente - Salto con elástico - Senderismo - Escalada | - Descens en barrancs i engorjats. - Espeleobusseig (espeleologia subaquàtica continental). - Salt de pont. - Salt elàstic. - Ciclocròs. - Marxa a cavall. - Quads. - Quatre per quatre. - Recorregut de bosc o de camp. - Trial. - Enduro. - Descens en bot (rafting). - Hidrotrineu (hydrospeed). - Piragüisme. - Bus-bob. - Esquí nàutic. - Fly-surf. - Motonàutica. - Parapent d'arrossegament. - Rem. - Surf. - Surf a vela. - Vela. - Wake Board. - Ala de pendent (Parapent). - Ala delta. - Baló aerostàtic (Aerostació, Baló). -Paracaigudisme. - Paramotor. - Ultralleugers. - Vol a motor. - Vol a vela. - Motos de neu. - Trineu amb gossos (Múixing). | 1.- Actividades subacuáticas: 1.1.- Escafandrismo: 1.2.- Snorkel: 2.- Actividades náuticas: 2.1- Navegación a vela. 2.1.1.- Vela ligera: 2.1.2.- Vela crucero: 2.1.3.- Windsurf: viento. 2.1.4.- Surf. 2.1.5- Rutas en barco a motor. 2.2.- Piragüismo: 2.2.1.- Kayak. 2.2.2.- Canoa. 2.2.3.- Rafting. 2.3.- Hidrospeed. 2.4.- Motos de agua. 2.5.- Esquí de río. 2.6.- Esquí acuático. 2.7.- Bodyboard. 2.8.- Wakeboard. 2.9.- Hidropedales. 2.10.- Donutski. 2.11.- Kite surf. 2.12.- Paradescenting. 2.13.- Ski bus. 2.14.- Busbob. 3.- Hípica. 3.1.- Turismo ecuestre. 4.- Ciclismo: 4.1.- Bicicleta de montaña. 4.2.- Cicloturismo. 5.- Actividades aéreas: 5.1.- Ala Delta. 5.2.- Parapente. 5.3. Velero (vuelo sin motor). 5.4.- Vuelo con ultraligero. 5.5.- Globo aerostático. 5.6.- Heliexcursión (excursión en helicóptero). 5.7.- Paracaidismo. 5.8.- Puenting. 5.9.- Banging (salto |

| | | | | |
|--|---|--|--|--|
| | <p>6.6. Travesía o Trekking. 6.7. Vías ferratas. 6.8. Puenting. 6.9. Circuito multiactividad. 7. Actividades de orientación. 8. Espeleología. 9. Tiro con arco. 10. Actividades de nieve: 10.1. Esquí alpino. 10.2. Esquí de fondo. 10.3. Heliesquí. 10.4. Esquí de montaña. 10.5. Esquí de fondo de paseo. 10.6. Esquí fuera de pistas. 10.7. Excursiones con raquetas de nieve. 10.8. Mushing. 10.9. Motos de nieve. 10.10. Escuelas de esquí. 11. Actividades con vehículos de motor: 11.1. Todoterreno con motor. 11.2. Quads. 11.3. Otros vehículos de motor. 12. Paintball y similares. 13. Otras actividades en la Naturaleza: 13.1. Turismo de caza. 13.2. Turismo de pesca.</p> | | | <p>con elástico). 5.10.- Vuelo con motor. 6.- Actividades de montaña y escalada: 6.1.- Montañismo. 6.2.- Barranquismo. 6.3.- Escalada en roca y rocódromo. 6.4.- Rápel. 6.5.- Tirolina. 6.6.- Alpinismo. 6.7.- Puente tibetano. 6.8.- Puente mono. 7.- Rutas a pie: 7.1.- Senderismo. 7.2.- Trekking. 7.3.- Itinerarios didácticos e interpretativos. 8.- Orientación 8.1.- Carrera de orientación. 8.2.- Actividad de orientación. 9.- Taller de medio ambiente y naturaleza. 10.- Ruta cultural. 11.- Agroactividad. 12.- Ruta temática. 13.- Espeleología. 14.- Tiro con arco. 15.- Actividades de nieve: 15.1.- Esquí de fondo. 15.2.- Esquí de travesía. 15.3.- Excursiones con raquetas de nieve. 16.- Rutas con vehículos a motor: 16.1.- Ruta en vehículo 4x4. 16.2.- Quads. 17- Circuitos o programas de multiactividad o multiaventura. 18- Paintball.</p> |
|--|---|--|--|--|

Tabla 15. Lista de actividades de turismo activo. Elaboración propia a partir de diversos decretos autonómicos.

A partir de estos decretos que hemos incluido como ejemplos, podemos realizar varias observaciones:

- Las actividades que se incluyen aquí hacen referencia a las actividades activas que se pueden realizar en la naturaleza (como veíamos anteriormente), pero no todos los decretos incluyen las mismas actividades, lo que da a entender que a) no existe una homogeneización en la consideración de las actividades que pertenecen al turismo activo en España y b) no todas las actividades tienen la misma relevancia en todas las comunidades autónomas, es decir, cada decreto incluye únicamente las actividades más practicadas en su región.
- Estos decretos tratan el turismo de aventura como parte del turismo activo y, al mismo nivel que otros segmentos turísticos, como son el turismo de nieve o el turismo náutico, lo que nos dificulta la distinción entre las actividades de aventura y el resto de actividades consideradas también activas.
- La presentación de las diferentes actividades se realiza de forma diferente, ya que unos decretos presentan las actividades en forma de lista, sin realizar ningún tipo de clasificación, como ocurre con los decretos de Andalucía o de Cataluña; mientras que otros las clasifican según ciertas características (el medio natural en el que se practican, el tipo de vehículo, etc.), como es el caso del decreto de Murcia o del de Aragón. De nuevo, vemos que no existe un consenso entre las diferentes administraciones autonómicas.

En nuestro trabajo de investigación, estos decretos, especialmente las listas que incluyen sobre las actividades del turismo activo, nos servirán de gran ayuda para concretar nuestra conceptualización y nuestra ontología, así como para realizar un análisis *in vitro* de los términos y sus definiciones, con objeto de establecer nuestras relaciones semánticas y posteriores definiciones.

Una vez vistas las características que presenta el segmento de aventura en España, así como el segmento del turismo activo en general según los diferentes decretos, pasaremos a estudiar las clasificaciones que se han realizado de sus deportes de aventura, con el fin de continuar con el análisis de este dominio para nuestro trabajo ontoterminográfico.

3.3.3.1.3. Actividades del turismo de aventura y propuestas para su clasificación

En las listas incluidas en los diferentes decretos autonómicos sobre el turismo activo que hemos presentado anteriormente, hemos podido observar dos aspectos relevantes sobre las actividades de este turismo: por un lado, la variedad de actividades que se pueden realizar en cada región; y por otro, las diferentes denominaciones para nombrar la misma actividad (por ejemplo, «descenso de barrancos, en el decreto andaluz, y «barranquismo», en el decreto aragonés). Este segundo aspecto, es decir, la existencia de varias opciones para denominar una misma actividad es muy característico de este segmento turístico, ya que tanto su reciente aparición como su continua evolución dan lugar a constantes neologismos (tanto por nueva acuñación como por modificación de términos anteriores). Esto da lugar a dos casos diferentes: por un lado, existe diversas variantes denominativas que hacen referencia al mismo concepto y que se pueden utilizar indistintamente, como por ejemplo «rafting» y «descenso en bote»; y por otro, a las variantes que hacen referencia a conceptos diferentes pero que se utilizan como sinónimos, como es el caso de «alpinismo» y «montañismo». En estos casos, el traductor debe estar muy atento y no caer en el error.

En este apartado, no trataremos esta variación, puesto que se verá más adelante en otros apartados y, profundamente, en el último capítulo con relación a la extracción terminológica (semi)automática (cf. 4.2.1.3.), aunque sí considerábamos relevante destacarla en este punto.

Hasta ahora hemos visto que en el segmento de aventura existen grandes controversias en cuanto a varios aspectos, como son la denominación y el marco conceptual del propio segmento, la denominación de las actividades que se realizan dentro de este segmento, etc. Además, otra dificultad que se encuentra aquí es la clasificación de las actividades, que trataremos a continuación a través de la revisión de diferentes clasificaciones que se han elaborado hasta el momento. Mediante este estudio, trataremos de profundizar en el conocimiento de este segmento turístico para elaborar nuestra ontología en los pasos siguientes.

Actualmente, existe un enorme abanico de actividades de aventura que aumentan constantemente ante el surgimiento de nuevas prácticas y, a lo largo de los años, han ido surgiendo numerosas clasificaciones que han pretendido agrupar y ordenar las diferentes actividades según diferentes criterios, como son por ejemplo: *la finalidad o propósito* que se busca con la actividad, distinguiéndose actividades recreativas, educativas y competitivas; *el elemento geográfico implicado*, se diferencian actividades de tierra (incluyendo las realizadas con animales y las que demandan nieve y hielo), de agua y de aire; *según el tipo de acciones*, se distinguen actividades individuales, imbricadas (que se realizan de manera individual pero que en determinados momentos necesita de la ayuda de los compañeros) y colectivas; según la *dificultad* de la actividad, se distingue entre actividades blandas (que no requieren de gran esfuerzo físico, como el cicloturismo o el senderismo) y actividades duras o extremas (que requieren de una forma física determinada y de práctica, como son la escalada o el alpinismo), entre otros.

De las clasificaciones que se han elaborado hasta el momento, el elemento geográfico implicado ha sido el criterio más utilizado, considerándose en unos casos como punto clave de la clasificación y en otros como criterio único. Como ejemplos de clasificaciones de este tipo, destacamos las siguientes:⁹²⁵

- Á. Miracle (1994) desarrolla una clasificación atendiendo al medio físico donde se desarrollan las actividades (agua, aire y montaña).
- F. Funollet (1995) propone una clasificación compleja basada en dos tipos de determinantes: por un lado, los determinantes de la actividad deportiva en el medio natural; y por otro, los determinantes didácticos de las actividades deportivas en el medio natural. Los primeros hacen referencia a la trayectoria (bidimensional o tridimensional), al plano (horizontal o vertical), al elemento (estable o inestable), al contacto (directo o indirecto), al desplazamiento (caminando, rodando, deslizando, cabalgando, nadando, flotando, escalando, rapelando, sumergiendo, volando). Y los segundos se refieren a la energía

⁹²⁵ Las clasificaciones que se presentan a continuación han sido extraídas de Gil Luque (2003), Peñalver Torres (2004) y Sáez Padilla (2005).

- (autogenerada o generada), acción (individual, imbricada o en equipo), ecosistema (pelágico, fluvial, lacustre, embalse, ciénaga, rupícola, cavernícola, forestal, arbustivo, praderativo, agrícola, desierto frío, desierto árido, alta montaña, urbanizado).
- Olivera y Olivera (1995) plantean la clasificación de actividades partiendo del entorno físico o medio donde se desarrollan (tierra, agua y aire) para seguir clasificándolas en el plano en el que esta se practica (vertical u horizontal), el entorno personal del participante (hedonistas o ascéticas o la combinación de ambos), el impacto medio ambiental (individual, en grupo con colaboración o en grupo sin valoración).
 - Page (1997) presenta una clasificación según el elemento natural donde se realizan las actividades (agua, tierra y aire).
 - Guillén et ál. (2000) se basan inicialmente en el medio geográfico implicado en la práctica de la actividad para su clasificación, aunque después realizan una segunda distinción en función del tipo de acciones que realizan las personas que las practican.
 - Peñalver Torres (2004) basa su clasificación también en el medio físico donde se realizan las actividades (tierra, agua y aire) y, dentro de esta clasificación, las agrupa según su afinidad en cuanto a los materiales, las técnicas y los entrenamientos empleados.

Como ejemplo de clasificación que incluye el criterio de clasificación más aceptado, a saber: el medio en el que se realizan las actividades, presentamos a continuación la propuesta de Peñalver Torres (2004: 185-198). Asimismo, permite observar la mayoría de las actividades de aventura más ofertadas en la actualidad.⁹²⁶

⁹²⁶ Por cuanto respecta al medio terrestre, se observan algunas actividades que se podrían considerar dentro del turismo de nieve, a saber: esquí de travesía o *randonée* y esquí de fondo o marcha. En este caso, estas actividades se pueden considerar también como parte del turismo de aventura, puesto que no requieren de instalaciones específicas como otras actividades del turismo de nieve, como son el esquí alpino o el snowboard.

| TIERRA | AGUA | AIRE |
|--|--|--|
| <ul style="list-style-type: none"> - Montañismo. Modalidades: <ul style="list-style-type: none"> - Área de marcha - Senderismo - Media montaña - Alta montaña - Expediciones - Área de escalada <ul style="list-style-type: none"> - Escalada en roca - Escalada en hielo - Escalada deportiva - Área de resistencia <ul style="list-style-type: none"> - Duatlón en montaña - Área de específicos <ul style="list-style-type: none"> - Barranquismo - Esquí de travesía o randonée <ul style="list-style-type: none"> - Orientación - Cicloturismo - Ciclismo de montaña o mountain bike <ul style="list-style-type: none"> - Espeleología - Esquí de fondo o marcha <ul style="list-style-type: none"> - Travesía (telemark) - Todoterrenos o 4x4 <ul style="list-style-type: none"> - Karting - Quads - Paseos a caballo - Tiro con Arco <ul style="list-style-type: none"> - Paintball - Outdoor Training - Bungee Jumping - Puenting | <ul style="list-style-type: none"> - Remo - Piragüismo. Modalidades: <ul style="list-style-type: none"> - Canoa - Kayak - Submarinismo <ul style="list-style-type: none"> - Rafting - Descenso de ríos <ul style="list-style-type: none"> - Hidrosped - Surf - Esquí acuático o náutico. Modalidades: <ul style="list-style-type: none"> - Wake Board, - Knee board, - Donut esquí, Ski bus, etc. <ul style="list-style-type: none"> - Kite Surf - Moto acuática | <ul style="list-style-type: none"> - Ultraligero - Parapente - Vuelo con motor <ul style="list-style-type: none"> - Volar en globo - Ala Delta - Paracaidismo. Modalidades: <ul style="list-style-type: none"> - Salto en tandem, - Bautismo aéreo, <ul style="list-style-type: none"> - free-style, - free-fly, - chute-assis, - head-down... - Vuelo sin motor |

Tabla 16. Elaboración propia a partir de la clasificación de Peñalver Torres (2004: 185-198).

En nuestro caso, también utilizaremos el criterio del medio físico en el que se practica una actividad como elemento clave para realizar nuestra propia clasificación (cf. Tabla 37), aunque también consideraremos el material que se emplea en cada actividad y el tipo de acción que se realiza en cada una de ella, ya que, en nuestra opinión, son los elementos básicos que intervienen en cualquier actividad. Y con respecto a la denominación de las actividades recogidas en las diferentes clasificaciones y decretos, realizaremos un estudio comparativo y determinaremos los términos que se pueden considerar como variantes denominativas y los que no, a la par que indicaremos el término o los términos de uso preferente.

Por último, cabe destacar que, a pesar de que esta clasificación basada en el medio geográfico es la más utilizada y, a simple vista, parece completamente objetiva, hemos encontrado cierta discrepancia en el tratamiento de algunas actividades, como es

el caso de «barranquismo» o «descenso de barrancos». Según se puede observar en la tabla anterior, esta actividad se incluye dentro del grupo de «Tierra», y lo mismo ocurre en la clasificación que realiza el Decreto de Aragón, de 2008, y el Decreto de Murcia, de 2007, así como en la clasificación de Vázquez (2002) y de Pascau Canales (1998), entre otros, donde en todos se considera el barranquismo como actividad terrestre. Sin embargo, otros autores defienden la inclusión de esta actividad dentro de las actividades acuáticas, como por ejemplo Olivera y Olivera (1996) y Fernández Arroyo (1996). De nuevo, encontramos la constante falta de homogeneización de este sector y discrepancias con aspectos básicos como puede ser este, lo que nos lleva de nuevo a la necesidad de tomar nuestras propias decisiones para nuestro trabajo de investigación. En nuestro caso, trataremos el barranquismo como una actividad terrestre por ser el medio físico principal donde se realiza la actividad, aunque consideramos imprescindible hacer mención a su doble naturaleza, ya que muchos de los tramos por donde se realiza están directamente relacionados con el agua (cataratas, ríos, pozas de agua, etc.).

Una vez que hemos profundizado en el sector turístico en su conjunto y, posteriormente, hemos realizado una incursión en el segmento del turismo de aventura (dominio de nuestro trabajo de investigación), pasaremos al estudio lingüístico de este dominio y, así, trataremos el discurso turístico como lenguaje de especialidad, determinaremos los principales elementos que lo caracterizan y terminaremos realizando un análisis de los recursos terminológicos disponibles para este segmento turístico. Con este estudio, pretendemos seguir ampliando nuestro conocimiento sobre el dominio del turismo, a fin de realizar un trabajo ontoterminográfico fiable, completo y documentado.

3.3.4. DISCURSO TURÍSTICO, LENGUAJE DE ESPECIALIDAD

Hasta hace relativamente poco tiempo, el discurso turístico no se había considerado como un lenguaje de especialidad y, por ello, no se le había prestado demasiada atención. A pesar de que investigadores como Cohen (1972), Cohen y Cooper (1986), Goffman (1974), Schudson (1979) hubieran tratado aspectos lingüísticos del turismo, concretamente sociolingüísticos, este no se ha tratado con la

seriedad ni con la profundidad requerida y no se ha considerado un lenguaje de especialidad hasta hace muy poco tiempo. A este respecto, y como nos indican Dann (1996) y Nigro (2006), podemos decir que el lenguaje del turismo cuenta con escasos estudios sistemáticos.

El turismo se ha estudiado desde diferentes aproximaciones: sociológica, antropológica y económica, entre otras, pero solo en los últimos años ha aumentado la percepción del turismo como lenguaje de especialidad, lo que ha dado lugar al comienzo de estudios desde un punto de vista lingüístico. Esto se debe principalmente al hecho de que el discurso parece, a simple vista, un lenguaje cercano y fácil de comprender, pues nos lo encontramos a diario y en cualquier parte (Dann, 1996) y, además, gran cantidad de su terminología proviene de otras disciplinas cercanas. En este contexto, algunos autores consideran el turismo como un lenguaje complicado de describir y de delimitar nítidamente, puesto que no cuenta con un contenido claramente definido ni con unas funciones textuales con límites claros debido a la influencia que recibe de otras disciplinas como son la historia, la geografía, el arte, etc. La alta interdisciplinariedad del discurso turístico, además de la presencia de una comunicación con receptores no especialistas hace que la especialización del discurso turístico no haya sido considerada como tal hasta hace relativamente poco tiempo. Sin embargo, los usos del lenguaje que se dan en este discurso, tanto lingüístico como no lingüístico, constituyen un tipo de comunicación especial y diferente de cualquier otra comunicación, puesto que cuenta con unas características y unas funciones propias que le confiere una identidad estructural y léxico-semántica que lo diferencia de otros lenguajes de especialidad, como puedan ser el económico, el jurídico, el político, etc. (Suau, 2006: 1).

Actualmente, podemos confirmar un consenso creciente dentro de la comunidad académica, que considera el discurso turístico como un lenguaje especializado, y, como tal, se considera un producto de la actividad comunicativa del ámbito del turismo.

3.3.4.1. Características del lenguaje del turismo como lenguaje de especialidad

[En las traducciones de textos turísticos] no se traducen lenguas, sino textos impregnados de referencias extralingüísticas, contruidos a partir de contenidos conceptuales cognitivos cargados de valor comunicativo (Jung, 2000: 56).

En la definición de turismo aportada por la OMT que veíamos al comienzo de este capítulo, se establecía el turismo como una actividad económica que ofrecía servicios a personas que se desplazan fuera de sus lugares habituales de residencia, así como los desplazamientos de dichas personas durante sus viajes. De esta forma, y partiendo de esta definición, se puede considerar el turismo como una actividad intercultural e interlingüística que “une” personas, culturas, lenguas y lugares mediante la provisión de servicios turísticos. Esta carga cultural, además de lingüística, que incluye el turismo le otorga al discurso turístico una dimensión propia y lo convierte en un lenguaje de mediación, tanto desde el punto lingüístico como desde el cultural. Debido a este carácter interlingüístico e intercultural, la traducción en este ámbito adquiere una posición imprescindible, ya que está presente en todo producto y texto turístico de una zona y sirve de mediación entre los turistas y las personas locales y su cultura.

En nuestra opinión, compartida con autores relevantes de este ámbito (Calvi, 2000; Kelly, 1997; Dann, 1996), las características léxicas, sintácticas y textuales propias del lenguaje del turismo lo convierten en un lenguaje de especialidad y, por tanto, no encontramos ninguna justificación para considerarlo como parte de la lengua general o de otro lenguaje de especialidad, como podría ser el económico, el jurídico o el artístico en algunos casos. Asimismo, debemos tener en cuenta la tipología textual existente en el discurso turístico, entendiendo «texto turístico», según la definición que ofrece Kelly (1997: 35), cualquier texto publicado por una entidad pública o privada de cualquier tipo cuyo objetivo sea a) proporcionar información al turista o b) promocionar un destino (ciudad, hotel, restaurante, etc.) y animar a los receptores para que lo

visiten.⁹²⁷ De esta manera, y puesto que cada género textual dentro de este campo de especialidad presenta unos objetivos comunicativos y funciones concretas que los pueden diferenciar del resto, es necesario especificarlo a fin de proporcionar una descripción más adecuada. Por lo que respecta a la traducción de este tipo de textos, esta variabilidad de géneros turísticos presenta un papel muy importante y, por tanto, será imprescindible que el traductor encargado conozca los géneros textuales y sus características propias para ser capaz de alcanzar buenos resultados en sus traducciones.

Partiendo de una clasificación general de los textos turísticos de cualquier segmento, podemos distinguir dos grupos: por un lado, los textos turísticos promocionales; y por otro, los textos turísticos no promocionales (Kelly, 1997: 34; Brosson, 2008: 10).

En el primer grupo, se encuentran los textos turísticos cuyo principal objetivo es la venta de productos y servicios en el lugar de destino y, para ello, utilizan un lenguaje directo, breve y persuasivo, lleno de optimismo y de belleza. Entre este tipo de textos, destacan la guía turística, con una función promocional y referencial; el folleto turístico, de carácter esencialmente promocional, publicado por entidades públicas o privadas; los catálogos, que contienen la lista de los programas vacacionales ofrecidos por agencias, touroperadores, cadenas de hoteles, etc. de carácter eminentemente promocional; los reportajes y artículos publicados en revistas especializadas, como *Cuadernos Turísticos* (en español), *International Journal of Tourism* y *Annals of Tourism Research* (en inglés) o *Tourismus Journal* (en alemán), y revistas no especializadas o de promoción, como revistas de líneas aéreas (*Ronda Iberia*, en español) o revistas de puertos (*Aqua Vista*, en español); los anuncios publicitarios y las páginas web de diferente naturaleza (directorios, como *Rural Ibérica*;⁹²⁸ portales, como la página de Turismo Andaluz;⁹²⁹ buscadores, como *Rutas y Viajes.net*,⁹³⁰ etc.) (Calvi, 2006: 272), todos ellos elaborados

⁹²⁷ Traducción propia de la definición original: «[...] any text published by a public or private organisation of any kind intended a) to give information to any kind of visitor or b) to advertise a destination (city, hotel, restaurant, etc.) and encourage visitors to go there».

⁹²⁸ <<http://www.ruraliberica.com/>>

⁹²⁹ <<http://www.andalucia.org>>

⁹³⁰ <<http://www.rutasyviajes.net/>>

por expertos en turismo y dirigidos a un destinatario general que está organizando sus vacaciones o ya se encuentra en el lugar de destino.

En el segundo grupo, encontramos los textos no promocionales y los «textos operativos para el turista» (Kelly, 1997: 35). En el primer caso, encontramos los reportajes, tanto en prensa escrita como en televisión o radio; las guías, ya sean estas culturales o prácticas en formato papel o electrónico; las audio-guías de museos o de otros lugares de interés; los mapas y planos de ciudad; las cartas de los bares y restaurantes; las condiciones generales de cualquier contrato turístico (alojamiento, transporte, viaje combinado, actividad de aventura, seguro turístico, etc.), etc.⁹³¹ Por su parte, el segundo caso se trataría de textos producidos por autoridades (aduana, policía, aeropuerto) y las instrucciones, recomendaciones de diferentes empresas del sector.⁹³² En general, el principal objetivo de estos textos deja de ser prioritariamente de persuasión para imperar el aspecto informativo sobre algún aspecto del lugar de destino.

En este contexto, es muy importante tener en cuenta que las páginas web se han convertido en un texto turístico promocional muy frecuente en la actualidad, especialmente durante la fase previa del viaje, puesto que el número de personas que las utilizan tanto para conocer y elegir lugares de destino como para contratar productos y servicios ha crecido exponencialmente en los últimos años (Tabla 17).

⁹³¹ A pesar de que estos textos son considerados como textos no promocionales, las guías turísticas, e incluso los mapas, incluyen a menudo un lenguaje similar al de los textos promocionales, en cuanto se refiere a la descripción de lugares, monumentos, etc., puesto que emplean un lenguaje lleno de positivismo, de adjetivos que indican la belleza, la autenticidad, la originalidad de lo que estén describiendo, lo que puede llegar a confundirlos con los textos promocionales. No obstante, su principal función es informativa y, así, su función apelativa no ocupa un lugar preponderante como en el otro grupo.

⁹³² Para profundizar en el tema de géneros textuales del sector turístico, véase Druetta (2007) y Mira Rueda (2008: 133-138).

| | Porcentaje de turistas | Tasa de variación interanual |
|------------------------|------------------------|------------------------------|
| Total | 100% | -9% |
| No usa Internet | 40% | -21% |
| Usa Internet | 60% | 1% |
| Consulta | 58% | 4% |
| Transporte | 40% | 3% |
| Alojamiento | 24% | 6% |
| Actividades | 50% | 5% |
| Reserva | 50% | 5% |
| Transporte | 46% | 7% |
| Alojamiento | 31% | 6% |
| Actividades | 6% | 10% |
| Pago | 46% | 6% |
| Transporte | 44% | 7% |
| Alojamiento | 23% | 10% |
| Actividades | 4% | 12% |

Tabla 17. Uso de Internet por parte de los turistas que visitaron España en 2009 (Ministerio de Industria, Turismo y Comercio, 2010b: 53).

Este acceso generalizado a Internet ha dado lugar, a su vez, a nuevas posibilidades de comunicación, con las que, por un lado, los géneros tradicionales se han vuelto más persuasivos, con mayor uso de recursos icónicos, más interactivos, más directos, etc.; y por otro, han surgido nuevas formas de comunicación donde los emisores son los mismos turistas que transmiten sus opiniones, recomendaciones, etc. a través de foros de viajeros, *blogs*, entre otros con una función principalmente informativa (Calvi y Bonomi, 2008: 185). De esta manera, presenciamos actualmente los cambios que se están produciendo en los géneros tradicionales promocionales y el surgimiento de un género nuevo no contemplado hasta el momento, donde la comunicación se produce de turista a turista, es decir, una comunicación no especializada. Tal y como afirma Dann (2007), este impulso de la Red y este nuevo género en el discurso turístico acabará menoscabando el papel de los expertos en la transmisión de información: «the traditional guidebook containing printed information

about a given place, by its own admission, may well be out-of-date from the very momento it is published» (2007: 15).

Asimismo, debemos tener en cuenta la elevada presencia de textos híbridos o multitextuales (Reiss, 1991) en el ámbito del turismo, esto es, textos que combinan en su estructura varios tipos de texto. De esta manera, un texto turístico, por ejemplo un folleto turístico, puede incluir una poesía, un menú gastronómico, resúmenes históricos y publicidad, todo ello combinado correctamente y de forma clara y precisa. En este sentido, no se puede determinar siempre de forma clara y precisa el género textual en cuestión que se está tratando y que, en su caso, se va a traducir.

El barrio de Gràcia

Un paseo por Barcelona nos permite completar una versión a pequeña escala de la 'vuelta al mundo'. Así, comenzando en Gràcia, un rápido vistazo nos descubre un barrio diverso en nacionalidades y en oferta lúdico-cultural. La calle Verdi es el centro neurálgico de esta antigua vila que, antes de su anexión a la gran Barcelona a principios del siglo XX, era el destino dominical para una parte de la rica burguesía barcelonesa.

Sobre todo a partir de media tarde, la calle de Verdi se convierte en un hervidero de gente abierta de miras y ávida de nuevos conocimientos. La oferta culinaria de esta calle es extensa. Los restaurantes de comida tradicional catalana se disputan la clientela con bares como el libanés Al-Waha o el egipcio Nut, espacios donde se puede hincar el diente a los típicos shawarma de carne de cordero y ternera o donde tomar un té, endulzado todo ello con baklawa (pequeños pastelitos). En la misma calle, en Extensions, dos jóvenes africanas hacen gala de su destreza como peluqueras para ofrecer un atrevido cambio de imagen. ¿La especialidad de la casa? Las rastas. Aunque para encontrar auténtico arte africano, mejor caminar unos metros más hasta llegar a Senkeur, en la calle de Torrijos. Allí, una pareja afro-catalana, a los que no es extraño encontrar en la misma calle tomando el té recostados sobre unas tumbonas, vende figuras y cerámica importadas del continente negro.⁹³³

Habida cuenta de la variación de géneros textuales que existe en el ámbito turístico, en nuestro trabajo de investigación nos centraremos únicamente en el primer grupo, a saber: los textos promocionales; y sobre ellos, realizaremos la caracterización que presentamos a continuación, siguiendo el esquema de planos (plano lingüístico, plano

⁹³³ <<http://www.bcn.es/turisme/castella/turisme/welcome.htm>>

pragmático y plano funcional) que utilizamos en el primer capítulo para tratar los lenguajes de especialidad (cf. 1.2.2.1., 1.2.2.2. y 1.2.2.3). A pesar de tratarse de una categoría general de textos turísticos, todos los tipos textuales que se incluyen en ella comparten las mismas características en cuanto a los tres planos que vamos a analizar, aunque varían en cuanto al formato de presentación y la longitud de los textos que lo forman.

3.3.4.1.1. Plano lingüístico

En este plano, podemos distinguir varios rasgos lingüísticos típicos del lenguaje turístico relacionados con su terminología, su sintaxis, su estilo y sus recursos estilísticos.

Como en cualquier lenguaje de especialidad, el discurso turístico se caracteriza principalmente por su terminología, que lo diferencia de otros lenguajes de especialidad. En ella, se refleja la variedad de los sectores profesionales relacionados con la industria del turismo y con la heterogeneidad de los componentes temáticos implicados en la descripción del producto turístico (Mira Rueda, 2008: 41).

En este nivel léxico, se pueden distinguir dos tipos de unidades terminológicas: por un lado, los términos específicos del sector turístico con una especialización fuertemente marcada, relativos a las organizaciones turísticas («agencia de viaje», «tour operador», etc.), a las estructuras («hotel», «restaurante», etc.), a los servicios («traslado», «pernoctación», «charter», etc.), a los aspectos profesionales («operador», «agente», «guía turístico», «animador», etc.) y a los diferentes segmentos turísticos que existen, como es el caso del segmento de turismo de aventura y los términos «multiaventura» y «team building». Por otro lado, se encuentran las unidades que pertenecen a la lengua común y que adquieren un significado especial en el lenguaje turístico, como es el caso de «turista», «alojamiento», «reserva», «viaje combinado», etc. Dentro del bloque terminológico, se incluyen también términos específicos pertenecientes a campos relacionados con el turismo, como son el campo del arte, de la arquitectura, la gastronomía, el deporte, etc., que aportan terminología al campo del

turismo y una unión entre estas disciplinas y el turismo, que a veces hace difusa la diferenciación entre ellos.

El grupo de unidades pertenecientes a la lengua común se utiliza sobre todo en la comunicación dirigida al público no especializado y suele tratarse de adjetivos calificativos que connotan la fascinación del lugar turístico con una función valorativa y afectiva («inmejorable», «espléndido», «magnífico», «wunderbar», «unvergleichlich», «einzigartig», «appealing», «captivating», «bewitching») y dan lugar a combinaciones recurrentes como «idílicas playas», «selvas inexploradas» o «breathtaking images». Algunos verbos de movimiento, como «venir», «escapar», «explorar», «kommen», «besuchen», «erkunden», «discover», «escape», «explore», son clave en los textos turísticos promocionales y en los anuncios publicitarios, ya que intentan atraer al lector a su lugar de destino. Asimismo, el uso de expresiones relacionadas con la localización espacial en la descripción de itinerarios («comenzamos el descenso hacia...», «nos dirigimos a...», «proseguimos nuestro camino a través de...», «start off at...») es muy frecuente en los textos turísticos.

A su vez, el lenguaje turístico presenta gran número de sinónimos o cuasisinónimos, como pueden ser los términos relacionados con el alojamiento («hotel», «parador», «hospedería», «posada», «hostal»)⁹³⁴ con la designación de algunos segmentos turísticos («turismo alternativo», «turismo activo», «turismo verde», «turismo sostenible», etc.) o con servicios turísticos y actividades ofertadas («trekking» y «senderismo») en los que en algunos casos no queda clara su denominación y pueden dar lugar a errores semánticos como es el caso del ejemplo de segmentos o el de las actividades de aventura que hemos indicado.

En este lenguaje especializado, hay gran abundancia de términos culturales y una estrecha relación con las tradiciones locales (Calvi, 2006: 273), que son los

⁹³⁴ En este caso, existe una clasificación de los establecimientos hoteleros basada en el Real Decreto 1634/1983 (Mesa del turismo, <http://www.mesadelturismo.com/common/mt/compendio/legislacionsectorial/hoteles/estatal/rdec_1634-83_esp.shtm>) para los criterios generales, y en las leyes autonómicas (Mesa del turismo, <http://www.mesadelturismo.com/common/mt/compendio/hoteles.shtm>), para la definición de las distintas tipologías.

llamados *realia* o *culturemas*⁹³⁵ («pueblos blancos», «costaleros», «almazara», «Glühwein», «Inn»). Estas palabras poseen un valor terminológico preciso pero pueden ser empleadas también con un significado valorativo, de acuerdo con las intenciones del texto, para sacarle partido a la trascendencia cultural como elemento de seducción (Mira Rueda, 2008: 47). Los textos turísticos que tratan sobre gastronomía o arquitectura presentan mayor número de términos culturales, ya que mantienen una estrecha relación con las tradiciones y la cultura local y, por tanto, la dificultad de su traducción es muy elevada. Es muy frecuente encontrar folletos turísticos u otros textos promocionales con préstamos de la lengua origen, en lugar de traducciones, perífrasis explicativas o equivalentes culturales, con el objetivo de implicar más al lector en el ambiente local, aunque estos a menudo van acompañados de explicaciones entre paréntesis para permitir la comprensión del lector.

Se tiende también a un uso abundante de comparativos y de superlativos, con lo que se le otorga al lenguaje una función apelativa y emotiva constantemente a la vez que se enfatizan las cualidades del producto. La promoción turística explota sistemáticamente esta estrategia y es muy usual en el lenguaje publicitario. El superlativo relativo es uno de los recursos más empleados, aunque a veces se encuentran formas atenuativas y la referencia resulta más genérica. En los siguientes fragmentos, podemos observar varios ejemplos:

(1) Andalucía, es **el principal destino vacacional** de los españoles y **uno de los principales** para los extranjeros.⁹³⁶ (Resaltados en negrita no pertenecen al original).

(2) Mit annähernd 200.000 Kilometern markierter Wanderwege finden sich hier **die besten** Voraussetzungen für ein flächendeckendes Wandervergnügen.⁹³⁷ (Resaltados en negrita no pertenecen al original).

⁹³⁵ Se han aportado diferentes definiciones sobre el término «culturema», pero una de las que nos parece más acertada es la que nos ofrece Vermeer (en Nord, 1997:34): «A cultureme is a social phenomenon of a culture X that is regarded as relevant by the members of this culture and, when compared with a corresponding social phenomenon in culture Y, is found to be specific to culture X. [...] For they may be different in form but similar in function (as in trains vs. cars vs. bicycles) or vice versa (for example, “to have coffee” in England, in the morning vs. “tomar café” in Spain, after dinner vs. “Kaffeetrinken” in Germany, in the afternoon»).

⁹³⁶ <<http://www.andalucia.org>>

⁹³⁷ <<http://www.deutschland-tourismus.de/>>

(3) Set in the beautiful rural Warwickshire countryside, on the banks of the river Avon, it is **one of the most important** tourist destinations in the UK.⁹³⁸ (Resaltados en negrita no pertenecen al original).

Otra forma muy utilizada de expresar la singularidad consiste en utilizar adjetivos como *único* u otras expresiones que niegan la existencia de parangones, así como el empleo verbos que indican la preeminencia dentro de un conjunto, como «destacar» o «distinguirse» y adverbios intensificadores, especialmente en inglés («amazingly», «extremely», «exceptionally»). En los siguientes fragmentos, se pueden encontrar ejemplos de algunas de estas características:

(1) En los Parques Naturales de Grazalema y Sierra de las Nieves y Sierra Bermeja, hay bosques de pinsapos **únicos** en todo el mundo. (Resaltados en negrita no pertenecen al original).⁹³⁹

(2) Ob man die Wälder Nordhessens durchstreift, auf der Mosel-Erlebnis-Route in Rheinland-Pfalz die Vielfalt von Flora und Fauna entdeckt oder **das einmalige** Biosphärenreservat Rhön in Hessen erkundet.⁹⁴⁰ (Resaltados en negrita no pertenecen al original).

(3) The Heart of England is a region steeped in history and heritage. It offers **an unrivalled** variety of countryside, towns, villages and places of interest.⁹⁴¹ (Resaltados en negrita no pertenecen al original).

Actualmente, se está produciendo un uso cada vez mayor de términos relacionados con el medio ambiente y con la ecología, lo que se conoce en alemán como *Ökosprache* (Martí Marco, 2006: 93). En España, se están empezando a utilizar cada vez más términos como «turismo sostenible», «turismo verde», «ecológico», «naturaleza», «protección», «tranquilidad», etc. y se están creando segmentos turísticos relacionados con este concepto ecológico, como el «turismo alternativo», «agroturismo», «ecoturismo», etc. Tanto es así, que estos segmentos relacionados con la naturaleza son los segmentos que se encuentran en auge en la actualidad, como veíamos más arriba. Por su parte, Reino Unido y Alemania (especialmente este último) tienen una tradición ecológica y de reciclaje mucho anterior a la nuestra, por lo que los términos

⁹³⁸ <<http://www.stratford-upon-avon.co.uk/>>

⁹³⁹ <<http://www.andalucia.org/>>

⁹⁴⁰ <<http://www.deutschland-tourismus.de/>>

⁹⁴¹ <<http://www.visitbritain.co.uk/>>

relacionados con este aspecto son más abundantes («Ökoreisen», «Ökotourismus», «naturnah», «verkehrberuhigt», «alternativ» «Begrünung», «eco-tourism», «green tourism», «eco-escape», «eco friendly travel», «eco holidays», etc.). Este cambio de mentalidad hacia lo ecológico ha dado lugar también a la creación de numerosos neologismos recientes en este ámbito, especialmente en lo que respecta al español («vacaciones sostenibles», «vacaciones verdes», por ejemplo).

Relacionado con el nivel léxico, nos encontramos el nivel morfológico, el cual aporta también unos rasgos distintivos al lenguaje del turismo. Los principales procedimientos de creación de este campo de especialidad son cuatro: la derivación, la composición, los préstamos y calcos y las siglas y acrónimos. A continuación, presentaremos brevemente estos cuatro fenómenos en el ámbito turístico, basándonos en la descripción que realiza Calvi (2000) y Mira Rueda (2008: 86):

- La formación de derivados con sufijos y prefijos es muy frecuente en español. Por ejemplo, encontramos sufijos, como -al («vacacional»), -dad («estacionalidad», «confortabilidad»), -ción («fidelización»), -ero («hotelero»), -ismo («turismo», «excursionismo»), y prefijos, como multi- («multipropiedad») y extra- («extrahotelero»). En alemán, también nos encontramos sufijos y prefijos frecuentes: -ung («Entwicklung», «Entdeckung»), -heit («Begebenheit»), -tät («Aktivität», «Lokalität»), -keit («Sehenswürdigkeit»), -schaft («Mitgliedschaft», «Landschaft»), gast- («Gasthaus»), -mit («Mitgliedschaft», «Mitglieder»), entre otros. Y, por último, en inglés, también hay sufijos y prefijos frecuentes: -ing («charming»), -ful («delightful»), -able («fashionable»), -al («cultural»), multi («multiadventure»), entre otros.
- Las siglas y acrónimos se observan en muchos ámbitos del sector turístico, como por ejemplo referidos a entidades, organizaciones y ferias sectoriales: Turespaña, Expovacaciones, CEHAT (Confederación Española de Hoteles y Alojamientos Turísticos), FEAAV (Federación Española de Asociaciones de Agencias de Viajes), IET (Instituto de Estudios Turísticos), Expotur (Feria Nacional de Turismo Rural, Deporte de Naturaleza y Aventura), Fitur (Feria Internacional de Turismo), Turisport (Salón de Turismo, Deportes y Desarrollo Rural), entre otros. En el sector del transporte, encontramos AVE, ICE

(InterCity Express), IC/EC (InterCity/EuroCity), RB (RegionalBahn). La clasificación oficial de los hoteles comprende un preciso repertorio de siglas: H (Hotel), HR (Hotel Residencia), M (Motel), P (Parador), CV (Ciudad de Vacaciones) etc. En impresos, billetes, reservas y documentos análogos las siglas sustituyen ampliamente palabras comunes, nombres de ciudades, aeropuertos, etc. (por ejemplo, AGP = Málaga), según esquemas aceptados internacionalmente. Además, cualquier guía turística dispone de su propia leyenda de siglas y abreviaciones utilizadas en las descripciones de hoteles e itinerarios. También en inglés y alemán son muy comunes este tipo de unidades: «B&B» (Bed and Breakfast), FB (Full Board), AGB (Allgemeine Geschäftsbedingungen), «inkl.» (inklusive), entre otros.

- La formación de términos compuestos se presenta normalmente en la forma sustantivo + sustantivo en las tres lenguas: en español, «aparto-hotel», «trenhotel», «barco-restaurante», «terrazza-mirador», «desayuno buffet», «cicloturismo», «agroturismo», «bono de hotel», «cupo de cama». No obstante, son muy frecuentes también las combinaciones sustantivo + adjetivo como en el caso de «turismo rural», «parque temático», «tiempo compartido», «viaje combinado». Tanto en alemán como en inglés, encontramos multitud de ejemplos de este tipo, especialmente del modelo sustantivo + sustantivo:⁹⁴² «Reiseinformation», «Ferienwohnung», «Hotelschiff», «Stadt-Entdecker», «Urlaubspaket», «Stadtrundgang», «Wohlfühl-Hotel» (en alemán), «youth hostel», «back road», «adventure tourism» (en inglés), etc. Asimismo, también se dan combinaciones de adjetivo+sustantivo y de verbo+sustantivo: «Vollpension», «Doppelzimmer», «Rundfahrt», «Kombiticket», «Sonderpreis», «Einzelleistung», «Schaufenster», «Mietwagen», «Wohnmobil», «Wanderreise», «Badeurlaub» (en alemán), «front office», «wildlife tourism», «cultural tourism», «electronic ticket», «e-ticket» (en inglés), entre otros.
- Los extranjerismos, especialmente los anglicismos, han adquirido gran importancia en el sector turístico, ya sea en modo de préstamos («overbooking»,

⁹⁴² Téngase en cuenta que la formación de compuestos nominales (*Komposita*) en alemán es un procedimiento muy habitual y se pueden encontrar diferentes variantes (cf. Vázquez de Francisco, 2006).

«check in», «check out», «rafting», «duty-free», «paintball», «mountain bike», «team building», «shuttle») o calcos⁹⁴³ («chárter», «camping», «tour operador» «tiempo compartido», «rocódromo», «surfista»). Además de los anglicismos, en el lenguaje del turismo se pueden encontrar términos provenientes del francés («rápel», «arnés», «parapente») y del italiano («tirolina»).⁹⁴⁴ En alemán, encontramos también multitud de estas formas de diferentes procedencias, aunque principalmente anglicismos: «Domizil», «Pension», «Hotel», «Service», etc. En inglés, en cambio, encontramos escasos extranjerismos, debido a que es el inglés la lengua de origen de la mayoría de estos términos. No obstante, encontramos también algunos, sobre todo galicismos, relacionados con la gastronomía: «chef», «sauté», «soufflé», «aperitif», «cafeteria», «carpaccio», etc.

Con respecto a la sintaxis, hay ciertos aspectos del lenguaje del turismo que lo caracterizan en este nivel lingüístico.

En primer lugar, encontramos abundantes formas impersonales que buscan la despersonalización del mensaje, especialmente en catálogos y programas turísticos donde abunda el estilo nominal y las descripciones de lugares, personas, situaciones y servicios. Asimismo, las formas imperativas también adquieren gran importancia en determinados textos turísticos, como los anuncios y los folletos turísticos. No obstante, las formas personales, especialmente la primera persona del plural (*nosotros*) y las dos formas de la segunda persona (*tú* y *usted*) también se utilizan en este tipo de textos especializados, dependiendo estas últimas del grado de acercamiento que se desee imponer en el texto.

(1) **Te invitamos** a pasear por estos lugares emblemáticos y por tradiciones y ritos de contenido mágico y simbólico.

⁹⁴³ La diferencia entre un préstamo y un calco reside en que el primero conserva inicialmente su forma originaria, y por lo tanto es a menudo identificable, mientras que el segundo es la traducción literal de una palabra de otra lengua, por lo que parece una palabra genuina (Cabré Castellví, 1993: 191).

⁹⁴⁴ El sector turístico donde se encuentra una mayor presencia de extranjerismos es el turismo activo o turismo de aventura, ya que se trata de un sector de reciente creación y que está experimentando una mayor expansión. La terminología de este sector se estudiará en el capítulo 4.

(2) **Viaja** a través del tiempo conociendo la Andalucía del pasado, del presente y del futuro. **Sumérgete** en un entramado de itinerarios cuyo trazado **te** descubrirá una tierra singular.

(3) Su historia **te** transportará a siglos de tradición...⁹⁴⁵ (Resaltados en negrita no pertenecen al original)

En alemán, predominan también las oraciones impersonales y las descripciones, al igual que en español. Asimismo, se utiliza la primera persona del plural (*wir*) y la segunda del singular; pero, al contrario que ocurre en los textos turísticos en español, la segunda persona del singular *du* no se suele emplear y se utiliza, en la mayoría de los casos, la forma formal *Sie*. Por su parte, en inglés, predomina también el estilo impersonal, especialmente con oraciones pasivas e imperativos, y la segunda persona (*you*), y al contrario que en español o en alemán, la aparición de la primera persona, ya sea en singular (*I*) o plural (*we*), es muy baja.

(1) Um **unseren** Internet-Service noch informativer und **Ihren** Wünschen entsprechend gestalten zu können, nehmen **wir** gern **Ihre** Ideen und Anregungen auf. Was können **wir** ändern oder ergänzen? **Wir** sind für **Ihre** Hinweise dankbar.⁹⁴⁶ (Resaltados en negrita no pertenecen al original)

(2) Herzlich willkommen! Erleben **Sie** Bonn von seinen schönsten Seiten.²⁵ (Resaltados en negrita no pertenecen al original)

(3) **Come wander** through these Stratford pages, **get a taste** of Olde England - and **make your** plans now.⁹⁴⁷ (Resaltados en negrita no pertenecen al original)

(4) If **you're** eating out in London, **check out** our recommended London restaurants and foodie events.⁹⁴⁸ (Resaltados en negrita no pertenecen al original)

El tiempo verbal también es una peculiaridad de los textos turísticos, ya que predomina el presente de indicativo, tanto en alemán e inglés como en español. Como hemos indicado, las descripciones son el principal discurso de los textos turísticos y, por tanto, el empleo de este tiempo verbal se ve condicionado a la presencia de estas

⁹⁴⁵ <http://www.andalucia.org/cms/turismo-cultural/sectionvideos_view?set_language=es>

⁹⁴⁶ <http://www.bonn.de/tourismus_kultur_sport_freizeit/tourist_information/index.html?lang=de>

⁹⁴⁷ <<http://www.stratford-upon-avon.co.uk/>>

⁹⁴⁸ <<http://www.visitlondon.com/attractions/eat/>>

descripciones. Se trata de un presente atemporal, que eterniza las peculiaridades y bellezas del lugar turístico, sea este monumento, paisaje o alojamiento. En el siguiente fragmento, se pueden observar varios ejemplos:

(1) La Jara **es** una bella comarca natural situada en una región de sierras, montes, dehesas, arroyos y ríos en la parte suroccidental de la provincia de Toledo, que **se extiende** sobre la cuenca meridional del río Tajo. La Jara **posee** una vegetación (...).⁹⁴⁹ (Resaltados en negrita no pertenecen al original).

(2) Cambridge **is** a delight to visit in any season: relax in its many pubs, restaurants and cafes whilst exploring the independent shops around the historic market place. There **are** brand new shopping areas too, with all the high street favourites that you would expect.⁹⁵⁰ (Resaltados en negrita no pertenecen al original).

El pasado y el futuro del indicativo también se encuentran en todo tipo de narraciones, aunque en menor medida.

(1) El suave clima andaluz, el aire puro de sus montañas o el contacto con la refrescante brisa del mar, te **ayudarán** a que disfrutes más mientras te cuidas.⁹⁵¹ (Resaltado en negrita no pertenecen al original)

(2) Im Zuge des Dreißigjährigen Krieges **wurde** der "Alte Zoll" 1642 unter Kurfürst Ferdinand als sogenannte "Drei-König-Bastion" Teil der Stadtbefestigung.⁹⁵² (Resaltado en negrita no pertenecen al original)

(3) We **will** plan your ideal walking holiday along Hadrian's Wall, you simply choose when to go!⁹⁵³ (Resaltado en negrita no pertenecen al original)

Debe tenerse en cuenta también que muchos textos turísticos explotan la alternancia presente/pasado como recurso estilístico, con objeto de añadir belleza e importancia histórica a las descripciones que se estén realizando.

Otro elemento sintáctico muy empleado en este tipo de textos son las oraciones condicionales y de relativo, que atenúan la fuerza prescriptiva del mensaje, a la vez que

⁹⁴⁹ Folleto turístico *Provincia de Toledo*, 2009. Instituto de promoción turística de Castilla-La Mancha.

⁹⁵⁰ <<http://www.visitcambridge.org/VisitCambridge/Home.aspx>>

⁹⁵¹ <<http://www.andalucia.org/cms/turismo-cultural>>

⁹⁵² <<http://www.bonn-region.de/deutsch/sightseeing/historisches-bonn/alter-zoll.html>>

⁹⁵³ <<http://www.celtrail.com/hadrian/>>

ofrecen un abanico de posibilidades, de acuerdo con los gustos y aficiones del destinatario (Mira Rueda, 2008: 106):

(1) **Was** es als Erstes zum Ausgehen, den Shows oder Konzerten in Berlin zu sagen ist: Es gibt irre viel zu sehen und zu erleben.⁹⁵⁴ (Resaltados en negrita no pertenecen al original)

(2) **Si** lo que te gusta es disfrutar del paisaje y la naturaleza te sugerimos algunos lugares especiales por su belleza, por su entorno o por la diversión que ofrece al visitante.⁹⁵⁵ (Resaltados en negrita no pertenecen al original).

(3) **If** you're planning a holiday in Scotland, or a visit to the UK from abroad you really should put Loch Ness at the top of your list of UK holiday destinations.⁹⁵⁶ (Resaltados en negrita no pertenecen al original).

Por último, es importante indicar otros rasgos característicos de este lenguaje especializado, como son el empleo frecuente de recursos estilísticos, como las metáforas, las hipérboles y las metonimias, con objeto de aportar belleza al texto y mantener su función poética y emotiva, y el uso de formas locativas («aquí», «hacia», «hasta», «allí», «en dirección a», «a través de», «towards», «here», «in Richtung», «hier», «da», «nach»).

(1) El **inicio** de este recorrido parte de la comarcal 340 **en dirección** Colmenar-Puerto de los Alazores, tomando un desvío **a la izquierda** mediante pista terriza a la altura del Punto kilométrico 520,5 (1). Esta pista nos **conduce** a las inmediaciones de una parcela (...).⁹⁵⁷ (Resaltados en negrita no pertenecen al original).

(2) Die Tour **startet am** Bahnhof Zoologischer Garten, wo man in den Bus 100 einsteigt. Berlins berühmteste Buslinie fährt durch den Tiergarten **in Richtung** Unter den Linden. (...) Auf der John-Foster-Dulles-Allee **am** Haus der Kulturen der Welt, der ehemaligen Kongresshalle, verlässt man den Bus. **Hier, am Rand** des Tiergartens - übrigens auch unter Politikern als Joggingstrecke beliebt -, entstand nach

⁹⁵⁴ <<http://www.visitberlin.de/young/index.php?bereich=nightlife&seite=nightlife>>

⁹⁵⁵ <http://www.lariojaturismo.com/rutas/index.php?Id_contenido=18503>

⁹⁵⁶ <<http://www.visitlochness.com/>>

⁹⁵⁷ <<http://www.rutasdelsur.es/>>

dem Umzug von Bonn **nach** Berlin das neue Parlaments- und Regierungsviertel.⁹⁵⁸ (Resaltados en negrita no pertenecen al original).

(3) We **depart from** Edinburgh city centre at 8.30 am and **on our way north** will stop to take in the peak of Ben Vrackie just outside Pitlochry - a small picturesque town that has been attracting visitors since the days of Queen Victoria's excursions to the Highlands. Ben Vrackie is a beautiful little mountain, with a good path all the way and tremendous views from the top - but there is always the lure of the Moulin Inn nestled back at the foot of the hill! After our day's walking we **continue northwards to** our hotel and base for the week where dinner awaits.⁹⁵⁹

En general, y para terminar con este plano lingüístico, podemos resumir los aspectos más relevantes en las siguientes ideas: predomina la originalidad y la creatividad expresiva, la concisión y la brevedad, la connotación positiva y eufemismos, los enunciados directos y descriptivos, así como la referencialidad y la preponderancia de los datos sensoriales, que procuran sumir al lector en las luces, colores, olores y ruidos del lugar evocado. En definitiva, un lenguaje lleno de belleza, poesía, armonía y persuasión para atraer al turista al lugar que se promociona a través del texto turístico en cuestión.

3.3.4.1.2. Plano pragmático

Según Cabré Castellví (1993: 148-151), en este plano influyen dos aspectos fundamentales, a saber: la temática y la situación comunicativa.⁹⁶⁰

En el discurso turístico, la temática es un aspecto que lo influye principalmente en dos aspectos: por un lado, existe una gran variedad de segmentos dentro de este ámbito que se clasifican según la tipología del destino (turismo de sol y playa, turismo de aventura, turismo rural, turismo cultural, entre otros); y, por otro lado, las temáticas procedentes de disciplinas relacionadas con el turismo, tales como la historia, la arquitectura, el deporte, la gastronomía y el arte. Este hecho caracteriza al discurso turístico y lo diferencia de otros lenguajes de especialidad, a la vez que dificulta

⁹⁵⁸ <http://www.berlin-tourist-information.de/deutsch/sightseeing/d_si_berlinprogramme_hauptstadt.php>

⁹⁵⁹ <http://www.walkaboutscotland.com/heart_of_the_highlands.php>

⁹⁶⁰ Estos aspectos del plano pragmático los hemos comentado anteriormente en la descripción realizada de los lenguajes de especialidad (cf. 1.2.2.).

sobremanera la traducción de este tipo de textos. En este sentido, para realizar una traducción de calidad, no solo se deberá estar familiarizado con el discurso turístico y sus características propias, sino que además será necesario conocer y dominar estas otras áreas de conocimiento relacionadas con el turismo.

Con respecto a la situación comunicativa, Calvi (2006: 271) distingue principalmente dos niveles comunicativos en el sector turístico comercial, que se caracterizan básicamente por los intervinientes en la situación:

- La comunicación profesional especializada (nivel especializado), que tiene lugar entre profesionales del sector (en ámbitos de agencias de viajes, hostelerías, ferias y reuniones profesionales, etc.).
- La comunicación entre profesionales y usuarios (nivel semiespecializado y divulgativo), que puede desarrollarse tanto de forma directa, en forma de conversaciones en agencias de viaje, hoteles y otras estructuras turísticas, interacciones con guías de turismo, acompañantes de grupo, animadores socioculturales, etc., como indirecta o mediada, a través de medios electrónicos o escritos donde se presenta y describe el producto turístico: guías turísticas, folletos, prensa especializada, páginas web, anuncios publicitarios, etc.

Un rasgo característico de esta comunicación indirecta en la comunicación entre profesionales y usuarios es la ausencia de la identidad del emisor (Dann, 1996: 89), ya que en la mayoría de los textos turísticos (folleto, catálogo, guía de viaje, etc.) se desconoce el autor del mismo y, por supuesto, su traductor. Esto es debido al hecho de que el contenido es realmente lo que interesa y lo que prima sobre cualquier otro aspecto y, por ello, el autor permanece en el anonimato. Además, en este tipo de comunicación el objetivo principal es el cliente, que es el sujeto al que se le pretende persuadir, captar y seducir para que adquiera el producto o el servicio turístico.

Dentro de este plano pragmático, encontramos uno de los causantes de la relegación del lenguaje del turismo a la traducción general y, por tanto, a la lengua general: el predominio de comunicación especialista - no especialista, es decir, el nivel de semiespecialización y divulgación del lenguaje del turismo. Este tipo de comunicación presenta una dificultad de comprensión menor que otras comunicaciones

realizadas entre especialistas (especialista-especialista); sin embargo, el discurso turístico mantiene sus características de lenguaje de especialidad, incluso en este nivel de comunicación, debido a la presencia de terminología específica, sus funciones principales y sus rasgos lingüísticos propios. En este sentido, seguimos defendiendo aquí la clasificación del lenguaje del turismo como lenguaje de especialidad, a pesar del predominio de este tipo de comunicación no especializada en la comunicación turística.

En este apartado, debemos añadir aquí una tercera situación comunicativa: la comunicación no especialista - no especialista presente cada vez más en los entornos virtuales como *blogs* y foros de viajes (Calvi y Bonomi, 2008). Esta comunicación divulgativa comparte algunos rasgos de la comunicación semiespecializada, concretamente con el uso de términos especializados del ámbito turístico, verbos en imperativo y recomendaciones, abundantes descripciones, función informativa y persuasiva del texto, aunque presentan una visión mucho más personal y más “real” del lugar que describen y, además, se trata de una comunicación personal y directa. Este tipo de comunicación es relativamente reciente, aunque está muy extendida por la Red actualmente, por lo que es necesaria tenerla presente a la hora de analizar el discurso turístico. En nuestro caso, no hemos tenido en cuenta este tipo de comunicación dentro de nuestro estudio por no tratarse de una comunicación especializada.

Por último, cabe destacar el canal de comunicación en estas situaciones comunicativas. A pesar de que se siguen utilizando textos turísticos en formato papel, sobre todo guías turísticas y folletos durante el viaje, la Red se ha convertido en el principal canal de comunicación en el ámbito turístico (cf. Tabla 17), especialmente antes y después del viaje, debido a los cambios de consumo y preferencias de los turistas, como veíamos anteriormente. En este contexto, se están desarrollando en el ámbito turístico herramientas y programas informáticos que facilitan la gestión de los servicios y productos turísticos, como es el caso de las centrales de reserva, portales de servicios turísticos y de empresas turísticas, pago electrónico, contratación directa, cambios en la reserva, etc. Asimismo, las entidades y empresas turísticas tienen una mayor presencia en la Web y ofertan cada vez más sus servicios a través de los medios electrónicos, debido principalmente al acceso que tienen los usuarios a estos medios. De esta manera, el usuario puede adquirir, reservar o solicitar información desde su casa a

través de Internet. Esta facilidad permite una mayor y más rápida difusión y propaganda de los servicios y productos turísticos, a la vez que se capta a un mayor número de turistas potenciales.

3.3.4.1.3. Plano funcional

Desde el punto de vista funcional, todos los autores (Dann, 1996; Kelly, 1997; Calvi, 2000, 2006, entre otros) están de acuerdo en que la función referencial o informativa y la función conativa o apelativa predominan en el discurso turístico. No obstante, las otras cuatro funciones establecidas por Jakobson (función expresiva o emotiva, función fática, función metalingüística y función poética) (1984) también se observan en el discurso turístico en mayor o menor medida (Dann, 1996: 35-6). El texto turístico, por tanto, se puede considerar como texto multifuncional, ya que puede combinar diferentes tipos de funciones, según el tipo textual en cuestión.

Con respecto a la función informativa, podemos decir que se trata de la función principal de los lenguajes de especialidad, ya que en ellos predominan las descripciones, las definiciones, las clasificaciones, las referencias, etc. (Cabré Castellví, 1993: 155). En el discurso turístico, por tanto, también se trata de la función predominante, ya que todo texto turístico busca transmitir una información nueva al receptor, es decir, al turista.

La función apelativa está orientada al receptor del mensaje y hace referencia a la influencia que ejerce el emisor del mensaje en las actitudes y comportamiento del receptor. Para Nord (1997: 42), la función apelativa «is designed to induce them to respond in a particular way». Para poner en práctica esta función, se utilizan principalmente los vocativos y los verbos en imperativo, con el objetivo de persuadir, recomendar, permitir, ordenar y aconsejar.

Como hemos indicado con anterioridad, las otras funciones que estableció Jakobson (1984) también están presentes en el discurso turístico, aunque en menor medida.

La función expresiva o emotiva tiene como referente al emisor del mensaje y hace referencia a sus emociones y sentimientos. Esta función se expresa a través de expresiones de disculpa, de elogios, de aprobación, de condena, etc.

La función fática se puede considerar un complemento de la función apelativa y se utiliza para garantizar que no se interrumpa la comunicación o la complicidad entre autor y receptor.

La función metalingüística es la habilidad del lenguaje para hablar de sí mismo y de los códigos que emplea para transmitir el significado (Dann, 1996: 36).

Por último, la función poética está presente en el acto de comunicación cuando este está centrado en el mensaje mismo, en su disposición, en la forma en cómo se trasmite. Entre los recursos expresivos utilizados están la rima, la aliteración, la metáfora, los paralelismos, etc. Esta función no es característica de los lenguajes de especialidad en general; sin embargo, es muy común en el discurso turístico, al igual que ocurre con el discurso publicitario. Esto es debido al hecho de que el lenguaje del turismo está orientado a la captación, atracción y seducción de los turistas potenciales, con lo cual necesita utilizar un lenguaje lleno de belleza y de optimismo. En este sentido, esta función adquiere relevancia en el contexto turístico frente a las otras funciones secundarias.

Dentro de esta multifuncionalidad, se podrá observar una jerarquía de funciones, entre las cuales se deberá buscar la función prioritaria en cada caso (Hurtado Albir, 2001: 486). Con aplicación al ámbito turístico, cada tipo de texto turístico contará con una función prioritaria o predominante y otras funciones secundarias. Concretamente, para los textos promocionales, que hemos estado analizando hasta ahora, las funciones prioritarias son las funciones referencial y apelativa, siguiéndolas de cerca las funciones poética y emotiva.

Estas funciones se pueden observar en cualquier texto promocional, aunque según el tipo de texto tendrá mayor o menor presencia. En este sentido, en un anuncio turístico o un eslogan imperará la función apelativa y las funciones poética y emotiva adquirirán importancia, con objeto de atraer al turista y dar a conocer su producto o

servicio. Asimismo, la parte visual tendrá un papel importante también, como es el caso del anuncio para visitar el Lago Ness en Escocia *Seeing is believing*.⁹⁶¹



Figura 41. Anuncio de Destination Loch Ness Ltd (DLN) para promocionar el Lago Ness, en Escocia.

Por su parte, un folleto turístico contendrá más información y, aunque permanezca el mismo lenguaje directo, optimista y con predominio de adjetivos valorativos, la función referencial ocupará el lugar predominante, seguida de la función apelativa, poética y emotiva, en este orden. En otras palabras, se destacará el contenido sobre la forma, aunque esta se cuidará mucho así como las posibles imágenes que la acompañen.

Por último, en lo que respecta a las páginas web, se puede encontrar una mezcla de ambas, es decir, se encontrarán apartados donde predominará la belleza del lenguaje y la atracción al lector apoyada por imágenes y, además, se encontrará texto donde se describan los lugares, los servicios y los productos que se ofrezcan. No obstante, predominarán, como en todos los textos promocionales, la función referencial combinada con la apelativa, puesto que son los principales objetivos de los textos turísticos promocionales, a saber: informar y persuadir.

3.3.4.2. Características formales del discurso turístico

Para concluir con la caracterización del discurso turístico, trataremos a continuación el aspecto formal de los textos turísticos.

⁹⁶¹ <<http://www.visitlochness.com/the-environment/green-tourism.php>>

Como veíamos anteriormente, dentro del discurso turístico, podemos encontrar diferentes géneros textuales, al igual que en el resto de lenguajes de especialidad. En el abanico de posibilidades textuales que nos ofrece el discurso turístico, podemos distinguir además diferentes funciones y objetivos comunicativos, que deben tenerse muy presente a la hora de traducir cada uno de estos géneros.⁹⁶² Como tipo textual y como objeto del proceso traslativo, el texto turístico está ligado a ciertas convenciones formales que lo hacen inmediatamente reconocible (Bugnot, 2005: 55).

Con respecto a su formato, Fischer (2000) nos indica que los textos turísticos constan de los siguientes elementos:

- la portada del tríptico/folleto, etc. con el nombre del museo, de la ciudad que se va a visitar o un eslogan publicitario.
- el texto propiamente dicho, formado por descripciones, información histórica, etc.
- la información práctica (sobre horarios, transportes, clima, moneda, etc.)
- las ilustraciones, que generalmente van acompañando el texto e incluso llegan a ocupar más espacio que este en el caso de textos con fines primordialmente publicitarios.
- y otros elementos de interés como, por ejemplo, planos, plantas de edificios, publicidad de empresas, lugares para comer y dormir, etc.

Otro aspecto importante con referencia a los textos turísticos es la presencia abundante de elementos no lingüísticos, como fotografías, dibujos y símbolos. Estos elementos no lingüísticos forman parte de los textos turísticos como representación de la realidad extralingüística y son elegidos de forma intensional y precisa por el autor para mostrar y magnificar la realidad que representan. De esta forma, la elección de una fotografía del destino que se está publicitando, por ejemplo, pretende obtener una idealización del referente que designa dicha fotografía, con objeto de persuadir y atraer al posible turista.

Como indicamos con anterioridad en el plano lingüístico del discurso turístico (cf. 1.2.2.1.), el lenguaje del turismo tiende a utilizar siempre connotaciones positivas y

⁹⁶² En el siguiente apartado, trataremos la traducción de los textos turísticos más ampliamente.

a evitar las negativas. De esta forma, se tiende a valorar la juventud, la belleza, la felicidad, la armonía, etc. como elementos claves en los textos turísticos, especialmente en los promocionales, con lo que estos elementos están presentes con gran frecuencia en este tipo de textos. Esta presencia se realiza tanto desde el plano lingüístico como desde el plano icónico, con fotografías, dibujos y símbolos.

La tipografía, es decir, los signos paralingüísticos como el formato de letra, el tamaño, el estilo, el color, etc., juega también un papel muy importante en los textos turísticos y, por tanto, es un aspecto característico de esta tipología textual. Estos signos permiten resaltar ciertas palabras o sintagmas claves para la coherencia textual y el sentido del texto. Así, es necesario tenerlos muy en cuenta a la hora de realizar y de maquetar una traducción turística, puesto que el traductor debe conservar estos signos para mantener las funciones determinadas por el autor del texto original, ya sea una función informativa, expresiva o apelativa. Asimismo, es necesario seguir las normas ortotipográficas de las diferentes lenguas, con objeto de evitar errores ortotipográficos o de estilo en las lenguas meta.

Como veremos a continuación, este aspecto formal necesita ser respetado durante el proceso de traducción, puesto que a través de él se pretende atraer y persuadir al turista para que visite el destino en cuestión a través de la transmisión de sensaciones, belleza, alegría, aventura, etc.

Hasta el momento, los textos turísticos que han recibido un mayor número de estudios han sido los de tipo promocional, ya sean folletos, guías de viaje o anuncios publicitarios (Nobs, 2005; Londero, 2006; Alonso, 2005, entre otros). La mayoría de estos estudios se han centrado en la descripción del formato y las características lingüísticas, pragmáticas y funcionales, tanto del discurso turístico como de sus géneros textuales. Sin embargo, en nuestra investigación pretendemos avanzar un poco más y ofrecer un estudio exhaustivo sobre la terminología del discurso turístico, en particular del segmento de turismo de aventura, y estudiar sus relaciones conceptuales, a fin de proporcionar una conceptualización completa del segmento y elaborar una herramienta terminográfica de calidad para el traductor de textos turísticos y los profesionales de este ámbito de especialidad.

3.3.5. LA TRADUCCIÓN DE TEXTOS TURÍSTICOS EN ESPAÑA

Language barriers are, as everyone knows, an important obstacle to transcultural communication. Tourists are strongly aware of this difficulty, which has important effects on their choice of prospective destinations, their preparations for a trip, the scope and content of their interaction with the locals, and the quality of their experience (Cohen y Cooper, 1986: 534).

Para Nobs (2005: 63), la traducción de textos turísticos es «ante todo una labor de mediación intercultural que se lleva a cabo en unas condiciones de trabajo que con frecuencia distan de ser ideales». Esta labor de mediación intercultural, inherente a la práctica de la traducción, se observa claramente en la traducción turística. El turismo es una actividad que implica el contacto constante y directo entre culturas y todo lo que ello incluye (folklore, costumbres, gastronomía, normas, etc.). En este sentido, debemos destacar el papel de mediador intercultural que adopta el traductor de textos turísticos, lo que lo sitúa en una posición primordial en cualquier actividad turística plurilingüe.

Teniendo en cuenta este papel de mediador lingüístico y cultural, la importancia del traductor y de la traducción de textos turísticos, así como el protagonismo que ha adquirido la internacionalización y la promoción exterior en este sector⁹⁶³ y la afluencia de turistas internacionales en nuestro país,⁹⁶⁴ se debería suponer que esta traducción especializada recibe la atención correspondiente por parte de las organizaciones y administraciones públicas o privadas, empresas de turismo, etc. Sin embargo, no siempre es así, y a menudo tenemos que encontrarnos con textos turísticos incomprensibles y de calidad cuestionable.

⁹⁶³ En los últimos años, la internacionalización del sector turístico ha adquirido un protagonismo muy notable y se ha convertido en una meta para las empresas españolas. Asimismo, se incluye como una de las líneas prioritarias de actuación en materia de turismo en los diferentes planes turísticos que desarrollan las diversas administraciones públicas, tanto regionales como estatales y europeas. Como ejemplo, podemos citar de nuevo algunos de los proyectos actualmente vigentes o recientes más importantes, a saber: el Plan Director de Marketing 2009-2012, de la Junta de Andalucía y el Plan General del Turismo Sostenible de Andalucía 2008-2011, de la Junta de Andalucía; el Plan Turismo 2020, del Ministerio de Industria, Turismo y Comercio; la Carta Europea de Turismo Sostenible, entre otros.

⁹⁶⁴ Véase Gráfico 1.

Este hecho es debido, tal y como nos indica Fuentes Luque (2005: 32), al descuido de las organizaciones y entidades turísticas a la hora de realizar los encargos de traducción, ya que a menudo encargan sus traducciones a personas con conocimientos escasos en unos determinados idiomas y sin ninguna experiencia en el mundo de la traducción o sin ningún aval académico, lo que provoca que la traducción sea deficiente. Bretthauer (2000: 147) ilustra esta idea de forma muy clara cuando representa una situación de un cliente hablando con su “traductor”: «*Sie können doch Französisch, übersetzen Sie doch mal bitte*».⁹⁶⁵ Además, actualmente se sigue sin considerar la traducción especializada de textos turísticos como género específico y prueba de ello es el hecho de que la bibliografía dedicada a este tema es muy escasa. Asimismo, los teóricos de la traducción, en su gran mayoría, la han ignorado y la han tratado siempre como parte de la traducción general o en relación con la traducción publicitaria. Tanto es así que todavía se sigue tratando como parte de la asignatura de traducción general de las licenciaturas o de los nuevos grados de Traducción e Interpretación y no como un tipo de traducción especializada en sí (cf. Durán Muñoz, 2008c; Castillo Rodríguez, 2010: 173 y ss.).

Dada la situación, no sorprende comprobar que las traducciones de textos turísticos adolezcan de la calidad requerida y presenten a menudo errores de diferente naturaleza. Tanto es así que sus más recurrentes desaciertos suelen utilizarse como ejemplos para ilustrar procedimientos erróneos en clases de traducción. Este hecho perjudica desde dos puntos de vista principalmente: por un lado, a los traductores y a su prestigio, y por otro, a la imagen del destino turístico en cuestión.

Además de la frecuente baja calidad en las traducciones de textos turísticos, otro aspecto que denota la falta de atención de los clientes de estas traducciones turísticas son las versiones reducidas, parciales, con información diversa, o incluso sin traducir, que encontramos de textos originales, que hacen que el contenido del texto meta no corresponda con el del texto origen y que, por tanto, produzcan en los lectores meta una desinformación en algunos aspectos de la información turística que quizá fueran relevantes para su completa comprensión y acercamiento.

⁹⁶⁵ «Usted sabe francés, ¿no? Pues, traduzca esto» [traducción de la autora].

La idea de calidad en traducción ha suscitado muchos comentarios al respecto a lo largo de la historia y ha motivado la publicación de numerosas obras,⁹⁶⁶ además de diferentes normas de organismos internacionales que pretenden regular la calidad de los servicios de traducción.⁹⁶⁷ A menudo, se equipara «calidad» a «adecuación» de la traducción, es decir, el texto meta debe adecuarse a la cultura meta para que proporcione la equivalencia comunicativa y funcione correctamente y, así, se pueda considerar como una traducción de calidad. En este sentido, Nord (1997: 137) nos ofrece una definición de adecuación (*Adequacy*):

In terms of functional approaches to translation, «adequacy» is used to describe the appropriateness of a translated text for the communicative purpose defined in the translation brief. Adequacy is thus a dynamic concept related to the process of translational action.

En esta línea, «calidad» o «adecuación» se conseguirían manteniendo en el texto de llegada la finalidad pragmática y funcional acordada al inicio de la traducción, es decir, el traductor debe ajustar el contenido y la forma del texto al contexto, sus usuarios y las funciones comunicativas que debe desempeñar el texto traducido en la cultura meta. De esta manera, en cuanto a la adecuación pragmática, la información que se ofrece deberá ser útil y accesible para un receptor del texto meta, deberá ajustarse a las normas y convenciones de la cultura meta, deberá ofrecer un estilo y un registro normal en la cultura de llegada para ese tipo textual y, por último, deberá contar con la terminología y fraseología propias del campo de especialidad en cuestión. Por su parte, en lo que respecta a la adecuación funcional, el traductor deberá tener en cuenta las funciones características de un determinado tipo de texto turístico en la cultura meta, con objeto de

⁹⁶⁶ En las obras de Nobs (2003) y de Conde Ruano (2009) se puede consultar una revisión exhaustiva de las diferentes propuestas de evaluación y del concepto de «calidad» que se han realizado a lo largo de la historia reciente y actualmente, tanto en Traductología como en un ámbito más general.

⁹⁶⁷ Entre estas normas, caben destacar la Norma EN-15038, comentada anteriormente, así como otras como la norma UNI-10547 (*Definizioni dei servizi e delle attività delle imprese di traduzione ed interpretariato*), aprobada en 1996 por el Ente Nazionale Italiano di Unificazione (UNI, <http://www.uni.com/it>); la norma alemana DIN-2345 (*Übersetzungsaufträge*), aprobada en 1998 por un comité técnico del Deutsches Institut für Normung (DIN); y las normas ÖNORM D 1200 (*Dienstleistungen – Übersetzen und Dolmetschen - Übersetzungsleistungen – Anforderungen an die Dienstleistungen und an die Bereitstellung der Dienstleistung*) y ÖNORM D 1201 (*Dienstleistungen – Übersetzen und Dolmetschen - Übersetzungsleistungen – Übersetzungserträge*), aprobadas en 2000, y la norma ÖNORM 1203 (*Dienstleistungen – Übersetzen und Dolmetschen - Dolmetschleistungen - Dolmetschverträge*), aprobada en 2003, elaboradas por el Österreichisches Normungsinstitut (ÖNORM, <http://www.as-search.at/>).

conseguir la adecuación funcional en el texto de llegada, al igual que se ha cumplido en el texto origen (Durán Muñoz, 2008a: 379-381).⁹⁶⁸

Por ejemplo, la traducción de un folleto turístico español que funciona correctamente en la cultura española, que presenta las funciones características de los folletos (referencial, apelativa, emotiva y poética) y que recibe la aceptación de los lectores españoles debe conseguir el mismo efecto en la lengua de llegada. Así pues, la traducción de dicho folleto turístico español al alemán o al inglés, debe ser bien recibido por la cultura meta, ya sea alemana o británica, y debe mantener las mismas funciones que el original y, por tanto, funcionar correctamente en la cultura de llegada. En definitiva, los lectores de la cultura de llegada no deben notar diferencias con respecto a los textos originales en su lengua, exceptuando las referencias propias de la cultura origen (en este caso, la española).

Además de estos dos criterios de calidad como son la adecuación pragmática y funcional, es importante destacar también otros aspectos relevantes como son los siguientes: la legibilidad del texto meta, es decir, la claridad y la comprensibilidad, coherencia y cohesión que muestra el texto de llegada en su totalidad; la corrección gramatical en cuanto a las normas gramaticales, ortográficas y ortotipográficas de la lengua de llegada; el tratamiento adecuado del material verbal y no verbal, es decir, las imágenes y símbolos que aparecen en la traducción, el tipo de letra y el color, etc.; y, por último, el comportamiento profesional del traductor en cuanto a las competencias técnica y documental (cf. 2.2. y 2.3.), así como con relación a los aspectos acordados en el encargo (plazos, presupuesto, herramientas requeridas, cumplimiento de normas, uso de glosarios y memorias de traducción, etc.).

En este contexto, la consecución de la calidad o adecuación de una traducción dependerá sobre todo de la capacidad del traductor para no cometer errores. Así pues, si el traductor profesional cuenta con las competencias de traducción que especifica la norma EN-15038 (cf. 2.2.) y sigue los criterios mencionados con anterioridad, podrá ofrecer una traducción de calidad, exenta de errores y con la equivalencia comunicativa adecuada. De esta manera, cuanto más preparado se encuentre el traductor profesional

⁹⁶⁸ En algunos casos, las funciones del texto origen y del texto meta pueden variar, pero esto deberá estar recogido en las especificaciones del encargo de traducción.

para afrontar textos turísticos, menos posibilidades habrá de encontrar dificultades de traducción y más seguridad tendrá a la hora de enfrentarse con problemas de traducción.

Como nos indicaba Hurtado Albir (2001: 288), así como Nord (1991: 151) y otros autores, podemos encontrar problemas y dificultades en cualquier tipo de traducción, tanto general como especializada, que suelen ser el origen del error, entendiendo por problema de traducción «*an objective problem which every translator (irrespective of his level of competence and of technical conditions of his work) has to solve during a particular translation task*» (Nord: *ibíd.*) y por dificultad de traducción «*[a] subjective [problem] and has to do with the translator himself and his specific working conditions*» (*ibíd.*).⁹⁶⁹ De esta manera, un problema de traducción concreto puede resultar de gran dificultad al principiante y permanecer así, aunque este haya aprendido a enfrentarse a él; mientras que otros problemas de traducción (en este caso, dificultades de traducción) perderán su dificultad una vez que el traductor haya adquirido las habilidades y conocimientos necesarios para abordarlo.

Por lo que respecta a los problemas de traducción, podemos citar como ejemplos de las diversas clasificaciones que se han realizado hasta el momento por ser unas de las más extendidas y reconocidas. Por un lado, Nord (1991: 151) presenta una clasificación basada en cuatro puntos diferentes, a saber: 1. pragmáticos (lugares, épocas, relaciones entre las parejas comunicativas, etc.), 2. interculturales (pesos y medidas, convenciones formales, tipos de texto, fórmulas de saludos, etc.), 3. interlingüísticos (vocabulario, sintaxis y rasgos suprasegmentales de los dos idiomas) y 4. específicos del texto (metáforas, juegos de palabras, figuras retóricas, aliteración, rima, etc.). Y por otro lado, Hurtado Albir (*ibíd.*: 640) incluye uno más y varía ligeramente la clasificación: 1. lingüístico (léxico, sintáctico, textual), 2. extralingüístico (cultural, temático, enciclopédico), 3. de transferencia (dificultad en encontrar la equivalencia dinámica), 4. psicofisiológico (relativo a la creatividad o al pensamiento lógico) y 5. profesional/instrumental (derivado del encargo de traducción o de dificultades documentales). En el caso de la traducción turística, encontramos ciertos problemas de traducción que se repiten en la mayoría de los textos turísticos: referencias

⁹⁶⁹ La cursiva se ha utilizado para resaltar la principal diferencia entre un concepto y otro.

culturales con equivalencia 0 (Kade, 1981: 78), es decir, los llamados *culturemas* o *realia*; la traducción de nombres propios de personas, museos, instituciones, etc.; la traducción de topónimos; la diferencia de convenciones textuales entre lenguas; la presencia de juegos de palabras y metáforas difíciles de conseguir en la lengua meta, entre otros. Asimismo, siguiendo la clasificación anterior de Hurtado Albir también se incluirían aquí como problemas de traducción la falta de recursos terminológicos, tanto por la cantidad como por la calidad, para satisfacer las necesidades de los traductores a la hora de realizar sus traducciones de textos turísticos.⁹⁷⁰

Por su parte, a menudo se encuentran dificultades que provocan errores en la traducción final de los textos turísticos, las cuales se deben a menudo a la falta de instrucciones en el encargo, al escaso tiempo del que disponen los traductores, a la falta de fuentes de referencia fiables y actuales, así como a otras dificultades producidas por la falta de competencias de los mismos traductores (competencia técnica, cultural, lingüística, etc.). Entre los errores más comunes que se pueden encontrar en este tipo de traducciones fruto de estas dificultades, se encuentran principalmente errores terminológicos, errores ortotipográficos y gramaticales, ruptura de convenciones culturales y del estilo propio del género turístico en cuestión, función comunicativa del texto origen ignorada, ruptura o incumplimiento de las expectativas, falta de naturalidad en el texto meta, etc.,⁹⁷¹ todo lo cual da lugar a una traducción de mala calidad, con falta de adecuación y de aceptabilidad por parte de los receptores. Estos errores se deben en su mayoría a la falta de profesionalidad que existe en el sector de la traducción de textos turísticos, debido, por un lado, a la falta de formación específica en esta materia, y, por otro, al intrusismo existente en este ámbito. En un estudio realizado mediante encuesta informal sobre la traducción en empresas turísticas, sobre todo hoteleras, descrito en Kelly (2005: 158-159), se puede observar este intrusismo y falta de profesionalidad en el sector turístico a la hora de realizar traducciones de textos turísticos, ya que todas las empresas encuestadas para dicho estudio admitieron no haber contratado a traductores profesionales para realizarlas y, además, algunas de ellas indicaron que habían empleado a personas con cierto conocimiento de las lenguas de trabajo, a miembros del departamento de marketing o al personal de otros departamentos de la empresa. En este contexto, se

⁹⁷⁰ Sobre este tema se profundizará en el apartado 3.3.6. y siguientes de este capítulo.

⁹⁷¹ En el siguiente apartado, trataremos con mayor profundidad los diferentes tipos de error encontrados en traducciones reales de textos turísticos.

plasma claramente la falta de interés y de puesta en valor que se le otorga a la traducción en el sector turístico.

No obstante, para nuestra investigación, también hemos pretendido comprobar esta realidad y, para ello, realizamos un estudio durante los meses de septiembre y octubre de 2009 y nos pusimos en contacto mediante correo electrónico con diferentes empresas turísticas, en concreto entidades públicas (Turespaña, Patronato de Turismo de Lanzarote, Patronato de Turismo Cabildo de La Palma, Patronato Provincial de Turismo de la Costa del Sol, entre otras),⁹⁷² así como con traductores y empresas de traducción especializados en el sector turístico. Después de informarles sobre la investigación que se estaba llevando a cabo y los objetivos de esta, se les remitió un formulario con varias preguntas⁹⁷³ y se les pidió información sobre los encargos de traducción de sus textos promocionales. Según este estudio, todas las entidades públicas que han contestado a la encuesta contratan los servicios de agencias de traducción externas o traductores autónomos, a excepción del Patronato de Turismo Cabildo de La Palma que emplea a su personal administrativo (con licenciatura en Traducción e Interpretación) para la realización de estas traducciones. A pesar de que no son demasiadas las entidades que han contestado a estas preguntas, se puede observar que, al menos por parte de las instituciones y entes públicos, se está empezando a cuidar la calidad de la traducción de textos turísticos mediante la contratación de servicios de profesionales de la traducción.

Un aspecto que influye a menudo en la falta de calidad de las traducciones turísticas es la direccionalidad, ya que estas traducciones, en su mayoría, se realizan de forma inversa, es decir, realizadas de la lengua materna del traductor a su lengua extranjera. Según las entidades públicas y traductores profesionales contactados en el estudio anteriormente mencionado, todo el volumen de trabajo consiste en traducciones

⁹⁷² Se contactó con un mayor número de entidades públicas, pero solo se obtuvo respuesta por parte de las que se mencionan expresamente.

⁹⁷³ Las preguntas eran las siguientes: 1) ¿qué tipo de textos turísticos traduce normalmente (textos promocionales, académicos, etc.)?; 2) ¿traduce de forma inversa o directa (cuál es el porcentaje aproximado de cada una)?; 3) ¿cuáles son los recursos que utiliza en la actualidad cuando trabaja con este tipo de traducción?; 4) ¿qué opina de ellos (si son de calidad, si presentan carencias)?; 5) ¿a menudo no encuentra la información que busca?, ¿qué hace cuando esto sucede?; 6) ¿que sugeriría como mejora para este tipo de recursos? En este apartado, comentaremos las preguntas y respuestas relacionadas con la profesionalidad de los traductores de textos turísticos. Las respuestas sobre los recursos del sector turístico (3-6) se tratarán en los siguientes apartados de este capítulo.

del español a otras lenguas (especialmente inglés y alemán), que en gran proporción se lleva a cabo por traductores españoles que realizan una traducción inversa. Esto se debe principalmente al hecho de que las traducciones turísticas se llevan a cabo en el país de destino de los turistas, en nuestro caso en España. Esta direccionalidad provoca a menudo falta de precisión en los textos meta, así como numerosos errores de los citados anteriormente. En sí mismas, las traducciones inversas no deberían presentar ningún tipo de errores si se hicieran por traductores profesionales con una formación adecuada y con las competencias necesarias para realizarlas. De hecho, la mayoría de los expertos (Pedersen, 2000: 109; Machenzie y Vienne, 2000: 125) también defienden la traducción inversa, aunque sostienen que es necesaria una revisión exhaustiva por parte de hablantes nativos para mejorar la calidad de este tipo de traducciones. En lo referente a la traducción turística, pocas veces se encargan a traductores profesionales o se realizan este tipo de revisiones. No obstante, como veíamos anteriormente, por parte de las instituciones y entes públicos se está empezando a cuidar la calidad de la traducción de textos turísticos y, para ello, se encargan las traducciones a traductores profesionales, ya sea para traducción directa o inversa, así como servicios de revisión exhaustivos realizados por hablantes nativos de la lengua a la que se está traduciendo de forma inversa.

La falta de cooperación entre los autores del texto original y el traductor dificulta también la labor traductora. A menudo, el traductor recibe el texto de origen sin información básica de apoyo (fotografías, planos, indicaciones, etc.), lo que dificulta el proceso traductor. Aunque este hecho pueda parecer insignificante, no lo es. La traducción turística es una traducción subordinada (Fuentes Luque, 2005: 88), es decir, existe una relación directa entre el texto y los elementos no verbales, como son las imágenes, los símbolos, los dibujos, los planos, etc., especialmente con respecto a los textos turísticos en papel. En este sentido, los elementos verbales constituyen solo una parte del texto, ya que estos elementos verbales y los visuales forman un todo indisoluble. Este hecho hace imprescindible conocer dichos elementos no verbales que van a acompañar al texto para, así, poder realizar una equivalencia entre el código escrito y el código icónico.

Todas estas dificultades presentes en la traducción de textos turísticos y los errores derivados de ellas influyen negativamente en lo que se conoce en *marketing* turístico como la «marca» (*branding*) de un destino o de un país, en este caso España. La marca de un país hace referencia a la imagen que se vende de ese país y con la que se conoce a ese país en el extranjero. La marca, además de representar y diferenciar los productos, consigue crear, de forma natural, una imagen psicológica basada en percepciones que aporta un valor añadido a los bienes que acompaña, algo que influye directamente en las decisiones de compra. Por ejemplo, España ha realizado numerosas campañas promocionales en el exterior con un eslogan (*Freu Dich! Du bist in Spanien*, en alemán, y *Smile! You are in Spain*, en inglés)⁹⁷⁴ para atraer y seducir a los turistas potenciales extranjeros.

Debido a la importancia que adquiere la «marca» como la imagen de un país, es necesario que todos los textos turísticos, incluido los carteles promocionales, estén adecuadamente traducidos para captar a los futuros turistas, ya que, según Kelly (1997: 34), los textos turísticos constituyen la fuente a través de la cual los turistas reciben la primera impresión de un lugar turístico. En este sentido, este tipo de promoción turística ocupa el lugar más importante de todos, ya que se trata de atraer a los lectores del texto turístico. Así, la traducción debe ser atractiva, directa, adecuada, bien formulada y exenta de errores, puesto que cualquier error en dicha traducción puede provocar el desinterés de los futuros visitantes a nuestro país y podría dar lugar a una pérdida de reputación de la visión de un lugar por los turistas, tanto antes como durante su visita.

Una vez que el turista es real, es decir, visita nuestro país, también debe contar con textos turísticos correctamente traducidos. En muchas ocasiones, se ha producido falta de concreción, de comprensión o de explicación que ha dado lugar a equívocos y a una frustración por parte del turista.

Por todo esto, la traducción turística es esencial, especialmente en un país como España, donde su economía se sustenta en el sector turístico. Sin embargo, la traducción de textos turísticos no ocupa el lugar que le correspondería según su peso y queda

⁹⁷⁴ Este eslogan forma parte de la campaña de turismo internacional desde el año 2005. Para más información, consulte la página web de Turespaña: <<http://www.tourspain.es/es/TURESPANNA/Ejes+operativos/Imagen/Listado+imagenes+Publicidad+internacional.htm>>.

relegada a la traducción general y a la ejecución de personas no profesionales en la materia.⁹⁷⁵ Por ese motivo, una imagen que se trabaja tanto desde el punto de vista del texto origen en el *marketing* turístico se pierde en el proceso de traducción por la falta de interés y valor que se le presta a esta fase de la promoción turística. En este contexto, consideramos indispensable otorgar mayor importancia a la traducción turística, a fin de alcanzar una traducción de calidad que satisfaga las necesidades de los turistas extranjeros.

Para conseguir la adecuación en una traducción de textos turísticos, el traductor profesional debe contar con las competencias citadas en el capítulo anterior (cf. 2.2.), comunes a toda traducción de textos especializados, a saber: las competencias traductológica, lingüístico-textual, heurístico-documental, cultural y técnica.

La competencia traductológica hace referencia a la adaptación que realiza el traductor a la hora de traducir un texto turístico con una finalidad concreta. Por ejemplo, en el caso de un folleto turístico, el traductor debe mantener la función informativa y apelativa propias de este tipo de textos turísticos, así como cualquier instrucción proporcionada al comienzo de la traducción con referencia a la finalidad.

La competencia lingüístico-textual es una de las competencias más importantes en el ámbito de la traducción turística, puesto que hace referencia a la terminología específica, a las características morfosintácticas y ortotipográficas y a las convenciones textuales del discurso turístico.

Con referencia a la terminología específica, ya hemos visto que el discurso turístico presenta términos procedentes de otras disciplinas relacionadas, términos específicos del discurso turístico («tour operador», «agencia de viaje»), términos propios de una cultura o culturemas («chiringuito»), neologismos («rafting»), extranjerismos («kayak»), fraseología («turismo de sol y playa»), falsos amigos («mapa», en español – «*Mappe*», en alemán, cuando debería ser «*Landkarte*»), etc., lo que obliga al traductor profesional a realizar una labor documental muy amplia a la hora de traducir cualquier texto turístico y a consultar fuentes terminológicas en el lenguaje

⁹⁷⁵ Aunque, como veíamos, se están viendo indicios por parte de las entidades públicas de un cambio a este respecto y, por tanto, del reconocimiento de la traducción de textos turísticos como un elemento esencial en este sector.

turístico. Por este motivo, la competencia documental es una de las competencias prioritarias a la hora de traducir un texto turístico.

Asimismo, la competencia documental también puede ayudar al traductor a resolver problemas textuales relacionados con las convenciones y características textuales de los diferentes géneros turísticos, con objeto de intentar reproducirlos y recrear textos que sean fieles al patrón cultural y lingüístico original, así como a solventar dudas pragmáticas sobre las dos lenguas para su correcta adecuación a la cultura de llegada.

Actualmente, como veremos a continuación, las fuentes terminológicas existentes en materia de turismo en general son escasas y de dudable calidad, especialmente en la combinación alemán-español, por lo que el traductor no tiene a menudo posibilidad de consultar una obra terminológica fiable que le pueda solucionar cualquier duda que tenga con referencia a la terminología de un texto turístico. Por ello, nuestro principal objetivo en esta tesis doctoral consiste en elaborar una base de datos ontoterminográfica apropiada, fiable y de calidad junto con una herramienta de consulta y búsqueda terminológica acorde a sus necesidades para facilitar la labor traductora con respecto a la documentación y mejorar, así, la calidad de las traducciones turísticas, concretamente del sector del turismo de aventura.

En esta situación, en la que se encuentran escasas fuentes terminológicas disponibles de calidad, se vuelve a menudo imprescindible la compilación de corpus *ad hoc* (Corpas Pastor, 2001: 155) o *mini-corpora* (Maia, 2002: 219) para la labor traductora. Se trata de un corpus elaborado en un momento determinado para responder a dudas producidas durante la traducción de un texto concreto. Su objetivo es recopilar el mayor número de documentos concretos en el menor tiempo posible (Corpas Pastor, 2001: 164) y, así, observar la terminología y las características de textos similares al que es objeto de traducción, para adecuarlo correctamente a la lengua de llegada. En el caso de la traducción de un folleto turístico de español a alemán o al inglés, se deberán seleccionar documentos del mismo tipo textual en la lengua de llegada, con objeto de poder observar el formato y las características que presenta este género textual.

La compilación de este tipo de corpus se está aplicando actualmente en la didáctica de la traducción y también se ofrece como una estrategia muy recomendable en la traducción profesional. Se trata de un trabajo previo a la traducción que, por un lado, dilatará la consecución del texto de llegada al tener que dedicar tiempo a buscar y seleccionar los documentos concretos que formarán parte del corpus y, por otro lado, agilizará la labor traductora una vez compilado el corpus y garantizará una mayor adecuación a la cultura de llegada.

Para la selección de las páginas que formarán el corpus, es necesario disponer de conocimientos de búsqueda y de evaluación de recursos electrónicos así como de almacenamiento y de herramientas de análisis y extracción de términos. Estos conocimientos documentales y técnicos forman parte de la competencia documental y técnica de los traductores profesionales, que, además de conocer y de disponer de herramientas propiamente para la traducción, como, por ejemplo, las memorias de traducción, deben conocer también las técnicas de búsqueda y de consulta en Internet así como de evaluación de recursos electrónicos.⁹⁷⁶

Por último, es muy importante también en la traducción turística el aspecto cultural y, por ello, el traductor debe contar con la competencia cultural. Como hemos indicado anteriormente, la traducción turística se considera como una traducción de mediación cultural, ya que la cultura de la lengua origen está muy presente en los textos turísticos, tanto en la terminología utilizada como en el estilo, los contenidos, el formato, etc. En este sentido, el traductor debe conocer ambas culturas para poder trasponer correctamente la traducción del texto de origen a la cultura meta siguiendo sus patrones propios y manteniendo su finalidad comunicativa.

Las traducciones de las referencias culturales en los textos turísticos se consideran a menudo problemas de traducción, ya que en su gran mayoría se trata de realidades inexistentes en la cultura de llegada (culturemas) o de nombres propios o topónimos. Ante situaciones de este tipo, los traductores deben tomar decisiones para

⁹⁷⁶ Existe mucha bibliografía sobre este asunto. Para ampliar sobre este tema, véase Palomares Perraut (1998, 2000); Palomares Perraut y Gómez Camarero (2004); Gonzalo García (2004) y Seguri Domínguez (2006). En los siguientes apartados de este capítulo, se realizará una compilación de corpus según el protocolo de selección y evaluación presentado por Seguri Domínguez (2006), a fin de continuar con la fase de recopilación de información de la metodología de trabajo ontoterminográfico.

evitar incurrir en errores de traducción y deben adoptar las estrategias o técnicas traductorales que mejor se adapten al término en cuestión, pero siempre teniendo en cuenta las indicaciones iniciales del encargo de traducción o las pautas establecidas por el cliente y el espacio disponible para dicha traducción.⁹⁷⁷

En el caso de las traducciones de referencias culturales en los textos turísticos, una de las técnicas más utilizadas consiste en mantener el término original, siempre y cuando se estime que el lector va a comprenderlo. Por ejemplo, algunos términos gastronómicos populares como «paella» o «sangría» se mantienen a la hora de traducirlos a otra lengua. Junto a esta técnica, otra muy utilizada en la traducción de estos culturemas es la explicitación, que consiste en la ampliación de la información que aparece en el original para facilitar la comprensión del lector del texto meta. Por ejemplo, al traducir el término «chiringuito», se puede mantener el término y añadir una descripción breve de este para facilitar su comprensión, como podría ser «*Beach restaurant*».

El objeto de cumplir con todas estas competencias a la hora de traducir un texto turístico recae en el hecho de que el texto pueda cumplir su función comunicativa. En caso contrario, lo que se consigue es un texto correctamente traducido en cuanto a su gramática y sus funciones comunicativas (en su caso), pero deficiente en la traducción de rasgos propios del lenguaje turístico. El resultado será un texto de llegada al que no están acostumbrados sus lectores y, por tanto, la finalidad comunicativa puede quedar mermada o deteriorada, ya que el lector no identificará ese texto con el modelo que tiene interiorizado y se sentirá ajeno a él (Suau Jiménez, 2006: 2).

Con la variedad de aspectos que hay que tener en cuenta a la hora de realizar una traducción turística, podemos asegurar que no resulta nada fácil trasladar un texto turístico de una lengua a otra, especialmente por las abundantes referencias culturales del texto origen. Por ello, es imprescindible que el traductor profesional cuente con las competencias citadas anteriormente.

⁹⁷⁷ El tema del espacio en la traducción turística presenta a menudo dificultades, ya que la traducción va destinada a un espacio reducido y limitado (dígase un folleto turístico, anuncio, guía turística, etc.) y el traductor debe ceñirse a este espacio sin opción a ampliarlo.

3.3.5.1. Características de la traducción de textos turísticos promocionales: problemas, dificultades y errores frecuentes

The complexity of promotional tourist discourse is underestimated by clients and translators: it may appear to be deceptively easy to translate with its extensive use of general language; yet, it is a specialized discourse with specific linguistic/cultural features (Pierini, 2007: 99).

La traducción de textos promocionales ha sido, y sigue siendo, la traducción turística más investigada, tanto desde el punto de vista terminológico como de sus convenciones textuales y sus funciones. Esto es debido, por un lado, al hecho de que los textos turísticos promocionales tienen una mayor presencia en este ámbito, ya sea en forma de anuncios, carteles, folletos o páginas web, y también al hecho de que este tipo de textos presentan la finalidad principal del discurso turístico, la persuasión.

Como veíamos anteriormente, durante el proceso de traducción de este tipo de textos turísticos encontramos a menudo tanto problemas como dificultades traductológicas, siendo los primeros intrínsecos del discurso turístico y los segundos, dependientes del contexto de la traducción (encargo, competencias del traductor, etc.). Entre los principales problemas que encontramos en este tipo de traducción, destacamos los siguientes:

- La traducción turística como traducción subordinada. En el lenguaje turístico promocional, las imágenes e iconos de los textos representan una parte muy importante del mensaje que se desea transmitir y, por ello, a menudo se le denomina traducción subordinada. En este tipo de traducción, como hemos visto anteriormente, los elementos verbales y los no verbales se entrelazan y dan lugar a un todo indisoluble (Fuentes Luque, 2005: 88). Esta característica de la traducción de textos promocionales se puede convertir en una dificultad a la hora de traducir si el traductor no cuenta con el material de referencia, es decir, si el traductor no conoce las imágenes, los planos, los dibujos, etc. a los que se está refiriendo el texto. A esta dificultad le podemos añadir también la dificultad derivada de las restricciones de espacio por la publicación del texto en un medio

determinado y de los ajustes al formato del texto. Como ocurre con la traducción subordinada, el traductor debe limitarse a un espacio determinado y ajustarse a un formato en concreto. Esta dificultad se podría solucionar si existiera una colaboración real entre el cliente y el traductor y a este se le facilitara toda la información necesaria, especialmente con referencia a los materiales no verbales, el lugar de publicación, etc. Sin embargo, como ya hemos visto anteriormente, esta colaboración es prácticamente inexistente y muchos traductores trabajan a ciegas y sin instrucciones claras ni material de apoyo, todo lo cual perjudica la calidad de la traducción.

- Frases ambiguas o de doble sentido. A menudo, nos encontramos con frases ambiguas o con dobles sentidos que no funcionan en otras lenguas y que acarrearán pérdida de sentido a la hora de traducirlas. Por ejemplo, el eslogan de Andalucía «Andalucía te quiere». En este caso, estamos tratando con dos acepciones del verbo «querer»: 1. Desear o apetecer. y 2. Pretender, intentar o procurar.⁹⁷⁸ En alemán, la traducción «Andalusien liebt dich» se acoge a la primera acepción del verbo «querer», pero no hace referencia a la segunda, «brauchen» (necesitar) y en inglés ocurre lo mismo con el verbo «need» o «like». En este caso, la traducción al alemán o al inglés de una de las dos acepciones del verbo «querer» a las que hace referencia el eslogan andaluz, provocaría pérdida del sentido del original y no presentaría la misma fuerza en las dos lenguas. Lo mismo ocurre con un eslogan más actual como es el siguiente: «Andalucía. Sabe», donde el verbo «saber» hace referencia al sabor de la región y a su conocimiento o habilidad, es decir, de nuevo tiene más de una acepción que se perdería al traducirlo a otros idiomas: en el caso del alemán, tendríamos que elegir entre «schmecken» y «wissen»; y en el caso del inglés, entre «taste» y «know». Así pues, el traductor se encuentra a menudo en una disyuntiva entre dobles sentidos en la lengua original y falta de concordancia en la lengua de llegada y, por tanto, debe decidir qué estrategia seguir a la hora de traducir el eslogan y optar por una omisión de sentido, una compensación de

⁹⁷⁸ Estas acepciones han sido sacadas de la versión en línea del Diccionario de la Real Academia de la Lengua <<http://www.rae.es>>

sentido o un cambio del eslogan, de acuerdo con las instrucciones iniciales del encargo. Por lo general, esta dificultad se encuentra a menudo en los eslóganes de los anuncios, que se caracterizan por ser breves, directos, atractivos y sugerentes. Se trata de fórmulas concisas y directas que concentran la idea clave de la estrategia publicitaria con estética y sonoridad (Martí Marco, 2007: 127). Además del eslogan, en un texto turístico promocional nos encontramos con el cuerpo del texto, que tiende a ser más descriptivo y referencial, aunque el lenguaje atrayente y emotivo se refleja a lo largo de todo el texto.

- **Culturemas o *realias*.** Este hecho es muy característico de los textos turísticos promocionales, puesto que son los textos que acercan al turista a la cultura original del texto en todos sus aspectos (gastronómico, musical, de costumbres, fauna autóctona, etc.) y busca la diferenciación del destino que está ofreciendo con el de la cultura del turista. Anejo a esta presencia de culturemas, se encuentran las explicitaciones o descripciones en las traducciones de textos turísticos. El traductor se ve obligado a explicarle al turista qué es lo que se va a encontrar cuando pida un plato gastronómico típico de un lugar, el tipo de música típico de un destino o el posible significado del nombre de algún lugar o monumento. En este caso, cuando un traductor se encuentre con el término «chiringuito», deberá decidir qué estrategia seguir, según las instrucciones iniciales. En cualquier caso, y como hemos visto con anterioridad, se suele mantener el término en español y proporcionar una explicitación de qué significa ese vocablo en español, ya que es un término propio de la cultura origen.
- **Lenguaje positivo y poético constante.** Con el objetivo de enaltecer el producto y de llamar la atención del lector, el discurso turístico en este tipo de textos promocionales mantiene un lenguaje lleno de connotaciones y valoraciones positivas. Esta finalidad se consigue con el uso de sustantivos cargados de un alto grado de subjetividad («Natur», «Kultur», «Schatz», «Geschichte», «aventura», «obra de arte», «historia», «History», «nature»), términos específicos del ámbito turístico («Kulturerlebnis», «Rundgang», «actividades de tiempo libre», «lugar de destino», «Health and Beauty», «Bed and Breakfast») y con extranjerismos («souvenir», «Shopping», «en-suite»). Asimismo,

predominan los adjetivos con significado positivo («gesund», «lecker», «spannend», «alegre», «juvenil», «delicious», «beatiful») y con valor de intensificación («vielfältig», «einmalig», «magnífico», «sorprendente», «extraordinario», «wonderful», «breathtaking»), se sustituyen los verbos con significado amplio por verbos que despiertan la curiosidad del lector («kennen lernen», «betrachten», «descubrir», «conquistar», «discover», «experience») y se emplean verbos de movimiento («wandern», «fahren», «segeln», «dirigirse», «visitar», «come», «move»).

- *Ökosprache*. Esta «lengua ecológica» es muy común en la lengua alemana y cada vez más en la lengua inglesa y española. Como veíamos anteriormente (cf. 3.3.4.1.1.), esta «lengua ecológica» está muy presente en los hablantes alemanes y en su ideología, ya que la ecología, el reciclaje, la protección del medio ambiente, etc. forman parte de su cultura desde ya bastantes años. Sin embargo, en España esta idea de proteccionismo natural está estableciéndose en la actualidad, con lo que los términos de esta «lengua ecológica» no son tan comunes aún. En este sentido, el traductor de un texto turístico promocional de español a alemán tendrá que tener en cuenta este aspecto particular del lenguaje turístico alemán y amoldar su traducción a las expectativas de los lectores en este sentido. En caso contrario, no se producirá ningún error de traducción, pero sí que quedará alejada la traducción de los lectores del texto de llegada.
- Lenguaje confuso del texto origen. En algunas ocasiones, los traductores de textos turísticos se encuentran con textos mal redactados en español que pueden resultar confusos y con errores ortográficos, lo que provoca un doble trabajo al traductor: por un lado, debe comprender el texto origen confuso y, por otro, traducirlo correctamente sin contar con un original adecuado. En la mayoría de los casos, el traductor debe ponerse en contacto con el cliente y requerir información sobre la duda surgida y, además, podría proponer frases para reemplazar el texto confuso.

Por su parte, además de los problemas de traducción, también se dan las dificultades de traducción, aunque estas dependen más del contexto del encargo y de las competencias

del traductor encargado que del texto y del discurso turístico en sí. Las dificultades más comunes son las siguientes:

- Nombres propios de personas, museos, instituciones, etc. Los traductores pueden encontrar dificultades a la hora de traducir estos nombres, puesto que no siempre está claro si se deben traducir o no. En algunos casos, se deben traducir, como ocurre con los pintores famosos, sobre todo de la Edad Media: Albrecht Dürer = Alberto Durero; Hieronymus Bosch = El Bosco; o con los nombres de los cuadros más famosos: La Joconde / La Gioconda = die Mona Lisa; El jardín de las delicias = Der Garten der Lüste. Sin embargo, en lo que respecta a los nombres propios de personas, en español ha sido normal adaptar los nombres de escritores, músicos, reyes, etc. (Carlos Marx, Juan Sebastián Bach, Isabel II), en alemán o en inglés se suelen dejar sin traducir, aunque sí se adaptan los nombres de los Papas: Johannes Paul II = John Paul II = Juan Pablo II.
- Topónimos. Para algunos topónimos existen equivalencias en otras lenguas, como ocurre con New York (Nueva York) o Köln (Colonia), pero para otros no (Düsseldorf = Düsseldorf). Además, también se pueden encontrar algunas formas en desuso, que pueden dar lugar a equívocos (Neu-York, en alemán de principios de siglo).
- Falta de conocimientos sobre el lugar, sus características, su cultura, etc. Para poder traducir textos turísticos, que se consideran como herramientas de mediación entre la cultura de llegada y la cultura local, es imprescindible una amplia competencia cultural de ambas culturas y, por tanto, no sería suficiente contar con gran competencia lingüística.
- Falta de instrucciones. Para cualquier tipo de traducción es imprescindible contar con un encargo explicitando la función de la traducción y sus destinatarios potenciales. Si el encargo de un texto turístico promocional no va acompañado de dichas indicaciones y el traductor no tiene posibilidad de preguntar ni de acceder al material que acompaña al texto (imágenes, vídeos, carteles, etc.), la equivalencia comunicativa se verá gravemente afectada.

- Fuentes de referencia. El traductor deberá tener acceso a gran cantidad de material turístico relacionado con la traducción que está realizando (atlas, guías, enciclopedias, etc.), a fin de adquirir los conocimientos básicos que necesite para poder realizar la traducción. Además, deberá utilizar recursos terminológicos (diccionarios especializados, glosarios, enciclopedias, etc.), así como textos comparables, para consultar la terminología del texto, profundizar en el contenido semántico, adquirir la fraseología típica de este discurso, etc. Por lo que respecta al material de referencia y a los textos comparables, el traductor dispone de numerosas fuentes de consulta, especialmente en la Red; sin embargo, en cuanto a los recursos terminológicos, el traductor de textos turísticos se encuentra limitado a la consulta de un número muy reducido de fuentes de calidad.

Estos problemas y dificultades que encontramos en las traducciones de textos promocionales, así como la gran cantidad de traducciones realizadas por personas ajenas a la traducción profesional provocan frecuentes errores de diferente naturaleza. A continuación, comentaremos los más habituales encontrados entre los textos promocionales del segmento del turismo de aventura:⁹⁷⁹

- Errores ortográficos. Estos errores se encuentran tanto en los textos de origen como en los textos meta y, a menudo, se deben a la falta de revisión posterior:
 - «Eine der beliebtesten Modalitäten ist das *Besteigen von steilen Felsen» (*Bergsteigen*).
 - *«Risokoliehaber» (*Risikoliehaber*)
 - *«Hydrospeed» (*Hidrospeed*)
 - *«Canyonig» (*Canyoning*)
 - «Horse *Ridding» (*Horse Riding*)
 - *«Waterskiiing» (*Waterskiing*)
 - Falta de acentuación: *«Numero», *«mas», *«esqui», etc.

⁹⁷⁹ Los ejemplos de errores se han extraído de los textos que forman el corpus compilado para la elaboración del trabajo ontoterminográfico (cf. 3.4.).

- «Desde los montes más *exhuberantes del archipiélago» (*exuberantes*).
- <Rápel> se encuentra como «rápel», «*rappel*», *«rapel» y sus plurales «rápeles», *«rapels» y *«rapeles».
- Errores gramaticales:
 - Uso incorrecto de preposiciones: On Mondays, the visit *of [to] the Royal Palace of La Granja will be *substituted by* the visit *of [to] the Royal palace of Ríofrío.
- Errores sintácticos:
 - *«a part of which looking»
 - «Everyone has *practise before going into the light-hearted competition».
 - «After a *practise session, the object is to complete [...]».
 - «A simple cave system putting the SRT skills into *practise».
- Errores de terminología:
 - *«Fischerei», para traducir «pesca» (deportiva) (*Angeln*)
 - *«Freestyle-Klettern», para traducir «Escalada libre». en este caso, no tiene nada que ver con el estilo libre, sino con la limitada utilización de elementos artificiales.
 - *«Rohr», para traducir «tubo de buceo» (*Snorkel*)
 - *«Schwimmen», para traducir «bucear» (*Tauchen*)
 - *«Descargar el folletín» (*Folleto*)
 - *«Trekking», para traducir «senderismo» (*Hiking*)
 - *«Regulator», para traducir «regulador» (*Atmenregler*)
 - «Abseilen», para traducir «descenso de barrancos» (*Canoying*)
- Omisión de contenido y falta de explicitación:
 - «Empfohlen wird das vorherige Einholen von Informationen über die verschiedenen Klettergebiete bei der Federación Canaria de

Montañismo y Escalada oder die Begleitung eines örtlichen Bergsteigers» (TM) → «Se recomienda ampliar o actualizar información sobre zonas de escalada, normativas, equipo, etc. a través de la Federación Canaria de Montañismo y Escalada o acompañarse de escaladores locales (TO)».

- «Die Sportfans können...» (TM) → «Los aficionados a los deportes náuticos pueden...» (TO)
- «Active Tourismus» (TM) → «Turismo activo y rutas de senderismo» (TO).

- Falta de coherencia y repetición:

- Diferentes términos para la misma idea en una descripción de una ruta: «Gehzeit» / «Fahrzeit» / «Fahrdauer».
- Incoherencia a la hora de traducir nombres de lugares: unas veces se incluye la traducción entre paréntesis («Cueva bonita» se traduce como «Cueva Bonita (Schöne Höhle)», y otras se dejan en español.
- Incoherencia con el término utilizado en el epígrafe, «Hidrospeed» y en el cuerpo del texto, «hydrospeed».
- En versiones traducidas de un mismo texto origen, se encuentran diferentes opciones de traducción de nombres propios. Por ejemplo, en el texto sobre barranquismo de la página web de Turismo Andaluz encontramos que la traducción en alemán se mantiene el nombre propio: «Garganta Verde», «Río Verde»; mientras que en inglés se opta por su traducción: «Verde Gorge», «Verde River».
- Incoherencia a la hora de traducir nombres propios en un mismo texto. Por ejemplo, en la versión alemana del mismo texto anteriormente mencionado encontramos «Río Verde» y «Fluss Guadalmina» cuando aparecen en la versión original referencias a estos ríos. Otros ejemplos pueden ser: «Die Schlucht von Marcos y Cordero im Herzen des Weltbiosphärenreservats Los Tilos, das Barranco de Las Grajas [...]», «The Barranco de la Madera (La Madera ravine), next to the las Nieves church in Santa Cruz de La Palma» o «Das Barranco de la Madera bei der Kirche von Las Nieves in Santa Cruz de La Palma».

- Repetición de los términos «Straßen» y «roads»: «caminos, senderos, cañadas y calzadas» (TO) = «Straßen, Wanderwege, Bäche und Straßen» (TM) // «roads, trails, creeks and roads» (TM).
- Repetición de la modalidad de escalda: «rock climbing, as much in rock as in ice».
- Incoherencia a la hora de utilizar el concepto <Rápel>, en su forma singular y plural, en un mismo texto: «rápel», *«rapel», «rápeles», *«rapels» y *«rapeles».

Como corolario, podemos resumir exponiendo que el lenguaje turístico promocional es un lenguaje eufemístico y positivo que pretende despertar el interés del turista potencial para consumir el producto o el servicio que ofrece y que magnifica, por tanto, el destino al que promociona. Por este motivo, la traducción turística debe quedar exenta de errores y debe mantener la misma equivalencia comunicativa que su texto origen, así como las características propias de los textos turísticos promocionales. En este contexto, el traductor de este tipo textual se convierte en un «poeta» turístico con el difícil cometido de transmitir todo el contenido referencial y formal del texto original sin cometer errores como los que hemos visto anteriormente. Actualmente, se observa cierta tendencia a la mejora de las traducciones turísticas y al interés por conseguir un texto meta de buena calidad por parte de las entidades públicas del ámbito del turismo, como hemos observado en nuestro estudio, así como por parte de las agencias y traductores autónomos especializados en este tipo de traducción.⁹⁸⁰ Sin embargo, aún encontramos textos turísticos que presentan numerosos errores de diferente naturaleza debido a menudo a la falta de formación específica de los traductores que realizan la traducción y a la falta de atención e interés por parte de los clientes, lo que perjudica la imagen de nuestro país e incluso provoca interpretaciones erróneas y desinformación por parte de los turistas que reciben los textos turísticos. En nuestra opinión, sería necesario un aumento de interés por parte de todas las entidades, públicas y privadas, para alcanzar un nivel elevado de calidad en los textos turísticos y formación específica

⁹⁸⁰ Hemos tenido acceso a guías de estilo y recomendaciones que elaboran algunas agencias de traducción especializadas en el sector turístico, para guiar a los traductores a la hora de traducir ciertos elementos típicos de la traducción de textos turísticos, como son los nombres de monumentos, de personas, de lugares, de ciudades, de calles; las direcciones; las abreviaturas, etc.

de traductores para la traducción de textos turísticos para, así, evitar una posible imagen negativa de nuestro país.

3.3.6. RECURSOS TERMINOLÓGICOS PARA LAS LENGUAS INGLÉS, ALEMÁN Y ESPAÑOL EN EL SEGMENTO DEL TURISMO DE AVENTURA

Una vez realizado el análisis exhaustivo del dominio del turismo en general por cuanto se refiere a sus características discursivas, sus géneros textuales, las dificultades y aspectos relevantes de la traducción de textos turísticos y los diferentes segmentos turísticos que lo forman, especialmente el segmento del turismo de aventura, hemos alcanzado una visión más completa sobre este sector, sobre las carencias que presenta en la actualidad y sobre las posibles necesidades que se pueden cubrir con el recurso ontoterminográfico que estamos elaborando en esta investigación.

Como hemos visto, una de las características de este sector es la baja calidad en las traducciones debido a errores de diferente naturaleza: gramaticales, semánticos, ortográficos, etc. Afortunadamente, las instituciones y empresas, públicas y privadas, están prestando cada vez más atención a la traducción de material promocional, aunque todavía seguimos encontrando material con traducciones deficientes. Con objeto de evitar los posibles errores en este tipo de material, consideramos necesario realizar una distinción entre las dos causas fundamentales de esta problemática: por una parte, los errores gramaticales, ortográficos, ortotipográficos, etc. que muestran las traducciones de este sector se deben principalmente a la formación del traductor que las realiza o a la falta de tiempo y de revisión del trabajo; y por otro lado, los errores semánticos, de variantes lingüísticas, etc. son debidos a la falta de recursos terminográficos adecuados y fiables para este segmento en concreto, aunque también a la falta de formación del traductor en el sector en el que trabaja.

Esta segunda causa queda patente en las respuestas obtenidas por los traductores de textos turísticos contactados en nuestro estudio indicado con anterioridad (cf. 3.3.5.1.) cuando se les realizaron las siguientes preguntas: 1) ¿cuáles son los recursos que utiliza en la actualidad cuando trabaja con este tipo de traducción?; 2) ¿qué opina de ellos (si son de calidad, si presentan carencias)?; 3) ¿a menudo no encuentra la

información que busca?, ¿qué hace cuando esto sucede?; 4) ¿qué sugeriría como mejora para este tipo de recursos?⁹⁸¹ Todos los encuestados han confirmado que existen carencias de recursos para la traducción de textos turísticos, especialmente para la combinación de alemán-español, y que, por tanto, a menudo no encuentran en los diccionarios la información que necesitan. Por ello, deben recurrir principalmente a diccionarios generales y búsquedas en Internet mediante buscadores generales, como *Google*, para resolver sus dudas, así como al empleo de textos comparables o paralelos como apoyo a la traducción, a fin de buscar las opciones más satisfactorias para los equivalentes de traducción y registro del dominio. Con respecto a la última pregunta realizada, todos están de acuerdo en la necesidad de recursos más completos y específicos del sector turístico y, además, una de las encuestadas indica la necesidad de elaboración de un libro de estilo para regular en cierta manera la traducción en este sector y que sirva como guía tanto a traductores como a clientes.

Volviendo a los problemas planteados anteriormente, para el primer tipo de problema consistente en los errores gramaticales, ortográficos, etc., no podemos hacer nada más que pedir a los clientes que solicitan traducciones de material promocional turístico que valoren este tipo de traducción como una traducción de textos especializados y que, por tanto, contraten a profesionales capacitados para desempeñar la labor de forma satisfactoria y les proporcionen todo el material necesario para llevar a cabo su trabajo.

Por su parte, el segundo problema, que radica principalmente en los recursos terminológicos que existen en el sector turístico, sí podría ser solventado por los terminógrafos, ya que se debe a dos causas principalmente y directamente relacionadas con los recursos terminológicos disponibles, a saber: por un lado, la falta de obras especializadas que necesitan consultar en sus combinaciones de idiomas⁹⁸² y, por otro,

⁹⁸¹ Este estudio constaba de seis preguntas, que se pueden consultar en la nota al pie n.º 74, de las cuales dos de ellas se han tratado en el apartado 3.3.5.1.

⁹⁸² Como veremos más adelante, los recursos monolingües son los que tienen mayor presencia en este sector, así como los recursos con la combinación español-inglés, dejando a un lado otras combinaciones como el alemán-español (cf. Tabla 20).

la respuesta inadecuada de las existentes para todas las necesidades que se presentan.⁹⁸³ Todo esto provoca grandes perjuicios al traductor que está realizando una traducción en este sector al no disponer siempre de información terminológica fiable, adecuada y completa. En nuestra opinión, para subsanar este problema en el sector turístico, sería necesario tener en cuenta tanto las necesidades de los traductores que trabajan en este sector como la multidisciplinariedad e interdisciplinariedad que presenta y, por consiguiente, elaborar recursos específicos para los diferentes segmentos turísticos y campos en los que se compone este campo de especialidad según las necesidades de los traductores. En otras palabras, consideramos que no es suficiente la elaboración de recursos generales (Tabla 18) que pretenden abarcar todo el sector, ya que estos no pueden recoger la terminología tan variada que presenta y se basan principalmente en los términos relacionados con la gestión, el transporte, el alojamiento turístico, el aseguramiento turístico y cierta terminología de los segmentos turísticos más tradicionales (segmento de sol y playa, segmento cultural, etc.), sino que es necesario crear recursos que traten específicamente y de forma independiente el segmento de turismo de aventura, el segmento de sol y playa, el segmento de salud y belleza, el ecoturismo, etc., a fin de recopilar toda la terminología propia de estos segmentos y ofrecer a los usuarios, en nuestro caso, a los traductores profesionales, recursos completos y adecuados para las diferentes traducciones a las que se deben enfrentar durante su trabajo. De esta manera, se facilitaría y agilizaría su trabajo y se evitarían las múltiples consultas que realizan a diferentes fuentes (diccionarios, glosarios, sitios web, etc.) antes de poder tomar decisiones traductológicas.

⁹⁸³ Recordemos que el lenguaje del sector turístico se caracteriza por su multidisciplinariedad, ya que está en constante contacto con otras disciplinas cercanas (arquitectura, arte, geografía, economía, transporte, etc.), y por su interdisciplinariedad, ya que está formado por diferentes segmentos turísticos que presentan su terminología propia. En la Tabla 18 se pueden observar ejemplos de diccionarios del sector turístico, que incluyen una terminología general de este ámbito de especialidad pero que carecen a menudo de la terminología propia de cada segmento, limitándose sobre todo a la terminología relacionada con la gestión, el transporte, el alojamiento turístico, entre otras.

| RECURSOS DEL SECTOR TURÍSTICO EN GENERAL (inglés, español y alemán) | | |
|---|--|--|
| 1 | Diccionario turístico | español |
| 2 | Diccionario de terminología turística y cultural | español |
| 3 | Léxico de términos turísticos. Diccionario turístico profesional | español |
| 4 | Turismo D+ | español |
| 5 | Glosario de Turismo y Hostelería | español |
| 6 | Glosario de Iconos | español |
| 7 | Tourism Dictionary online | inglés |
| 8 | WebTerm: Tourism, Travel, Management and EDP Glossary | inglés |
| 9 | A dictionary of Travel and Tourism Terminology | inglés |
| 10 | Dictionary of Travel, Tourism and Hospitality | inglés |
| 11 | Dictionary of Leisure, Travel and Tourism. | inglés |
| 12 | Tourism Glossary | inglés |
| 13 | Lexikon der Tourismuswirtschaft | alemán |
| 14 | Langenscheidt Praxiswörterbuch Tourismus Spanisch | alemán, español |
| 15 | Bilingual Spanish-English and English-Spanish lexicons (INCYTA) | español, inglés |
| 16 | Diccionario turístico inglés-español | español, inglés |
| 17 | Diccionario de términos de turismo y de ocio: inglés-español, Spanish-English | español, inglés |
| 18 | Travel Industry Dictionary | inglés, alemán |
| 19 | Online Wörterbuch Tourismus | inglés, alemán |
| 20 | Online Dictionary German-English | inglés, alemán |
| 21 | PONS Fachwörterbuch, Touristik, Hotellerie, Gastronomie, Englisch-Deutsch / Deutsch-Englisch | inglés, alemán |
| 22 | TourisTerm | inglés, francés, español, árabe, ruso |
| 23 | Tourist dictionary | inglés, francés, alemán, holandés |
| 24 | Fachwörterbuch für Hotellerie & Gastronomie | alemán, inglés, francés, italiano |
| 25 | LORETO Thesaurus | inglés, francés, alemán, italiano, polaco, portugués, español, turco |
| 26 | Lexmatch Wörterbuch Tourismus | español, inglés, francés, alemán, italiano, portugués |
| 27 | Standardisiertes Wörterbuch Tourismus: Deutsch, Englisch, Französisch | alemán, inglés, francés |

Tabla 18. Algunos recursos terminológicos del sector turístico en general para las lenguas inglés, alemán y español.

En este contexto, se enmarca uno de nuestros objetivos de esta investigación, a saber: la elaboración de un recurso útil y fiable para apoyar las necesidades de los traductores a la hora de realizar traducciones de material promocional para el segmento del turismo de aventura para las lenguas español, inglés y alemán, ya que, según una de nuestras hipótesis planteadas en esta investigación, los recursos para el segmento del turismo de aventura son escasos, incompletos y de baja calidad.

Como parte de la fase inicial de la metodología propuesta en nuestra investigación (cf. 2.6.1.) y como complemento al análisis del dominio objeto de estudio que hemos venido realizando en este capítulo, probaremos esta hipótesis anterior mediante el estudio de los diferentes recursos turísticos disponibles en la actualidad para el segmento del turismo de aventura para las lenguas inglés, español y alemán. Mediante dicho estudio tendremos la oportunidad de conocer los recursos existentes, las necesidades actuales en este segmento, así como las deficiencias que presentan para los traductores profesionales, a fin de elaborar un recurso adecuado para nuestro objetivo.

Para su consecución, este estudio se ha dividido en tres fases: la primera, que consiste en la búsqueda y selección de las fuentes terminológicas disponibles para las lenguas indicadas (inglés, alemán y español) en papel y en formato electrónico; la segunda fase, que consiste en la clasificación de las fuentes seleccionadas según la clasificación de fuentes terminológicas establecida con anterioridad (cf. 2.4.1.); y, por último, el análisis cuantitativo de los recursos seleccionados y clasificados según los criterios de la plantilla de evaluación elaborada también anteriormente (cf. 2.4.2.1.). De esta manera, hemos pretendido realizar una valoración objetiva y sistemática de las fuentes terminológicas para el segmento del turismo de aventura seleccionadas.

3.3.6.1. Búsqueda y selección de los recursos terminológicos turísticos para el segmento del turismo de aventura

Una vez presentados algunos de los recursos terminológicos del sector turístico disponibles actualmente y después de haber indicado su ineficacia para llevar a cabo la traducción de textos turísticos promocionales del segmento del turismo de aventura, limitaremos nuestra búsqueda a los recursos que versan específicamente sobre este

segmento turístico, así como el procedimiento seguido para detectar y seleccionar los recursos seleccionados.

Antes de empezar con la búsqueda de recursos, hemos determinado unos criterios mínimos que deben cumplir para su selección, a saber: a) deben ser recursos publicados (en caso del formato en papel o en CD-ROM) o actualizados (en caso de recursos en línea) después del año 2000, a fin de recopilar posibles recursos adecuados para su empleo en la traducción, y b) deben tratar el segmento del turismo de aventura o las actividades que lo caracterizan.

Para realizar la detección de estos recursos terminológicos, hemos utilizado los tres tipos de búsquedas en Internet que diferencia Autermühl (2001: 52 y ss.) y que ya comentamos en el capítulo anterior (cf. 2.6.1.2.2.), a saber: *búsqueda institucional*, que hace referencia a la búsqueda que se realiza en sitios web de instituciones u organizaciones internacionales, como la Organización Mundial del Turismo (OMT)⁹⁸⁴ o la Biblioteca de la Universidad de Málaga;⁹⁸⁵ *búsqueda temática*, normalmente mediante directorios, como *Onelook.com*⁹⁸⁶ o *Yourdictionary.com*⁹⁸⁷ y, por último, *búsqueda por palabra clave*, a través de buscadores, como *Google*.⁹⁸⁸

A pesar de que nuestra plantilla de evaluación está diseñada para evaluar recursos tanto en formato electrónico como en papel, hemos realizado la búsqueda a través de Internet, ya que en la Red se pueden encontrar recursos en línea, es decir, los recursos disponibles a través de Internet, así como datos de referencia de recursos electrónicos y recursos en papel. Estos últimos se encuentran disponibles en catálogos de bibliotecas, en portales de librerías, en catálogos de librerías, entre otros muchos sitios.

Una vez establecidos los tres tipos de búsquedas que se pueden realizar en la Red, iniciamos nuestra andadura con la *búsqueda institucional*, a través de instituciones, organizaciones, federaciones, asociaciones y escuelas. Al contrario de lo que

⁹⁸⁴ <http://www.unwto.org/index_s.php>

⁹⁸⁵ <<http://www.uma.es/servicios/biblioteca/>>

⁹⁸⁶ <<http://www.onelook.com>>

⁹⁸⁷ <<http://www.yourdictionary.com/>>

⁹⁸⁸ <<http://www.google.com>>

pensábamos antes de comenzar, este tipo de búsqueda no fue muy fructífera, ya que la mayoría de las organizaciones, instituciones y asociaciones turísticas no disponen de recursos terminológicos y, en el caso de incluir algunos, se tratan de recursos sobre el turismo en general y, además, normalmente trabajan con las lenguas inglés y español, pero no con la combinación alemán-español. Por ejemplo, la OMT dispone de una base de datos de terminología turística multilingüe, *Turisterm*,⁹⁸⁹ que trata sobre el dominio del turismo en general y que, además, actualmente no incluye el idioma alemán; *Turespaña*⁹⁹⁰ o el Ministerio de Industria, Turismo y Comercio⁹⁹¹ no disponen de ningún tipo de recurso terminológico turístico; la base de datos de la Unión Europea, *IATE*,⁹⁹² no incluye el turismo como un campo de especialidad; la mayoría de las asociaciones y federaciones de deportes relacionados con el turismo de aventura (senderismo, piragüismo, deportes acuáticos, etc.) no incluyen recursos terminológicos o no incluyen la combinación de lenguas alemán-español; algunas universidades y facultades de Turismo⁹⁹³ incluyen una recopilación de recursos turísticos, pero no incluyen recursos terminológicos, y así sucesivamente.

La búsqueda institucional se ha combinado a menudo con la *búsqueda por palabras clave* para ampliar el área de búsqueda, ya que nuestro objetivo ha consistido en consultar tantas fuentes institucionales como fuera posible para encontrar sus recursos terminológicos y realizar, así, su posterior evaluación. En este sentido, se ha utilizado el buscador *Google*,⁹⁹⁴ y las técnicas booleanas propias de este buscador,⁹⁹⁵ para crear ecuaciones de búsqueda que nos ayudaran a detectar un mayor número de

⁹⁸⁹ <<http://www.unwto.org/trad/index.php>>

⁹⁹⁰ <www.tourspain.es/>

⁹⁹¹ <<http://www.mityc.es/es-ES/index.htm>>

⁹⁹² <<http://iate.europa.eu/iatediff/SearchByQueryLoad.do?method=load>>

⁹⁹³ La página web de una asignatura de la Universidad de Sevilla incluye enlaces a recursos turísticos (<<http://www.aloj.us.es/gist/utilidades.htm>>), la biblioteca de la universidad de Cádiz dispone de una recopilación de recursos de turismo (<http://biblioteca.uca.es/sbuca/bibcjer/recursos_tur.asp?capbd=4>), el Centro de Documentación europeo en la Universidad de Barcelona dispone de una página web con recursos sobre turismo (<<http://selene.uab.es/ce-documentacio-europea/ue/tem-turisme.htm>>).

⁹⁹⁴ Según la opinión de numerosos analistas, *Google* es el mejor buscador con relación a la calidad en el resultado de las búsquedas, tanto en precisión como en recuperación (Lorán, 2005: 2; Radev et ál., 2005: 580; Murphy et ál., 2007: 301).

⁹⁹⁵ Para profundizar en las técnicas aplicables para realizar búsquedas en este buscador, véase URL: <<http://www.google.es/support/bin/static.py?page=searchguides.html&ctx=basics>>.

instituciones turísticas. Entre las ecuaciones empleadas, citaremos las siguientes a modo de ejemplo: *institución AND turismo*, *institución OR turismo de aventura*, *institución AND turismo de aventura*, *institución AND turismo activo*⁹⁹⁶ o “*institución turística*”, “*organización turística*”, *organización + turismo*, con objeto de recuperar un mayor número de instituciones turísticas. Con esta combinación de búsquedas, hemos recuperado un mayor número de instituciones, asociaciones y organismos turísticos, por lo que hemos comprobado si tenían disponibilidad de recursos terminológicos propios. Sin embargo, como ya hemos indicado anteriormente, estas páginas institucionales carecen en su mayoría de recursos de términos especializados, especialmente con relación al par de lenguas español-alemán y al segmento del turismo de aventura.

En segundo lugar, hemos realizado la *búsqueda temática*. Al igual que la anterior, esta búsqueda puede ir acompañada de una búsqueda por palabras clave, para ampliar el número de resultados obtenidos. Con la combinación de ambos tipos de búsquedas, se aumentó el número de resultados y, por tanto, se obtuvo un mayor número de directorios turísticos. Este es el caso, por ejemplo, del directorio de senderismo y montañismo de *Labota*⁹⁹⁷ o el directorio turístico *Masactividades.com*, con actividades y ofertas turísticas de todo el mundo.⁹⁹⁸ Sin embargo, y como ocurría también en el tipo de búsqueda anterior, estos directorios especializados en turismo, no contienen recursos terminológicos y, por tanto, no son útiles para nuestro cometido. Debido a este hecho, dejamos de buscar directorios especializados en este campo de especialidad y comenzamos a realizar búsquedas de directorios de fuentes y recursos lingüísticos.

En la Red, existe gran cantidad de directorios que recogen recursos lingüísticos de todo tipo, tanto generales como específicos, y, entre estos últimos, se localizan fácilmente las fuentes especializadas en turismo y en deportes relacionados con el turismo de aventura. Así pues, encontramos numerosos diccionarios, glosarios, léxicos, bases de datos, etc. incluidos en este tipo de directorios. Por ejemplo,

⁹⁹⁶ Recordemos que en muchas ocasiones el turismo de aventura se considera sinónimo del turismo activo o también como incluido dentro de este último.

⁹⁹⁷ <<http://groups.msn.com/labota/otrasdireccionessenderistas.mswnw>>

⁹⁹⁸ <<http://www.masactividades.com/nego.php?ver=1>>

Yourdictionary.com,⁹⁹⁹ *Onelook.com*,¹⁰⁰⁰ *Diccionarios en Internet de las Ciencias Económicas y Sociales*,¹⁰⁰¹ *Dictionary.com*,¹⁰⁰² *Lexicoool*,¹⁰⁰³ *Glossare und Lexica*,¹⁰⁰⁴ entre otros.



Figura 42. Página principal del directorio *Yourdictionary.com*.

Estos directorios de diccionarios disponen de gran cantidad y variedad de diccionarios, glosarios y fuentes terminológicas en general; sin embargo, nos encontramos los mismos problemas que anteriormente: por un lado, hay una disponibilidad escasa de fuentes terminológicas sobre turismo en general y, especialmente sobre el segmento de aventura, y por otro, no cuentan con obras de este ámbito en el par de lenguas alemán-español, aunque sí se encuentran algunas para las lenguas inglés y español.

⁹⁹⁹ <<http://www.yourdictionary.com>>

¹⁰⁰⁰ <<http://www.onelook.com>>

¹⁰⁰¹ <<http://www.eumed.net/coursecon/dic/index.htm>>

¹⁰⁰² <<http://dictionary.reference.com/>>

¹⁰⁰³ <<http://www.lexicoool.com/index.asp?IL=3>>

¹⁰⁰⁴ <<http://www.mm-translations.com/DE/glossare-lexika.htm>>

En tercer y último lugar, realizamos la *búsqueda por palabras claves*. Esta búsqueda se basa en la utilización de palabras clave del tema concreto sobre el que se esté buscando información para realizar búsquedas en buscadores o motores de búsqueda en la Red. Este tipo de búsqueda es el más recurrente para realizar cualquier búsqueda en Internet, debido a sus numerosas ventajas, especialmente a su flexibilidad, su fácil manejo y su rapidez de respuesta. Además, se puede combinar con los otros tipos de búsquedas para ampliar el número de resultados, se pueden utilizar técnicas complejas como operadores booleanos o truncamientos para afinar la búsqueda, se pueden modificar los criterios de búsqueda¹⁰⁰⁵ para acotar los resultados, se recuperan gran cantidad de páginas, se pueden utilizar diferentes buscadores con indexaciones de páginas diferentes, se pueden utilizar numerosas ecuaciones de búsqueda, etc. Sin embargo, también cuenta con el inconveniente de presentar un exceso de documentos no pertinentes, lo que se conoce como ruido documental. Este ruido provoca una ralentización en la detección y selección de los recursos y obliga a menudo a desechar gran cantidad de páginas recuperadas. No obstante, si se realiza un buen uso de los buscadores y se conocen sus características y sus posibilidades de búsqueda, se pueden salvar en gran parte estos obstáculos.

En nuestra investigación, el uso de este tipo de búsqueda ha sido esencial para detectar y seleccionar las fuentes terminológicas recogidas en este apartado, y el ruido documental ha sido menor gracias al ajuste de criterios de búsqueda y al uso de los operadores booleanos del buscador.

A la hora de seleccionar las palabras claves, un proceso esencial en esta búsqueda, hemos utilizado las diferentes denominaciones del segmento del turismo de aventura, que hemos ido encontrando durante el estudio realizado a dicho segmento, además de aplicar nuestra propia intuición.¹⁰⁰⁶ Este tipo de búsqueda depende principalmente de las palabras clave utilizadas para realizar las búsquedas; por este motivo, hemos pretendido ser lo más objetivo posible y hemos realizado multitud de

¹⁰⁰⁵ En *Google*, se pueden modificar criterios de búsqueda en el apartado de «Búsqueda Avanzada». Por ejemplo, se puede determinar el país, el idioma, la fecha de publicación, el formato, etc. de los documentos que recupera.

¹⁰⁰⁶ Almahano Güeto (2002: 287-306) y Seghiri Domínguez (2006: 435) abogan por la intuición para la búsqueda de documentación mediante palabra clave.

búsquedas utilizando diferentes términos relacionados con el turismo de aventura en combinación con los términos «diccionario», «glosario», «Wörterbuch», «Glossar», «lexicon», «dictionary», «glossary», etc. y las técnicas booleanas propias del buscador. A modo ilustrativo, presentamos a continuación algunos de los términos utilizados para realizar las búsquedas por palabras clave:

| ESPAÑOL | ALEMÁN | INGLÉS |
|------------------------------------|--------------------|---------------------|
| Turismo de aventura | Abenteuertourismus | Adventure tourism |
| Turismo activo | Aktivtourismus | Active tourism |
| Deporte turismo deportivo | Sporttourismus | Sport tourism |
| Turismo extremo, turismo de riesgo | Extremer Tourismus | Extreme tourism |
| Turismo sostenible | Sanfter Tourismus | Sustainable tourism |
| Deporte extremo | Extremer Sport | Extreme sport |

Tabla 19. Ejemplo de palabras clave utilizadas como descriptores para la búsqueda por palabras clave.

Estos términos representan las principales denominaciones que hacen referencia al turismo de aventura y, por tanto, son los que se han utilizado para realizar las búsquedas por palabras clave. A pesar de haber utilizado estos descriptores concretos para la localización de recursos terminológicos y de haber ajustado los criterios de búsqueda del buscador, se obtuvo ruido documental, como por ejemplo páginas web sin recursos, recursos en otros idiomas, páginas obsoletas, etc., por lo que fue necesario una gran labor de filtrado para seleccionar los recursos.

| RECURSOS DEL SEGMENTO TURÍSTICO DE AVENTURA (inglés, español y alemán) | | |
|---|---|------------------------------------|
| 1 | Glosario sobre deportes de aventura | Catalán, español, francés e inglés |
| 2 | Glosario del deporte extremo | inglés, español |
| 3 | Enciclopedia de los deportes de aventura | español |
| 4 | Glosario de Aventureros | español |
| 5 | Extreme Sports Terminology | inglés |
| 6 | Glossary of Extreme Sports Terms | inglés |
| 7 | Sports and Games - Visual Dictionary Online (Merriem-Webster) | inglés |
| 8 | Sports und Spiele - Bild Wörterbuch (PONS) | alemán, inglés |

Tabla 20. Recursos terminológicos del sector turístico de aventura para las lenguas inglés, alemán y español.

Como queda patente con la información contenida en la tabla, los recursos disponibles para el segmento del turismo de aventura son extremadamente escasos, ya que solo hemos podido localizar cuatro que puedan ser utilizados para la traducción de textos de este dominio. Asimismo, ninguno de estos recursos hace referencia explícita a la terminología utilizada en este tipo de segmento turístico, sino que se refieren a las actividades deportivas que se realizan en él, lo cual podría ocasionar cierta incoherencia con la terminología turística, así como carencias en cuanto a terminología no incluida. Además, se observa la ausencia del alemán en todos, excepto en uno, lo cual se convierte en un gran problema si tenemos en cuenta que el mercado alemán es un mercado emisor de turistas muy fuerte, especialmente para España, donde este ocupa el segundo puesto y uno de los principales puestos en cuanto al turismo de aventura. A este respecto, podemos confirmar que los traductores profesionales de textos sobre el turismo de aventura carecen de recursos específicos para realizar sus traducciones directas o inversas para la combinación español-alemán. Por su parte, los traductores que trabajen con la combinación de inglés-español sí tienen a su disposición algunos recursos, ya sean estos monolingües o bilingües, que pueden utilizar para sus traducciones, aunque veremos más adelante el resultado de su evaluación mediante nuestra plantilla de evaluación y, por tanto, determinaremos si son o no fiables.

En este contexto, cabe preguntarse cómo proceden los traductores profesionales cuando han de traducir un texto turístico perteneciente a este segmento del turismo de aventura. Generalizando las respuestas obtenidas de los traductores que respondieron a las preguntas de nuestro estudio sobre los recursos terminológicos del sector turístico, podemos determinar que utilizan en gran medida los diccionarios generales, las búsquedas en *Google* y los textos paralelos o comparables. Además de estas fuentes de información, los traductores de textos turísticos de este segmento también podrán utilizar recursos terminológicos específicos de las diferentes actividades deportivas que se realizan dentro de este sector, como pueden ser las actividades subacuáticas, el barranquismo, etc., aunque estos recursos no satisfarán todas sus necesidades terminológicas y conceptuales, debido a las posibles variaciones que se pueden encontrar entre la terminología utilizada en el dominio turístico y la empleada en los campos específicos de dichas actividades.

En nuestra investigación, tras haber observado el número tan limitado de recursos relacionados con el segmento del turismo de aventura en su conjunto, consideramos necesario tener en cuenta este tipo de recursos también y, para ello, vamos a abrir nuestra búsqueda a los recursos terminológicos que traten sobre las diferentes actividades que se realizan en este segmento del turismo de aventura. De esta manera, ampliaremos el número de palabras clave utilizadas como se indica en la Tabla 24, combinadas, como antes, con palabras como «lexicon», «glosario», «terminología», «Wörterbuch», etc.

| ESPAÑOL | ALEMÁN | INGLÉS |
|-------------------------------------|---------------------|-------------------------|
| Buceo, snorkeling | Tauchen, Snorkeling | Diving, snorkeling |
| Montañismo, alpinismo | Bergsteigen | Mountaneering |
| Mountain Bike, bicicleta de montaña | Mountain biking | Mountain bike |
| Piragüismo, canotaje | Paddlen, Kanusport | Paddling, canoeing |
| Trekking, Senderismo | Trekking, Wanderung | Trekking, hiking |
| Globo, Globo aerostático | Heißluftballonflüge | Hot balloon, ballooning |
| Rafting | Rafting | Rafting |
| Escalada | Bergsteig | Climbing |
| Vuelo en ala delta | Drachenfliegen | Hang glider |
| Windsurf | Windsurf | Windsurfing |
| Kitesurf, flysurf | Kitesurf | Kitesurfing, flysurfing |

Tabla 21. Ejemplo de palabras clave utilizadas como descriptores para la búsqueda por palabras clave ampliada.

Con esta nueva búsqueda por palabras clave, se obtuvieron numerosos resultados que, de nuevo, tuvieron que ser filtrados para evitar el ruido y seleccionar los recursos pertinentes para cada actividad de aventura. A continuación, se pueden observar los recursos seleccionados, agrupados según la actividad a la que hacen referencia.

| RECURSOS DE MONTAÑISMO (inglés, español y alemán) | | |
|---|---|-----------------------------------|
| 1 | Glosario de Andarines.com | español |
| 2 | Glosario montañoero | español |
| 3 | Diccionario de terminología montañoera | español |
| 4 | Glosario de montaña y senderismo | español |
| 5 | Glosario de montaña | español |
| 6 | Vocabulario referido a la montaña (o el monte) | español |
| 7 | Glossary of Swansea University Mountaineering Club | inglés |
| 8 | Glossary of Rock Climbing and Mountaineering Terms | inglés |
| 9 | Mountaineering Dictionary | inglés |
| 10 | German - English Mountaineering Dictionary | alemán-inglés |
| 11 | Alpines Wörterbuch | alemán, inglés, italiano, francés |
| 12 | Terminología en la Normativa de Espacios Naturales Protegidos | español |

Tabla 22. Recursos terminológicos encontrados para la actividad de montañaismo.

| RECURSOS DE ESCALADA (inglés, español y alemán) | | |
|---|--|--|
| 1 | Climbing Glossary | inglés, alemán, francés, italiano, español |
| 2 | Glossar - Klettern | alemán |
| 3 | Lexikon von Fachbegriffen rund um's Klettern | alemán |
| 4 | Lexikon der Kletterbegriffe | alemán |
| 5 | Kletter - Lexikon | alemán |
| 6 | Climbing Glossary | inglés |
| 7 | Climbing Glossary of Terms | inglés |

Tabla 23. Recursos terminológicos encontrados para la actividad de escalada.

| RECURSOS DE ESPELEOLOGÍA (inglés, español y alemán) | | |
|--|---|---|
| 1 | Glossary of caving terms | inglés |
| 2 | A Lexicon of Cave and Karst Terminology with Special Reference to Environmental Karst Hydrology | inglés |
| 3 | Caving Definitions - A Glossary of Terms | inglés |
| 4 | Espeleocedario | español |
| 5 | Glosario de Turismo Subterráneo | español |
| 6 | Lexikon der Begriffe des Karstes | alemán |
| 7 | The Caver's Multi-Lingual Dictionary | inglés, búlgaro, alemán, español, francés, croata, húngaro, italiano, portugués, rumano, serbio |

Tabla 24. Recursos terminológicos encontrados para la actividad de espeleología.

| RECURSOS DE BICICLETA (inglés, español y alemán) | | |
|---|-----------------------------------|---|
| 1 | Diccionario de la lengua ciclista | español |
| 2 | Sheldon Brown's Bicycle Glossary | inglés |
| 3 | Mountain Bike Terms | inglés |
| 4 | Bike Lexicon | alemán |
| 5 | Bicycle Dictionary | inglés, alemán, francés, italiano, español, holandés, etc. |
| 6 | Diccionario ciclista plurilingüe | español, alemán, esperanto, francés, inglés, polaco, portugués y ruso. |
| 7 | Multilingual Bicycle's Vocabulary | español, alemán, esperanto, francés, inglés, polaco, portugués, sueco, etc. |

Tabla 25. Recursos terminológicos encontrados para la actividad de bicicleta.

| RECURSOS DEL MUNDO ECUESTRE (inglés, español y alemán) | | |
|---|--|---|
| 1 | Glosario de términos ecuestres | español |
| 2 | Equine Multilingual Dictionary | inglés, francés, español, alemán |
| 3 | Pferd & Reiter - Wörterbuch | alemán, inglés, francés, italiano y español |
| 4 | Diccionario multilingüe del mundo ecuestre | español, inglés, francés y alemán |
| 5 | WebTerm - Pferde | alemán, inglés |

Tabla 26. Recursos terminológicos encontrados para la actividad del mundo ecuestre

| RECURSOS DE VUELO SIN MOTOR (inglés, español y alemán) | | |
|---|---|---------------|
| 1 | Diccionario de la Lengua Española (Parapente) | español |
| 2 | Skydiving Glossary | inglés |
| 3 | Skydiving Information - Glossary | inglés |
| 4 | WebTerm - Gliders | alemán-inglés |
| 5 | English-German Sailplane Dictionary | alemán-inglés |

Tabla 27. Recursos terminológicos encontrados para la actividad de vuelo sin motor.

| RECURSOS DE BUCEO (inglés, español y alemán) | | |
|---|---|--|
| 1 | Tauchsport Lexikon | alemán |
| 2 | WebTerm - Tauchen | alemán, inglés |
| 3 | Diccionario multilingüe de buceo y actividades subacuáticas | español, inglés, francés, italiano, holandés, alemán |
| 4 | Bienvenue dans le dictionnaire multi-langues de la plongée | francés, alemán, inglés, español, noruego, sueco |

Tabla 28. Recursos terminológicos encontrados para la actividad de buceo.

| RECURSOS DE PIRAGÜISMO, WINDSURF Y ACTIVIDADES ACUÁTICAS (inglés, español y alemán) | | |
|---|--|--|
| 1 | Paddlers Wörterbuch | alemán |
| 2 | Kanu Begriffe | alemán |
| 3 | Deutscher Kanu Verband | alemán |
| 4 | Paddlers Wörterbuch | alemán |
| 5 | Paddling Terms | inglés |
| 6 | Paddlers' Dictionary | inglés |
| 7 | Diccionario de Deportes acuáticos | español |
| 8 | Vocabulario del Windsurfista | inglés, español |
| 9 | Windsurf-Dictionary | inglés, alemán, polaco, italiano, francés, español |
| 10 | Glosario sobre deportes náuticos ¹⁰⁰⁷ | catalán, español, francés, inglés |

Tabla 29. Recursos terminológicos encontrados para las actividades de piragüismo, windsurf y actividades acuáticas en general.

Con la clasificación de los recursos de las actividades que se ofrecen en el segmento del turismo de aventura, podemos observar con claridad la variedad de actividades que se practican dentro de este segmento y que, además, crecen constantemente. Asimismo, encontramos los diferentes recursos disponibles que tienen los traductores profesionales a su alcance para realizar traducciones de textos turísticos en este dominio. A pesar de que estos recursos ofrecen cierta ayuda al traductor de textos turísticos en este segmento, es necesario tener en cuenta que su terminología es demasiado amplia y específica del sector al que se dedican, por ejemplo, un diccionario de escalada incluirá terminología útil para el traductor de textos turísticos como los tipos de escalada posibles, el equipamiento necesario, etc., pero, a su vez, gran cantidad de terminología que pertenece al ámbito de la escalada como deporte (tipos de suelo,

¹⁰⁰⁷ El segmento del turismo náutico y, por tanto, los recursos referentes a él se consideran independientes del segmento del turismo de aventura y de sus actividades. No obstante, en este caso se ha considerado oportuno incluir este recurso sobre las actividades náuticas, puesto que incluye las actividades acuáticas propias del segmento del turismo de aventura.

equipamiento para profesionales, etc.) y que no son pertinentes para este tipo de traducciones.

A continuación, pasaremos a clasificar estos recursos encontrados y mostraremos la calidad de cada uno y su adecuación para la tarea traductora en este segmento y, así, constatar nuestra hipótesis de que en el segmento del turismo de aventura existe una carencia de recursos a la par que una baja calidad en los existentes.

3.3.6.2. Clasificación de los recursos terminológicos encontrados

La clasificación de las fuentes terminológicas incluidas en las tablas anteriores (Tabla 20 y Tablas 22-29) se ha realizado conforme a los criterios de clasificación presentados en el capítulo anterior (cf. 2.4.1.), a saber: según la presentación, según el número de lenguas, según el tipo de soporte, según la ordenación de sus contenidos, según la función, según el grupo de destinatarios, según el tipo de unidades terminológicas, según la extensión en la especialidad, según la fuente de información utilizada y según la aplicación.

Esta clasificación, sin embargo, no ha podido realizarse por completo ni se ha podido aplicar a todas las fuentes, debido a la falta de información que ofrecían sobre algunos aspectos y que no ha sido posible averiguar.¹⁰⁰⁸ Este ha sido el caso de los criterios referentes al grupo de destinatario y a la fuente de información utilizada.

Con respecto al primer criterio de clasificación, los recursos terminológicos no indican en su gran mayoría al destinatario al que van dirigidos, por lo que se considera un destinatario general interesado en la práctica de estas actividades. Asimismo, ninguno de los recursos encontrados incluye a los traductores o a otros profesionales de la lengua entre su grupo de destinatarios, lo que, sumado al número tan bajo de estos recursos, indica la falta de atención que se les presta a los profesionales de la traducción en este sector.

¹⁰⁰⁸ En el Anexo V, se pueden consultar los recursos seleccionados y su clasificación completa según los diferentes criterios aplicados.

Otro factor importante que influye en la labor traductora consiste en la finalidad del recurso, es decir, en el uso para el que se ha elaborado. En general, los profesionales de la traducción necesitan que los recursos les proporcionen información para la codificación de los términos, es decir, necesitan que los recursos sirvan de ayuda para la redacción y la utilización del término de forma adecuada. Por ello, los traductores necesitan información lingüística, pragmática, etc., como ejemplos, usos, fraseología, dominio, sinónimos, etc., aparte de la definición y el equivalente de traducción, para facilitar su labor y optimizar sus resultados mediante el uso correcto del término en la traducción que estén realizando, ya sea directa o inversa. De esta forma, necesitan recursos especializados cuya aplicación vaya destinada a la codificación. En el caso de los recursos seleccionados para este estudio, todos están destinados a la descodificación, o dicho de otro modo, a la comprensión del término turístico y, por tanto, no a la codificación. Este hecho se debe fundamentalmente a que los recursos seleccionados no ofrecen información de codificación (contexto, frecuencia de uso, marcas pragmáticas, etc.), a excepción de la definición y de hipervínculos a otros conceptos (en su caso), lo que dificulta en gran medida el proceso traductor, como ya hemos comentado.

Además de la falta de información acerca de los destinatarios que veíamos anteriormente, los recursos no indican tampoco el proceso de elaboración al que han sido sometidos, es decir, no se indica la fuente originaria que se ha utilizado para elaborar el recurso ni tampoco los autores de este, lo que reduce en gran medida la calidad del recurso así como su fiabilidad. En algunos casos, incluso la elaboración es colaborativa, es decir, los términos incluidos y sus equivalentes o definiciones han sido gestionados por varias personas diferentes sin especificar, como es el caso del *Glosario de Aventureros* y del *Mountaineering Dictionary*.

Con respecto al número de lenguas, se puede observar en las tablas anteriores que predominan los recursos monolingües de español, inglés o alemán y multilingües de más de 4 lenguas diferentes, incluyendo las tres lenguas de esta investigación (español, inglés y alemán). De esta manera, tenemos 15 en inglés, 14 recursos en español, y 11 en alemán y 12 recursos multilingües que combinan diferentes lenguas, además de las tres

requeridas.¹⁰⁰⁹ En este sentido, podemos afirmar que, a pesar de no haber localizado apenas recursos bilingües, la comprensión de términos y la traducción de equivalentes podrían estar garantizada, siempre y cuando sea positiva la evaluación de estos recursos.

Con respecto al soporte de los recursos, también se puede comprobar la existencia casi absoluta de recursos en línea, frente a los recursos en papel y en otro formato electrónico, como es el CD-ROM. De los recursos que se han recopilado por seguir los criterios iniciales, a saber: estar publicados o creados a partir del año 2000 y versar sobre alguna de las actividades de este segmento turístico, solamente dos se han seleccionado en formato papel (Tabla 30), mientras que el resto se pueden consultar en línea de forma gratuita.¹⁰¹⁰ En este sentido, podemos corroborar que Internet ocupa el lugar preferido para crear recursos terminológicos, debido a las ventajas que presenta frente a los recursos en papel o en otro formato. Entre las ventajas de este medio, destacan su facilidad de acceso, de publicación, de actualización, de creación y de mantenimiento, lo que lo convierten en favorito para la mayoría de usuarios. Además, es necesario tener en cuenta la rápida aparición y evolución de este tipo de actividades, lo que hace que necesite de un soporte fácil de modificar y de actualizar.

| RECURSOS EN PAPEL (inglés, español y alemán) |
|--|
| Enciclopedia de los deportes de aventura. La guía más completa de los deportes de aventura y riesgo (2004) |
| Diccionario multilingüe del mundo ecuestre (2004) |
| Tauchsport Lexikon (2005) |

Tabla 30. Recursos seleccionados en formato papel.

Según la ordenación de sus contenidos, los recursos siguen en su gran mayoría una ordenación alfabética, aunque nos encontramos con algunos recursos que ofrecen una ordenación diferente. Este es el caso de los diccionarios visuales de *Merrien-*

¹⁰⁰⁹ Estos recursos son los que han obtenido mayor puntuación. No obstante, contamos, además, con dos recursos bilingües de inglés-español y cuatro, de alemán-inglés y con tres recursos multilingües que no contienen alguno de las tres lenguas de trabajo.

¹⁰¹⁰ Entre los recursos terminológicos que versan sobre el dominio turístico en general en estas tres lenguas de trabajo (Tabla 18) se encuentran muchos más recursos en formato papel, aunque estos no han sido tenidos en cuenta.

Webster (Figura 43) y de *PONS*, que ofrecen una ordenación temática y visual de forma monolingüe y bilingüe, respectivamente. Asimismo, otros recursos como el *Glossary of Extreme Sports Terms* ofrecen una búsqueda temática inicial por actividad, aunque su ordenación dentro de cada actividad es de forma alfabética. Por su parte, el resto de recursos optan por una ordenación alfabética o mediante búsqueda por términos.

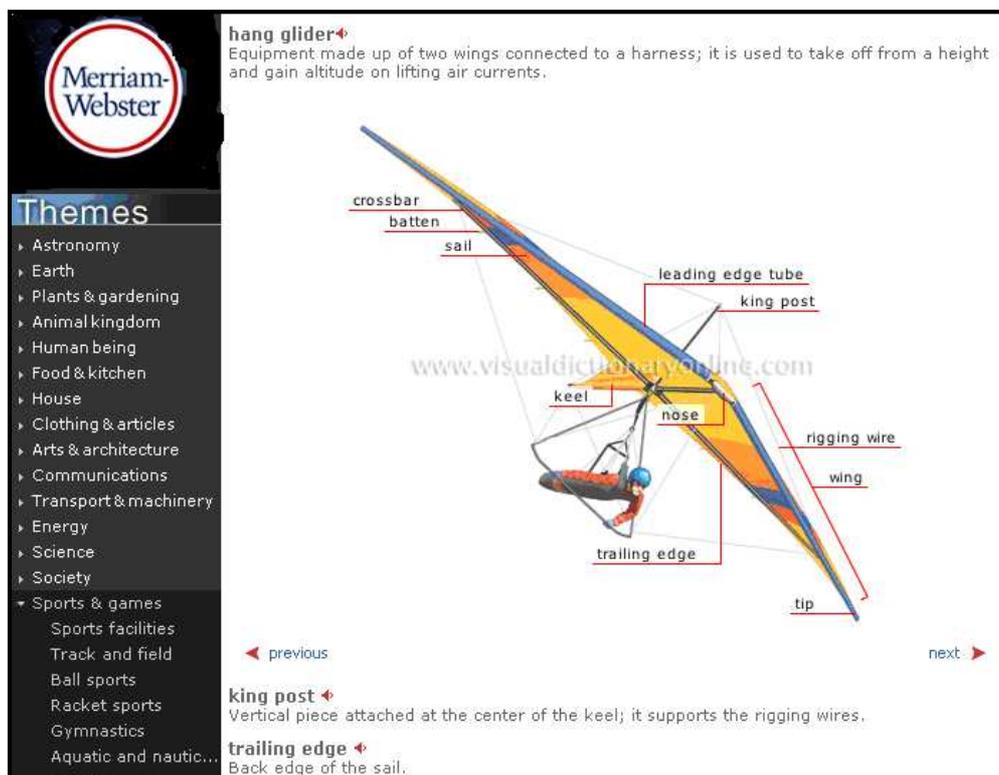


Figura 43. Pantalla de la búsqueda «Aerial Sports» en el diccionario visual de *Merriem Webster* (sección *Games and Sports*).

Asimismo, todos los recursos son descriptivos según su normalización. En este sentido, no existe ningún recurso que haya sido publicado o elaborado por una institución oficial que prescriba la terminología que se debe utilizar en el ámbito turístico, por lo que se puede constatar la ausencia de normalización en la terminología empleada en este sector. Esta falta de normalización, o al menos de cierta homogeneización en la terminología utilizada, provoca el empleo de gran cantidad de anglicismos, así como la presencia elevada de variación denominativa, lo que se conoce como cuasisinónimos, como es el caso de «turismo de aventura», «turismo activo» y

«turismo de deportes»; «barranquismo» y «descenso de barrancos»; «senderismo» y «trekking»; «flysurf» y «kitesurf», entre otros muchos, aunque no todos son correctos.

Con respecto al tipo de unidades terminológicas que forman los recursos, debemos confirmar que los recursos seleccionados no presentan por lo general información paradigmática ni sintagmática, es decir, no incluyen sinónimos, antónimos, fraseología, etc., por lo que el criterio de clasificación en esta ocasión no puede ser utilizado. Esto se debe principalmente al hecho de que los recursos están diseñados para ayudar a la descodificación de los términos y no a la codificación, por lo que de nuevo se presenta un obstáculo para los traductores especializados en este ámbito.

El criterio sobre la extensión en la especialidad sí es más relevante en este caso, ya que encontramos recursos que hacen referencia al conjunto de actividades de aventura practicados en este segmento, recopilados en la Tabla 20, y recursos que hacen referencia a una o varias actividades, que se pueden consultar en las Tablas 22-29. Como indicábamos anteriormente, los recursos que tratan una o varias actividades son menos útiles para la traducción de textos turísticos promocionales de este segmento, puesto que están formados por terminología demasiado específica y exclusiva de dicha actividad. No obstante, al haber pocos recursos que traten todas las actividades en su conjunto, consideramos que podrían servir de cierta ayuda a los traductores.

Por último, clasificaremos los recursos según la presentación que ofrecen. Como veíamos en el capítulo anterior, los recursos terminológicos se pueden dividir en ocho grupos, a saber: diccionarios especializados, diccionarios visuales, enciclopedias, vocabularios, glosarios, léxicos, tesauros, siglas y acrónimos, abreviaturas y símbolos, bases de datos y corpus textuales (cf. 2.4.1.1.).

En nuestra búsqueda de recursos terminológicos del segmento del turismo de aventura, hemos encontrado principalmente léxicos y glosarios, aunque también hemos localizado diccionarios especializados, dos diccionarios visuales y una enciclopedia.¹⁰¹¹

¹⁰¹¹ Cabe recordar qué se entiende por estos tipos diferentes de recursos: los diccionarios especializados contienen un número de entradas muy elevado y proporcionan una información muy precisa sobre el ámbito de especialidad, los diccionarios visuales son diccionarios especializados que indican los términos sobre una imagen, los léxicos son recursos que ofrecen equivalentes de traducción en una o más lenguas sin definiciones, los glosarios tratan sobre ámbitos y jergas muy concretos y contienen un menor número

Por su parte, también se ha encontrado un corpus textual que incluye textos de este segmento turístico, enmarcado en el proyecto *La contratación turística electrónica multilingüe como mediación intercultural: aspectos legales, traductológicos y terminológicos* (N.º de ref.: HUM-892. Junta de Andalucía), pero no se ha incluido en este estudio por tener acceso restringido solo para los investigadores del proyecto y, por tanto, no encontrarse al alcance de los traductores profesionales.¹⁰¹² En lo referente a siglas y acrónimos, abreviaturas y símbolos, no hemos detectado ningún recurso de esa naturaleza basado en este segmento turístico, aunque sí hay disponibles fuentes generales de siglas y acrónimos, abreviaturas y símbolos¹⁰¹³ que incluyen términos utilizados en el sector turístico y que pueden servir de ayuda, como por ejemplo la sigla «AGB» («Allgemeine Gesellschaftsbedingungen», en alemán). En el Anexo V, se puede observar la distribución de estos recursos según el tipo al que pertenecen.

Llegados a este punto, hemos finalizado la clasificación de los recursos según los diferentes criterios de clasificación expuestos y pasaremos al apartado de evaluación de los recursos localizados y seleccionados. Como indicamos con anterioridad, la completa clasificación de los recursos se puede consultar en la tabla incluida como Anexo V.

3.3.6.3. Evaluación de los recursos terminológicos turísticos

En este apartado, nos dedicaremos a la evaluación de los diferentes recursos terminológicos que hemos seleccionado y clasificado en los pasos anteriores. Para ello, presentaremos, a modo de ejemplo, una evaluación de dos de los recursos seleccionados, con objeto de ejemplificar los pasos seguidos con todos los recursos terminológicos que se han indicado en el apartado anterior, mostrar cómo se ha aplicado

de términos que los vocabularios, y los vocabularios son una recopilación de términos específicos de un ámbito de especialidad que pueden ofrecer equivalentes de traducción en una o más lenguas.

¹⁰¹² <<http://clg.wlv.ac.uk/malaga/>>

¹⁰¹³ Algunos de los recursos de abreviaturas y acrónimos disponibles actualmente en español, inglés y alemán: *Abkürzungen* (alemán): <<http://www.abkuerzungen.de/>>, *Acronyms and Abbreviations* (multilingüe): <<http://www.silmaril.ie/cgi-bin/uncgi/acronyms>>, *Acronym Finder* (multilingüe): <<http://www.acronymfinder.com/>>, *Spanisch-Deutsch Abkürzungen*: <http://www.anabell.de/spanien/spanische_deutsch_abkuerzungen.html>, *Abbreviations* (multilingüe): <<http://www.abbreviations.com/>>.

la plantilla de evaluación que se presentó en el capítulo anterior (cf. 2.4.2.) y cómo se ha realizado la valoración del 0 al 2 de cada apartado, según las características de cada recurso.

Para realizar una demostración, hemos seleccionado dos recursos diferentes: por un lado, el *Glosario sobre deportes de aventura*, de TermCat, un recurso que ha obtenido una de las puntuaciones más alta de todos los recursos seleccionados (17 puntos), lo que la sitúa en el nivel medio de nuestra escala con una fiabilidad medio-alta; y, por otro, un recurso con una puntuación muy baja (4 puntos), el *Bicycle Dictionary*, que lo sitúa en el nivel más inferior de la escala, con una fiabilidad baja. De esta forma, se podrá comprobar las diferencias existentes entre dos recursos existentes así como el uso diverso de la puntuación según la calidad del recurso.

Comenzaremos con el *Glosario sobre deportes de aventura*, un glosario especializado, multilingüe, en soporte electrónico, específico de las actividades que se realizan en el segmento del turismo de aventura, descriptivo, con una ordenación alfabética y destinado a un público en general.

La primera parte de la plantilla está dedicada a la descripción del recurso y solo reciben puntuación tres indicadores: la *Autoría*, la *Fecha de publicación* o de actualización y la especialización (*Especializado en...*).

| DESCRIPCIÓN | | Puntuación |
|---|--|------------|
| Título | Glosario sobre deportes de aventura | |
| Autoría (0, 1, 2) | TermCat | 2 |
| URL (si procede) | http://www.termcat.cat/dicci/esports_aventura/index.html | |
| Fecha de publicación/actualización (0, 1, 2) | 2006 | 2 |
| Lenguas | <input type="checkbox"/> Monolingüe <input type="checkbox"/> Bilingüe <input checked="" type="checkbox"/> Multilingüe (español, inglés, francés y catalán) | |
| Soporte | <input type="checkbox"/> Papel <input checked="" type="checkbox"/> Electrónico | |
| Especialización en TURISMO DE AVENTURA (0, 1, 2) | <input checked="" type="checkbox"/> Sí, aunque no en el turismo sino en las actividades de aventura que se realizan en él. <input type="checkbox"/> No | 2 |
| Puntuación total = | | 6 |

Tabla 31. Valoración del primer parámetro de *Glosario sobre deportes de aventura*.

Este recurso ha recibido una puntuación máxima en la autoría, debido a dos razones: en primer lugar, se especifica su autor (TermCat), y en segundo lugar, su autor es un centro de terminología de reconocida relevancia en el ámbito de la creación de recursos lexicográficos y terminológicos, por lo que ofrece confianza en lo referente a la calidad. Asimismo, se le otorga también la puntuación máxima en el apartado de *Fecha de publicación*, puesto que indica la última actualización que se realizó en este recurso, la cual se puede considerar, además, reciente (año 2006). Por último, el apartado de especialización se puntúa también con la máxima puntuación, ya que el recurso trata sobre nuestro dominio de interés e incluye gran parte de los términos y actividades relacionados con el turismo de aventura y, por tanto, no se tiene en consideración que no haga referencia explícita al turismo de aventura.

En el segundo parámetro, la macroestructura, el recurso recibe la puntuación de 7 puntos (Tabla 32).

| MACROESTRUCTURA | | Puntuación |
|-------------------------------|---|------------|
| Indicaciones (0, 1, 2) | <input checked="" type="checkbox"/> Sí <input type="checkbox"/> No | 2 |
| Número de entradas (0, 1, 2) | 85 | 1 |
| Ausencia de errores (0, 1, 2) | <input checked="" type="checkbox"/> Sí <input type="checkbox"/> No | 2 |
| Ergonomía (0, 1, 2) | <input checked="" type="checkbox"/> Sí <input type="checkbox"/> No | 2 |
| Puntuación total = | | 7 |

Tabla 32. Valoración del segundo parámetro de *Glosario sobre deportes de aventura*.

Este recurso cumple en este apartado con todos los requisitos exigidos para considerarlo una fuente de calidad, a saber: incluye indicaciones sobre el uso del recurso y las abreviaturas y anotaciones utilizadas; su número de entradas no es muy elevado, pero contiene la terminología de los deportes de aventuras más importantes y, por tanto, se trata de un recurso muy completo; no presenta ningún tipo de error, ni gramatical ni de contenido u ortográfico; su acceso y manejo son muy sencillos y, además, permite realizar varios tipos de búsquedas (por orden alfabético o por búsqueda de términos). De esta forma, su puntuación debe ser muy positiva.

El tercer parámetro, la microestructura, recibe una puntuación de 5 puntos (sobre 16 puntos), por lo que su evaluación en este tercer parámetro es bastante baja.

| MICROESTRUCTURA | | Puntuación |
|---|---|------------|
| Definición (0, 1, 2) | <input checked="" type="checkbox"/> Sí, en catalán <input type="checkbox"/> No | 1 |
| Información gramatical (0, 1, 2) | <input checked="" type="checkbox"/> Sí (género y número, cuando es plural) <input type="checkbox"/> No | 2 |
| Uso de las unidades (0, 1, 2) | <input type="checkbox"/> Sí <input checked="" type="checkbox"/> No | 0 |
| Información sintagmática (locuciones, colocaciones y enunciados fraseológicos) (0, 1, 2) | <input type="checkbox"/> Sí <input checked="" type="checkbox"/> No | 0 |
| Información paradigmática (sinónimos, antónimos) (0, 1, 2) | <input checked="" type="checkbox"/> Sí, solo variantes denominativas <input type="checkbox"/> No | 1 |
| Ejemplos (0, 1, 2) | <input type="checkbox"/> Sí <input checked="" type="checkbox"/> No | 0 |
| Remisiones/hiperenlaces (0, 1, 2) | <input type="checkbox"/> Sí <input checked="" type="checkbox"/> No | 0 |
| Tipos diferentes de entradas (sustantivo, adjetivo, verbo, adverbio, acrónimos, abreviaturas) (0, 1, 2) | <input checked="" type="checkbox"/> Sí, una abreviatura <input type="checkbox"/> No | 1 |
| Puntuación total = | | 5 |

Tabla 33. Valoración del tercer parámetro de *Glosario sobre deportes de aventura*.

En este parámetro, el recurso recibe la puntuación máxima en solo un indicador de los ocho que lo forman, a saber: información gramatical, gracias a su inclusión de datos gramaticales de las diferentes unidades. Por su parte, el indicador de la definición recibe un punto, debido a que las definiciones están escritas en catalán, es decir, el recurso es multilingüe en lo que respecta a los equivalentes de traducción pero las definiciones de las entradas las especifica solo en catalán, por haber sido el idioma de trabajo del proceso terminológico. A pesar de la calidad de las definiciones, no podemos otorgarle la puntuación máxima por este hecho, ya que dificulta la comprensión del término y se la impide a personas que no conozcan la lengua inglesa. En nuestro caso, además, sería recomendable que las definiciones se ofrecieran en español, en alemán o en inglés para otorgarles la puntuación máxima de dos puntos, ya que, recordemos, los recursos tienen unos destinatarios y unos objetivos concretos. Asimismo, el parámetro de la información paradigmática y tipos de entrada diferentes obtienen también un punto en la evaluación, ya que el recurso incluye información de este tipo pero no es constante a lo largo de todo el recurso. Por ejemplo, en el caso de los sinónimos incluye varios ejemplos de este tipo de variación («tabla de windsurf» y «tabla deslizadora a vela» o «mushing» y «trineo con perros», etc.), y con respecto a las abreviaturas solo se incluye una, BTT. Por último, el recurso recibe la puntuación mínima en el resto de indicadores,

a saber: información sintagmática, uso de las unidades, ejemplos y remisiones/hiperenlaces. Esto se debe a la carencia de esta información en el recurso, es decir, este recurso no incluye ninguno de estos cuatro tipos de información, los cuales se consideran importantes a la hora de evaluar un recurso de esta naturaleza.

Tras el cómputo total de indicadores, este recurso consigue una puntuación de 18 puntos, lo que lo sitúa como uno de los recursos con mayor puntuación de los evaluados según la escala de fiabilidad propuesta en este estudio.¹⁰¹⁴

Una vez finalizada la evaluación del recurso *Glosario sobre deportes de aventura*, realizaremos la misma operación con el otro recurso seleccionado, el *Bicycle Dictionary*, que se trata de un léxico, multilingüe, en soporte electrónico, dedicado a la terminología de la bicicleta, descriptivo y ordenado de forma alfabética. En este recurso, no se indica el grupo de destinatarios al que va dirigido; no obstante, hace pensar que se dedica a un público lego en la materia.

Como en el recurso anterior, la primera parte de la plantilla está dedicada a su descripción y solo reciben puntuación tres indicadores: la autoría, la fecha de publicación y la especialización.

¹⁰¹⁴ Recordemos que la escala de fiabilidad propuesta está compuesta por cuatro niveles según puntuación: 1. Nivel bajo, de 0 a 8 puntos; 2. Nivel medio-bajo, de 9 a 15 puntos; 3. Nivel medio-alto, de 16 a 23 puntos, y 4. Nivel alto, de 24 a 30 puntos (cf. 2.4.2.1.). Véase también el Anexo V para comparar con otros recursos.

| DESCRIPCIÓN | | Puntuación |
|---|---|------------|
| Título | Bicycle Dictionary | |
| Autoría (0, 1, 2) URL (si procede) | Paul van Roekel y Anja de Graaf http://www.cyclingaroundtheworld.nl/tipsadvice/bikevoca_results.asp | 1 |
| Fecha de actualización (0, 1, 2) | No se indica | 0 |
| Lenguas | <input type="checkbox"/> Monolingüe <input type="checkbox"/> Bilingüe <input checked="" type="checkbox"/> Multilingüe (inglés, francés, alemán, italiano, español, finés, japonés, noruego, portugués, sueco, holandés) | |
| SopORTE | <input type="checkbox"/> Papel <input checked="" type="checkbox"/> Electrónico | |
| Especialización en TURISMO (0, 1, 2) | <input checked="" type="checkbox"/> Sí, en bicicleta <input type="checkbox"/> No | 1 |
| Puntuación total = | | 2 |

Tabla 34. Valoración del primer parámetro de *Bicycle Dictionary*

En este parámetro, el recurso ha recibido dos puntos: un punto en dos de los tres indicadores y no ha recibido ninguna puntuación máxima en ninguno de los indicadores que forman el primer bloque.

En el indicador de la autoría, ha recibido un punto, debido, por un lado, a que el léxico aparece en una página personal donde se indican los datos de las dos personas que lo han elaborado, Paul van Roekel y Anja de Graaf (por lo que hay que puntuarlo), pero por otro, estos autores no presentan ningún respaldo por parte de instituciones, organizaciones o entidades oficiales y de prestigio, lo cual le resta fiabilidad al recurso y no permite otorgarle la máxima puntuación. Asimismo, se le otorga también la puntuación de un punto en el apartado de especialización, puesto que el recurso trata sobre la terminología de la bicicleta y sus partes pero no incluye directamente terminología del sector turístico. A pesar de no incluir terminología del sector turístico, se ha seleccionado este recurso y se le aplica un punto en este apartado, debido a que el cicloturismo y el deporte de bicicleta son unas actividades muy frecuentes en el turismo de aventura.

En el segundo parámetro, la macroestructura, el recurso recibe un punto, ya que se le puntúa con la mínima puntuación de cero en tres de los cuatro indicadores que forman esta parte (Tabla 35).

| MACROESTRUCTURA | | Puntuación |
|-------------------------------|---|------------|
| Indicaciones (0, 1, 2) | <input type="checkbox"/> Sí <input checked="" type="checkbox"/> No | 0 |
| Número de entradas (0, 1, 2) | Aprox. 25 | 0 |
| Ausencia de errores (0, 1, 2) | <input type="checkbox"/> Sí <input checked="" type="checkbox"/> No | 0 |
| Ergonomía (0, 1, 2) | <input checked="" type="checkbox"/> Sí <input type="checkbox"/> No | 1 |
| Puntuación total = | | 1 |

Tabla 35. Valoración del segundo parámetro de *Bicycle Dictionary*

El recurso recibe estas puntuaciones tan bajas por no cumplir con los requisitos exigidos en este parámetro, a saber: no incluye indicaciones de uso, el número de entradas de este parámetro es muy inferior al recomendado y presenta gran cantidad de errores, tanto de ortografía («clavijga», «camera», «campainha», «Schraubschluessel») como de traducción («camera a aire», como equivalente a «Schlauch», cuando debería ser «tubo (de plástico)»; «bahco» como equivalente a «Schraubschluessel», cuando debería ser «llave de tuerca», o «campainha» como equivalente a «Glocke», cuando debería ser «bocina»), por citar solo algunos casos.

Por último, con respecto a su ergonomía, el recurso recibe un punto, ya que su consulta y acceso se consideran aceptables, aunque no es muy dinámico. Al incluir un número muy limitado de términos, el recurso los muestra todos directamente al seleccionar los idiomas en los que se desea realizar la búsqueda.

El tercer parámetro, la microestructura, recibe una puntuación muy baja también (0 de 10 puntos), debido a la falta de información en las entradas de los términos.

| MICROESTRUCTURA | | Puntuación |
|---|---|------------|
| Definición (0, 1, 2) | <input type="checkbox"/> Sí <input checked="" type="checkbox"/> No | 0 |
| Información gramatical (0, 1, 2) | <input type="checkbox"/> Sí <input checked="" type="checkbox"/> No | 0 |
| Uso de las unidades (0, 1, 2) | <input type="checkbox"/> Sí <input checked="" type="checkbox"/> No | 0 |
| Información sintagmática (locuciones, colocaciones y enunciados fraseológicos) (0, 1, 2) | <input checked="" type="checkbox"/> Sí <input type="checkbox"/> No | 0 |
| Información paradigmática (sinónimos, antónimos) (0, 1, 2) | <input type="checkbox"/> Sí <input checked="" type="checkbox"/> No | 0 |
| Ejemplos (0, 1, 2) | <input type="checkbox"/> Sí <input checked="" type="checkbox"/> No | 0 |
| Remisiones/hiperenlaces (0, 1, 2) | <input type="checkbox"/> Sí <input checked="" type="checkbox"/> No | 0 |
| Tipos diferentes de entradas (sustantivo, adjetivo, verbo, adverbio, acrónimos, abreviaturas) (0, 1, 2) | <input type="checkbox"/> Sí <input checked="" type="checkbox"/> No | 0 |
| Puntuación total = | | 0 |

Tabla 36. Valoración del tercer parámetro de *Bicycle Dictionary*.

En este parámetro, el recurso recibe la puntuación mínima (cero puntos) en los ocho indicadores que se incluyen, a saber: definición, información gramatical, uso de las unidades, información sintagmática y paradigmática, ejemplos, remisiones/hiperenlaces y tipos de entrada diferentes, lo que indica la baja calidad de este recurso. Esto se debe al hecho de que este recurso carece de estos tipos de información, es decir, no proporciona definiciones de los términos, ni información gramatical o paradigmática, ni ejemplos, etc., los cuales se consideran elementos esenciales a la hora de evaluar un recurso terminológico.

En total, el recurso obtiene la puntuación de tres puntos, por lo que se encuentra en la escala de fiabilidad baja (de 0 a 8 puntos). En este sentido, no se recomienda el uso de este recurso a los profesionales de la traducción por una falta de calidad y de fiabilidad, como ya hemos visto durante su evaluación.

Una vez expuesta la metodología seguida para la evaluación de los recursos terminológicos, consideramos que se ha esclarecido la aplicación de nuestra plantilla de evaluación, así como la valoración mediante las puntuaciones de los diferentes indicadores. Para conocer las puntuaciones obtenidas por el resto de recursos seleccionados para esta investigación, remitimos al Anexo V, donde se podrá encontrar

la clasificación de los recursos según los criterios seguidos y el resultado de la puntuación de cada recurso.

En general, después de haber realizado la evaluación de todos los recursos seleccionados, siguiendo la metodología expuesta anteriormente, podemos confirmar nuestras hipótesis iniciales, a saber: los recursos disponibles para el segmento del turismo de aventura son escasas y de baja calidad. La primera hipótesis se constata con el resultado infructuoso de la recopilación de recursos realizada sobre el turismo de aventura. Tanto es así que nos hemos vistos obligados a ampliar la búsqueda a los deportes de aventura y extremos en general, así como a los recursos que tratan la terminología de las diferentes actividades que se ofrecen en este segmento turístico.

Por su parte, la segunda hipótesis se verifica con los resultados obtenidos después de aplicar la plantilla de evaluación y la escala de fiabilidad establecida en esta investigación. Se aprecia que las puntuaciones obtenidas después de haber realizado la evaluación a los 66 recursos seleccionados son muy bajas y que la mayoría de ellos (42 recursos) se encuentra en la escala de 9 a 15 puntos, lo que nos indica que la calidad de esos recursos resulta medio-baja. El segundo lugar lo ocupan los 13 recursos que han obtenido una puntuación en la escala de fiabilidad medio-alta, de 16 a 23 puntos; y en tercer lugar, nos encontramos los recursos de la escala baja (11 recursos). Por último, no se ha conseguido ningún recurso que haya obtenido una puntuación del 24 al 30 para situarse en la escala de fiabilidad alta. Debido a estos resultados, podemos confirmar, por tanto, que los recursos disponibles actualmente para los traductores de textos turísticos reciben una puntuación muy inferior a la recomendada y, por tanto, solo el 20% de todos recursos evaluados consiguen una puntuación apta para su recomendación (Gráfico 7).

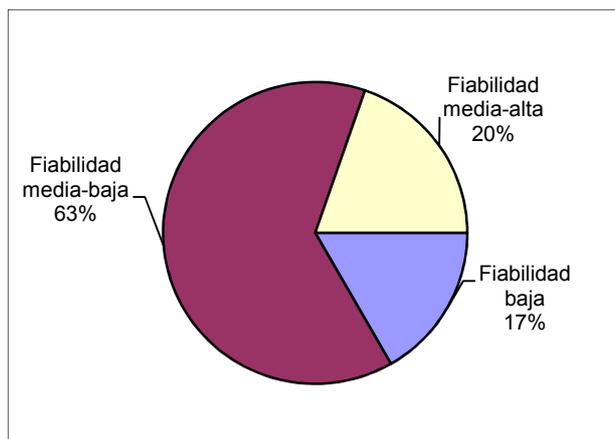


Gráfico 7. Escala de fiabilidad de los recursos evaluados en porcentajes.

En nuestra opinión, las deficiencias que presentan estos recursos y que, por tanto, nos obligan a puntuarlos con una puntuación inferior son debidas a dos factores principalmente: por un lado, estos recursos están elaborados con el objetivo de descodificar la información, es decir, de comprender únicamente el término desconocido en un momento dado y no de ayudar a la redacción; y por otro lado, los destinatarios a los que van dirigidos estos recursos son en su mayoría turistas y, en algunos casos, profesionales del turismo, por lo que no necesitan más información que el equivalente de traducción o la definición de un término desconocido. En estos casos, no son necesarias las indicaciones de uso, los ejemplos, la posible fraseología, sinónimos, etc., sino simplemente una ayuda para permitir la comprensión del mensaje en la lengua original. En cambio, cuando hablamos de traductores u otros profesionales de la lengua, esta información se vuelve imprescindible, ya que es esencial para realizar la codificación del mensaje con el uso adecuado y preciso de los términos.

Como veíamos en el capítulo anterior (cf. 2.5.3.), los traductores presentan unas necesidades y preferencias específicas a la hora de consultar recursos terminológicos, que son básicamente la inclusión de información lingüística y pragmática en los recursos, es decir, necesitan información sobre las características, el dominio, el contexto, el significado, etc. que presentan los términos y, por tanto, no pueden limitarse a la búsqueda del equivalente. Sin embargo, los resultados de la evaluación de los recursos nos muestran que ninguno ofrece la información requerida por los traductores, a excepción de la inclusión de definiciones de algunos recursos, ya que,

como decíamos anteriormente, su objetivo es simplemente de comprensión del término original, ya sea a través de definiciones en los recursos monolingües con o sin equivalencias o mediante equivalentes si son léxicos.

Por ello, es necesario crear recursos que tengan en cuenta a los traductores profesionales como destinatarios meta y que cubran sus necesidades a la hora de consultarlos, tanto con relación a la macroestructura como a la microestructura. Por todo ello, se confirma aún más la necesidad de nuestra investigación, a fin de ocupar un hueco vacío en este segmento turístico en general y, especialmente, para nuestro grupo de destinatarios.

Una vez finalizado el análisis del dominio del segmento de aventura en todos sus aspectos (lingüístico, traductológico, legislativo y terminológico), pasaremos a compilar el corpus de trabajo.

3.3.7. ESQUEMA CATEGORIAL DEL DOMINIO

Como vimos en el capítulo anterior (cf. 2.6.1.1.), el análisis del dominio de trabajo tiene como resultados: a) el conocimiento general y específico del terminógrafo sobre el dominio de trabajo y b) la conceptualización de dicho dominio. Concretamente, el punto b) es resultado también del primero, ya que, gracias al conocimiento que adquiere el terminógrafo sobre el dominio es capaz de realizar una conceptualización inicial del dominio de trabajo que, conforme vaya avanzando en el estudio de la terminología del dominio, podría sufrir cambios posteriores.

En nuestro caso, llegados a este punto consideramos que hemos adquirido la formación necesaria para plantear un esquema categorial inicial del segmento del turismo de aventura, que, como decíamos, podrá ser modificada y ampliada conforme vayamos avanzado en las fases del trabajo ontoterminográfico. Para ello, hemos utilizado los marcos semánticos, según la Semántica de Marcos (cf. 1.3.2.3.3.1.), a fin de organizar el dominio de trabajo basándonos en los conceptos básicos y las relaciones que establecen entre ellos. El resultado de esta organización se puede observar en la siguiente figura:

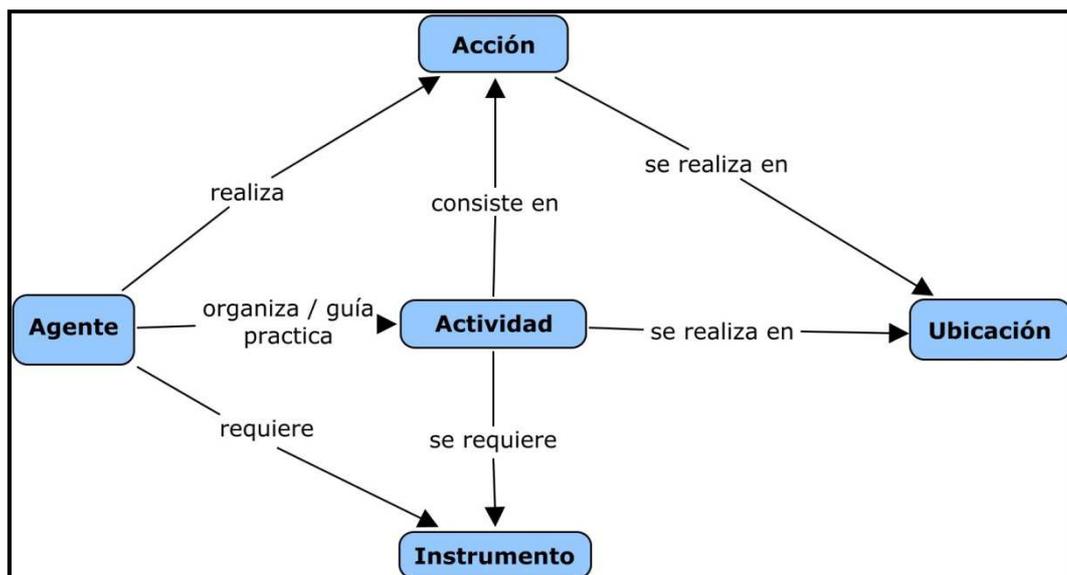


Figura 44. Marco semántico del segmento del turismo de aventura.

En esta representación, se definen las cinco categorías principales de este dominio que hemos detectado durante la fase de análisis, a saber: Agente, que hace referencia a la persona que realiza, organiza o dirige una actividad; Actividad, que se refiere a la práctica que realiza, guía o dirige un agente; Acción, que se dirige a la acción que se realiza en la actividad; Ubicación, que engloba el medio natural en el que se realiza una actividad, e Instrumento, que hace referencia a los objetos que se necesitan en una actividad. Además, se incluyen en el esquema las diferentes relaciones que vinculan las categorías semánticas entre sí. Mediante esta metodología, queda patente que, para entender y definir cualquier concepto dentro de este marco, se requiere la presencia de los otros cuatro conceptos que lo forman. De esta manera, un agente, es decir, una persona que realiza, guía o practica una actividad se definirá según la actividad y la acción que realice y el instrumento que utilice; para una actividad, será necesario explicitar la acción que se realiza, así como su ubicación y el instrumento que se utilice, y así sucesivamente. En definitiva, se concreta la posición que ocupa un concepto dentro de una situación comunicativa concreta y, por tanto, elimina, por un lado, la posible ambigüedad que pueda surgir en el nivel lingüístico con los términos del dominio de trabajo y, por otro lado, facilita la comprensión de las relaciones establecidas entre los diferentes conceptos que forman parte de dicho dominio.

Estas relaciones incluidas en el marco semántico representado aparecen como relaciones unidireccionales, es decir, relaciones que van de una categoría a otra y no a la inversa. No obstante, según nuestra metodología, cualquier relación que se represente debe contar con su inversa, ya que los vínculos establecidos entre categorías no funcionan de forma estanca, sino de forma dinámica y bidireccional, como por ejemplo en los dos modelos de relaciones siguientes que puede establecer la categoría <Instrumento>:

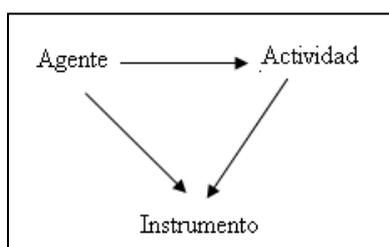


Figura 45. Modelo A.

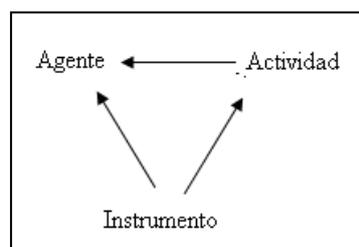


Figura 46. Modelo B.

En estos ejemplos, se observa cómo la categoría <Instrumento> está vinculado a las categorías <Agente> y <Actividad> de forma indirecta en el modelo A: <Agente> requiere <Instrumento> y <Actividad> requiere <Instrumento>, y de forma directa en el modelo B: <Instrumento> es requerido por <Agente> y <Instrumento> es requerido en <Actividad>. De esta manera, para cualquier relación que se estableciera entre conceptos automáticamente debería crearse su inversa, a fin de configurar una representación categorial de este dominio más completa y eficiente, lo que daría como resultado la siguiente conceptualización del dominio del turismo de aventura partiendo de las relaciones inversas (Figura 47):

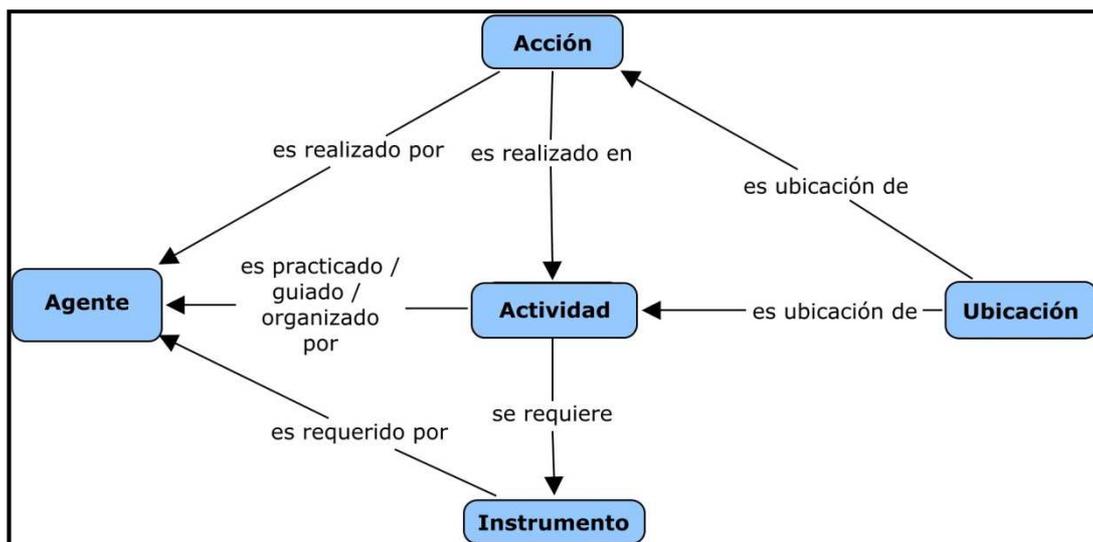


Figura 47. Marco semántico del segmento del turismo de aventura con relaciones inversas.

Esta representación inicial que incluye las principales categorías del dominio del turismo de aventura ayudará a la organización y adquisición de conocimiento de forma más adecuada en las siguientes fases y, por su parte, podrá ir modificándose o actualizándose según se vaya avanzando en el análisis del dominio. No obstante, recordemos la imposibilidad de establecer definitivamente y de forma estable una representación conceptual para algún dominio, puesto que estos avanzan constantemente creando, modificando o eliminando conceptos. De esta manera, debemos ser conscientes de que la estructura que creemos deberá ser lo más flexible posible para ajustarse al dinamismo del dominio en el que se trabaje.

3.4. COMPILACIÓN DEL CORPUS AVENTURIS

The beginning of any corpus study is the creation of the corpus itself. The decisions that are taken about what is to be in the corpus, and how the selection is to be organized, control almost everything that happens subsequently. The results are only as good as the corpus (Sinclair, 1991: 13).

Como veíamos en el segundo capítulo (cf. 2.6.1.2.), la siguiente fase a la delimitación del proyecto y análisis del dominio según nuestra propuesta metodológica

consiste en la compilación del corpus de trabajo que sirva de base para el posterior trabajo ontoterminográfico. Debido al objetivo principal de nuestro trabajo, que consiste en la realización de un trabajo ontoterminográfico, nuestro corpus especializado estará formado por dos subcorpus: por un lado, un subcorpus comparable, formado por textos originales en las lenguas de trabajo (inglés, alemán y español), que se ha venido a denominar AVENTURIS 1; y por otro, un subcorpus paralelo, constituido por componentes paralelos de textos originales y sus respectivas traducciones, que se ha denominado AVENTURIS 2.¹⁰¹⁵ En este apartado, trataremos la fase de compilación general de los dos subcorpus y, en las siguientes secciones, describiremos cada uno de forma independiente.

Esta fase de compilación se divide en tres etapas diferentes, a saber: el diseño del corpus, la búsqueda documental y la organización de los textos.

En la fase del diseño, es imprescindible tener en cuenta las variables lingüístico-pragmáticas determinadas en la fase anterior sobre la delimitación del trabajo, puesto que se trata de los puntos básicos para establecer el diseño de un corpus especializado:

- Lenguas de trabajo: español, inglés y alemán.
- Destinatarios: traductores profesionales.
- Finalidad: elaboración de un recurso ontoterminográfico multilingüe.
- Temática: segmento del turismo de aventura.

Asimismo, es necesario establecer los criterios externos que deben cumplir todos los textos que se seleccionen en la etapa de búsqueda documental, para garantizar su fiabilidad, su rigor y su cohesión interna, así como la representatividad del discurso especializado objeto de estudio. Para nuestra compilación, hemos considerado los siguientes criterios:

¹⁰¹⁵ El corpus AVENTURIS se inserta dentro del macrocorpus de turismo compilado en el marco del proyecto HUM-892: La contratación turística electrónica multilingüe como mediación intercultural: aspectos legales, traductológicos y terminológicos. URL: <<http://clg.wlv.ac.uk/malaga/>>.

- Criterio temporal: los textos que se seleccionen deberán ser lo más actuales posibles, a fin de mostrar la terminología actual del turismo de aventura.¹⁰¹⁶
- Criterio geográfico: los textos seleccionados están limitados geográficamente, puesto que el objetivo principal de este trabajo es asistir a traductores profesionales que traducen de una variedad concreta a otra. Por ello, los textos en español pertenecen a páginas web de España; los textos en inglés pertenecen a páginas de Gran Bretaña, y los textos en alemán, de Alemania.
- Criterio formal: con respecto a la variedad diastrática, los textos pertenecen todos a un nivel de abstracción medio-especializado y divulgativo, puesto que se basan en una situación comunicativa de especialista - no especialista. Y, con respecto a la variedad diafásica, los textos serán todos escritos, originales y completos, es decir, no se incluirán fragmentos de estos, y estarán publicados en páginas web.
- Tipología textual: los textos seleccionados pertenecen a la misma tipología textual, que en este caso se trata de textos promocionales publicados o elaborados en Internet por entidades públicas o empresas registradas que ofrecen actividades del turismo de aventura.

Una vez que hemos establecido todos los criterios iniciales para la compilación de nuestro corpus especializado, pasaremos a la búsqueda documental. Pero, antes de comenzar con esta búsqueda, debemos establecer específicamente las actividades que vamos a considerar como parte de este segmento, ya que, como veíamos en apartados anteriores, el turismo de aventura se caracteriza por su gran variedad de oferta de actividades. De esta manera, y después del análisis del dominio llevado a cabo y de la consulta de las diferentes clasificaciones elaboradas, presentaremos a continuación las actividades con las que trabajaremos en este dominio.

¹⁰¹⁶ Debido a la dinamicidad de este dominio, es importante que el proyecto tenga un tiempo limitado y que se preste atención a los posibles cambios sufridos por los términos.

| | |
|-----------------------------|-------------------------------------|
| 1. Paintball | 21. Windsurf |
| 2. Tiro con arco | 22. Kitesurf |
| 3. Excursiones con animales | 23. Boogiesurf |
| 4. Montañismo | 24. Esquí acuático |
| 5. Excursionismo a pie | 25. Esquí de río |
| 5.1. Travesía | 26. Moto de agua |
| 5.2. Trekking | 27. Bodyboard |
| 5.3. Sendensmo | 28. Wakeboard |
| 5.4. Nordic Walking | 29. Remo |
| 6. Escalada | 30. Salto con elástico |
| 7. Cicloturismo | 31. Puenting |
| 7.1. Bicicleta de montaña | 32. Parasailing |
| 7.2. Ciclismo en ruta | 33. Vuelo en globo |
| 9. Rápel | 34. Helie:ursión |
| 10. Tirolina | 35. Vuelo con ultraligero |
| 11. Espeleología | 36. Parapente |
| 12. Orientación | 37. Paracaidismo |
| 13. Piragüismo | 38. Ala delta |
| 13.1. Kayak | 39. Vuelo a vela |
| 13.1.1. Kayak de mar | 40. Paramotor |
| 13.1.2. Kayak de río | 41. Todoterreno |
| 13.2. Canoa | 42. Karting |
| 14. Rafting | 43. Buggy |
| 15. Hihospeed | 44. Motociclismo |
| 16. Hihobob | 45. Quad |
| 17. Barranquismo | 46. Motos de nieve |
| 18. Buceo | 47. Excursión con raquetas de nieve |
| 18.1. Buceo con bombona | 48. Mushing |
| 18.2. Buceo con snorkel | 49. Heliesqui |
| 19. Surf | 50. Teambuilding |
| 20. Paddle Surf | 51. Circuito multiaventura |

Tabla 37. Actividades del turismo de aventura.

A continuación, comenzaremos con la búsqueda documental para la compilación del corpus y, para ello, nos basaremos en los tres tipos de búsqueda que planteaba Autermühl (2001: 52 y ss.; cf. 2.6.1.2.2.), a saber: *búsqueda institucional*, que hace referencia a la búsqueda que se realiza en sitios web de instituciones u organizaciones internacionales, *búsqueda temática*, normalmente mediante directorios y, por último, *búsqueda por palabra clave*, a través de buscadores generales o específicos.

Con respecto al primer tipo de búsqueda, la *búsqueda institucional*, hemos utilizado como fuentes las páginas web oficiales de turismo de diferentes regiones, como por ejemplo, la *Web oficial de turismo en Andalucía*,¹⁰¹⁷ *The official website for English tourism*,¹⁰¹⁸ *Deutschland. Das Reiseland*;¹⁰¹⁹ páginas web de agencias de viajes

¹⁰¹⁷ <<http://www.andalucia.org/>>

¹⁰¹⁸ <<http://www.visitengland.com/>>

y tour operadores, como *Halcón Viajes*,¹⁰²⁰ *MuchoViaje.com*,¹⁰²¹ *Lastminute.com*¹⁰²² o *Destinia.com*.¹⁰²³

Las *búsquedas temáticas* se han realizado a través de los numerosos directorios de turismo, concretamente de turismo activo o de aventura que existen en la actualidad. Como ejemplo, podemos destacar el directorio de *VirtualTourist.com (Europe Travel Guide)*,¹⁰²⁴ la guía de turismo de aventura y actividades en Andalucía de la *Revista Ibérica*¹⁰²⁵ o el directorio de *TurismoActivo.com*.¹⁰²⁶

Por último, se han llevado a cabo también *búsquedas por palabras clave*, a fin de encontrar empresas de turismo activo o de aventura que ofrecían actividades de nuestro interés. Para ello, se han utilizado las mismas técnicas y palabras clave que se utilizaron en el segundo capítulo para la búsqueda de los recursos terminológicos (cf. Tabla 21) combinadas con palabras como «empresa», «asociación», etc. A pesar del ruido obtenido, se encontraron numerosas páginas web interesantes para la extracción de textos para la compilación del corpus textual. Algunas de ellas están recogidas en la siguiente tabla (Tabla 38):

¹⁰¹⁹ <<http://www.deutschland-tourismus.de/>>

¹⁰²⁰ <<http://www.halconviajes.com/>>

¹⁰²¹ <<http://www.muchoviaje.com/>>

¹⁰²² <<http://www.lastminute.com/>>

¹⁰²³ <<http://destinia.com/>>

¹⁰²⁴ <<http://www.virtualtourist.com/travel/Europe/TravelGuide-Europe.html>>

¹⁰²⁵ <<http://www.revistaiberica.com/directorios/aventura/andalucia.htm>>

¹⁰²⁶ <<http://www.turismoactivo.com/>>

| LENGUA | URL |
|---------|--|
| ESPAÑOL | http://www.pangeacentral.com/ http://www.pasoslargos.com/ http://www.laserrania.org/ http://www.andarines.com/ http://www.skydivespain.com/ http://www.rutacero.com/ http://www.ociosport.net/ http://www.reservatuaventura.com/ http://www.pangeacentral.com/ http://www.naturexplorer.com/ |
| INGLÉS | http://www.absoluteadventure.co.uk/ http://www.bachtravel.com http://www.visitwales.co.uk/ http://www.blackmountain.co.uk/ http://www.canoehirewales.co.uk/ http://www.callofthewild.co.uk/ http://www.wildernessscotland.com/ http://www.mountainbikescotland.com/ http://www.interactivities-uk.com/ http://www.contours.co.uk/ |
| ALEMÁN | http://www.wasungen.de/ http://www.mum-werra-kanu-tours.de/ http://www.kringeltours.de/ http://www.eisenach-tours.de/ http://www.krumos-kanus.de/ http://www.leineradweg.de/ http://www.geraradweg.de/ http://www.rennsteig.net http://www.tourist-schotten.de/ http://www.eder-kanu.de/ |

Tabla 38. Ejemplos de páginas web obtenidas con la búsqueda de palabras clave para el segmento de turismo de aventura.

Una vez que hemos seleccionado las fuentes de información, pasamos a homogeneizar y codificar los documentos, a fin de crear un corpus especializado ordenado y fácil de consultar. Para ello, deberemos realizar dos pasos diferentes: por un lado, registramos los datos de todos los documentos seleccionados en una tabla (Tabla 39) y le proporcionamos un código identificativo y único a cada documento, que se registrará también en la tabla junto con el resto de datos; y por otro lado, homogeneizamos el formato de todos los textos para su posterior gestión y, así, convertimos los diferentes formatos que encontramos en la Red (.pdf, .html, .rtf, etc.) en formato ASCII o de texto plano (con extensión .txt). De esta manera, conseguimos que todos los documentos que forman nuestro corpus especializado se encuentren en el mismo formato y se localicen de forma inequívoca. Asimismo, otra tarea que se debe

llevar a cabo una vez que se han registrado los textos y se han convertido a un formato plano, se deben limpiar de caracteres erróneos, espacios dobles, saltos de línea, etc. para evitar cualquier tipo de confusión y para permitir una consulta clara de los textos.

| ID | Título | Url | Género textual | Idioma | Traducción | Región | Autor |
|----------------------|-------------------------------|---|----------------|--------|------------|-----------|----------------------|
| 0001 TOES 1027 | Multiaventura en las Zorreras | http://www.pangeacentral.com/multi_zorreras.html | Promoción | ES | No | Andalucía | Pangea Active Nature |

Tabla 39. Ejemplo de tabla de registro.

Una vez que hemos presentado las tres fases de compilación de nuestro corpus, pasaremos a la descripción de los dos subcorpus, con objeto de concretar los aspectos específicos de cada subcorpus.

3.4.1. EL SUBCORPUS COMPARABLE AVENTURIS 1

El subcorpus AVENTURIS 1 se trata de un subcorpus comparable multilingüe formado por tres componentes de textos originales escritos en español, inglés y alemán. Como el corpus al que pertenece, cumple los criterios externos que se expusieron en el apartado anterior, a saber: criterio temporal, criterio geográfico, criterio formal y tipología textual. Así pues, se trata de un subcorpus virtual, formado por documentos íntegros en formato electrónico, de tipología promocional, que versan sobre el dominio de aventura, presentan una situación comunicativa de especialista - no especialista y una delimitación temporal y geográfica específica. Con estos criterios se pretende acotar la

¹⁰²⁷ El protocolo de codificación que se ha utilizado para este corpus ha consistido en un código alfanumérico, cuyas partes se indican a continuación: «0001» hace referencia al registro del texto que corresponde en cada momento (en el caso del ejemplo, sería el primer texto compilado) y que, por tanto, irá cambiando según se vaya registrando los diferentes textos que forman el corpus; «TO» indica que el texto compilado es un texto original y, por el contrario, «TM» indicaría que se trata de una traducción; por último, «ES» hace referencia al idioma del texto que, según el caso, debería seleccionarse entre ES para español, DE para alemán o EN para inglés.

búsqueda de documentos y compilar, así, un corpus lo más representativo posible del dominio objeto de estudio.

Como vimos en capítulos anteriores, el tamaño y la representatividad que debe tener un corpus se considera un tema polémico, debido a las diferentes opiniones que existen en torno a él (cf. 1.4.2.2.1 y 2.6.1.2.2.). Actualmente, es posible medir de forma cuantitativa y, por tanto, objetiva la representatividad de un corpus a posteriori mediante la aplicación *ReCor* (Seghiri Domínguez, 2006; Corpas Pastor y Seghiri Domínguez, 2006, 2007a, 2007b, 2011/en prensa; cf. 2.6.1.2.2.), un programa que realiza el análisis de la densidad léxica de un corpus compilado mediante el algoritmo N-Cor.¹⁰²⁸

En nuestra investigación, hemos utilizado este programa para medir la representatividad de nuestro corpus, que presenta las siguientes dimensiones:

| IDIOMA | N.º DOCUMENTOS | N.º PALABRAS |
|---------|----------------|--------------|
| ESPAÑOL | 455 | 252.878 |
| INGLÉS | 634 | 252.199 |
| ALEMÁN | 646 | 255.805 |

Tabla 40. Recuento del subcorpus AVENTURIS 1.

Y a continuación mostramos los resultados arrojados después de la aplicación del programa a los tres componentes que forman el subcorpus AVENTURIS 1:

¹⁰²⁸ Este método se ha visto implementado en la aplicación *ReCor*, que se encontrará disponible en breve en la dirección URL: <http://www.recorweb.com>.

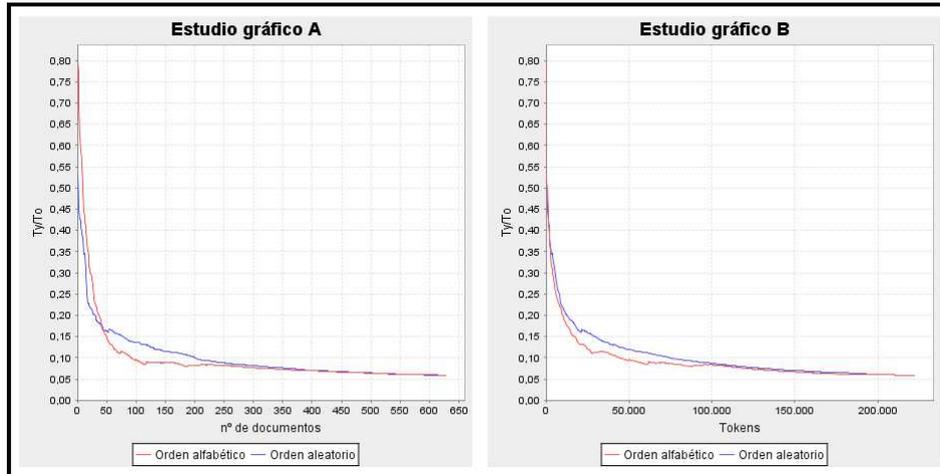


Gráfico 8. Resultados de *ReCor* para el corpus comparable de inglés.

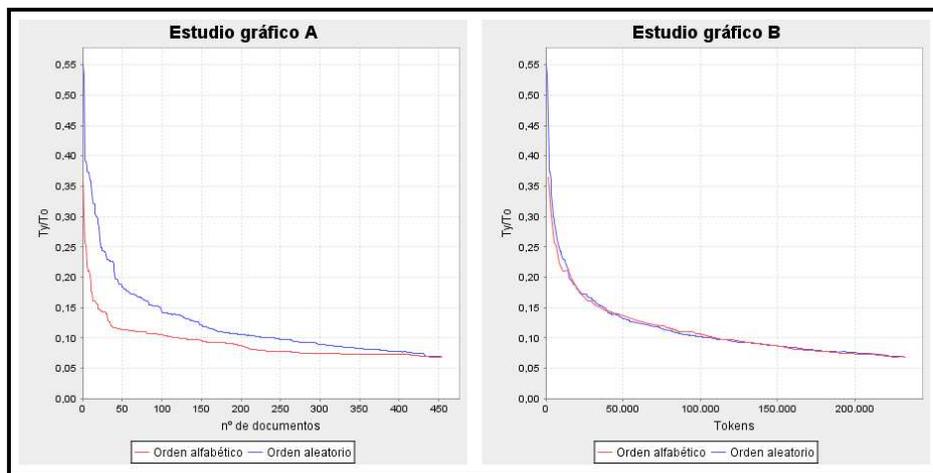


Gráfico 9. Resultados de *ReCor* para el corpus comparable de español.

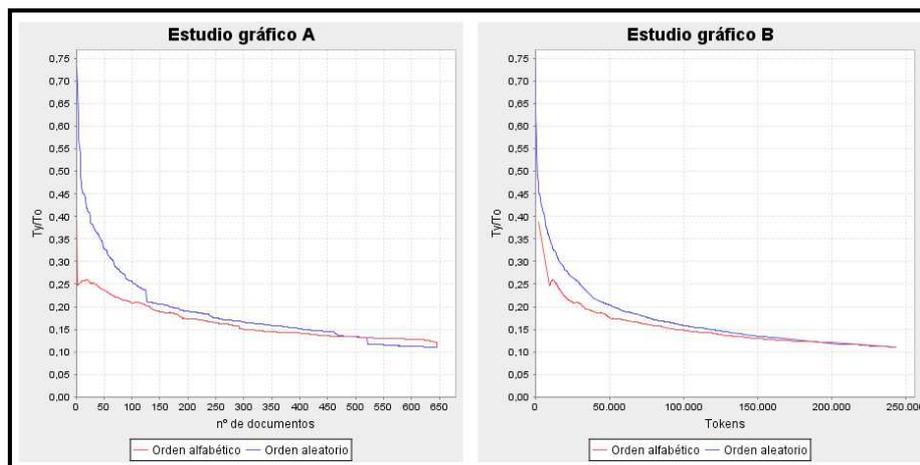


Gráfico 10. Resultados de *ReCor* para el corpus comparable de alemán.

En el primer gráfico, que muestra los resultados del componente inglés, tanto la hipérbola de color rojo, que representa el orden alfabético de aparición de las palabras (*tokens*) en el texto, como la hipérbola de color azul, que representa estos *tokens* con un orden aleatorio, se estabiliza en el valor 0,05 para los 634 textos (primera gráfica) y 252.199 *tokens* (segunda gráfica) que forman este componente. De esta manera, estas gráficas nos indican que, después de haber analizado el número de palabras que forman el componente inglés (252.199), existe una probabilidad de 0,05 de que aparezcan más palabras nuevas.

En cuanto a los otros dos gráficos, referentes a los componentes español y alemán, las curvas nos indican que existe una probabilidad de 0,07 aproximadamente de que en el componente español aparezcan palabras nuevas, y de 0,10 aproximadamente de que lo hagan en el componente alemán. Así pues, observamos que la representatividad es mayor en el componente inglés que en los otros dos componentes a pesar de contar con el mismo número de palabras.

Como conclusiones de estas gráficas, podemos extraer las siguientes:

- Existe cierta estabilidad en las hipérbolas de los tres componentes en valores próximos a cero, aunque los resultados demuestran que el subcorpus no alcanza la representatividad absoluta (cero o inferior a 0,05).
- El programa arroja resultados muy favorables en lenguajes muy acotados y restringidos como es el lenguaje jurídico (cf. Seghiri Domínguez, 2006), por su elevada densidad léxica y su nivel de repetición de estructuras y expresiones, pero en lenguajes más flexibles y con mayor variedad interna, como es el lenguaje utilizado en el segmento del turismo de aventura, no es posible medir su representatividad de forma tan eficiente.
- La variedad de actividades que se ofrecen en el turismo de aventura, unido a la flexibilidad y apertura de su lenguaje y al nivel menos especializado de la tipología textual seleccionada (textos promocionales) no permiten establecer un corpus representativo al 100%.

- El aumento de textos en los diferentes componentes mantiene la representatividad obtenida con las dimensiones establecidas del corpus e incluso disminuye la representatividad.¹⁰²⁹
- El componente alemán ofrece mayor variabilidad en cuanto a la terminología de este segmento, ya que sus términos aparecen en unas ocasiones como unidades monoléxicas y en otras como unidades poliléxicas en combinaciones con unidades diferentes. Es decir, el carácter aglutinador de la lengua alemana obstaculiza la consecución de la representatividad del componente.

Debido a la imposibilidad de medir la representatividad del subcorpus AVENTURIS 1 de forma cuantitativa con el programa *ReCor*, es necesario basarnos en los criterios externos establecidos en el apartado anterior y a comienzos de este, a fin de alcanzar la representatividad del dominio del turismo de aventura. Además de estos criterios, hemos pretendido compilar un corpus equilibrado en cuanto al número de palabras, aunque no en cuanto al número de documentos (Tabla 39), debido a varias razones: por un lado, los textos incluidos en el corpus español presentan por lo general una mayor extensión que los textos de las otras dos lenguas; y por otro, el español suele utilizar un mayor número de palabras que las otras dos lenguas por sus características lingüísticas y léxicas.

Además de mantener el equilibrio del número de palabras, hemos considerado relevante también mantener el equilibrio conceptual, es decir, hemos tenido en cuenta las diferentes actividades que se practican dentro del segmento de turismo de aventura (Tabla 40) y hemos incluido un número aproximado de textos para todas ellas, con objeto de ofrecer una homogeneización en la terminología extraída. No obstante, es necesario tener en cuenta a este respecto que no todas las actividades tienen la misma oferta y demanda, por lo que para algunas no ha sido posible obtener un número similar de textos, como ha sido el caso de la actividad de hidrobob, que se considera como actividad dentro de este sector.

¹⁰²⁹ Se realizaron varias pruebas de incremento de textos con los componentes español y alemán y los resultados obtenidos mostraron una disminución en su representatividad.

Como conclusión, podemos confirmar que el subcorpus AVENTURIS 1 se considera un corpus representativo según criterios cualitativos, ya que cumple todos los criterios externos establecidos en cuanto a la compilación de los documentos que forman parte de los tres componentes, así como por lo que respecta al equilibrio conceptual y número de palabras contenidas.

3.4.2. EL SUBCORPUS PARARELO AVENTURIS 2

El subcorpus AVENTURIS 2 se trata de un subcorpus paralelo multilingüe, más específicamente trilingüe (español, inglés y alemán) formado por tres componentes: un componente constituido con textos escritos originariamente en español y otros dos componentes formados por sus versiones traducidas al inglés y al alemán. Asimismo, al igual que ocurría con el otro subcorpus, se trata de un subcorpus virtual, formado por documentos íntegros en formato electrónico, de tipología promocional, que versan sobre el dominio de aventura y presentan una situación comunicativa de especialista - no especialista.

Un gran inconveniente que se ha encontrado a la hora de compilar este subcorpus y que nos ha obligado a limitar sus dimensiones ha sido la carencia de páginas web que ofrecían traducciones de su contenido en estos dos idiomas (alemán y español). La gran mayoría de las páginas solo ofrecen su contenido en español y, como bien confirma Mezyed Rivas (2005: 25) en su informe sobre el turismo rural y activo en España, esto provoca una falta de atención a la demanda de los principales mercados emisores de turistas. Por su parte, en caso de disponer de alguna versión traducida, esta era siempre en inglés y, ocasionalmente, también en alemán. Por este motivo, nuestro subcorpus paralelo presenta unas dimensiones muy limitadas, que se presentan a continuación:

| IDIOMA | N.º DOCUMENTOS | N.º PALABRAS |
|---------|----------------|--------------|
| ESPAÑOL | 40 | 58.643 |
| INGLÉS | 40 | 51.573 |
| ALEMÁN | 40 | 51.076 |

Tabla 41. Recuento del subcorpus AVENTURIS 2.

De esta manera, se concluye también que la representatividad y adecuación de este subcorpus no se puede alcanzar mediante medios cuantitativos, sino, de nuevo, mediante criterios cualitativos, es decir, a través de los criterios externos establecidos con anterioridad.

Llegados a este punto, ya habríamos finalizado las dos primeras fases de trabajo según la metodología propuesta y, por tanto, pasaríamos a la gestión ontoterminográfica y la elaboración propia de la base de datos, cuyo desarrollo se presenta en el siguiente capítulo.

CAPÍTULO 4.

REALIZACIÓN DEL TRABAJO ONTOTERMINOGRÁFICO Y SU EDICIÓN

ÍNDICE DE CONTENIDOS

- 4.1. Introducción
- 4.2. Realización del trabajo ontoterminográfico con *ProTermino*
 - 4.2.1. Gestión del trabajo ontoterminográfico
 - 4.2.1.1. Aspectos previos
 - 4.2.1.2. Extracción de candidatos a término
 - 4.2.1.2.1. Validación y eliminación de candidatos a término
 - 4.2.1.3. Extracción de relaciones semánticas
 - 4.2.1.4. Elaboración de la ontología de dominio
 - 4.2.2. Elaboración de la base de datos ontoterminográfica
 - 4.2.2.1. Campos terminológicos
 - 4.2.2.1.1. Campo *Definition*
 - 4.2.2.1.2. Campo *Equivalence*
 - 4.2.2.1.3. Campo *Context* y *Collocations*. Módulo de concordancias
 - 4.2.2.2. Resultado de entrada terminológica y validación
- 4.3. Exportación y edición del trabajo ontoterminográfico
 - 4.3.1. *Trandix*: herramienta para la búsqueda terminológica en contexto
 - 4.3.1.1. Antecedentes
 - 4.3.1.1.1. *Locolex* y *COMPASS*
 - 4.3.1.1.2. *MobiMouse Plus*
 - 4.3.1.1.3. *Smarty*
 - 4.3.1.1.4. *Evroterm* y *TERMINATOR*
 - 4.3.1.2. Objetivos
 - 4.3.1.3. Descripción
 - 4.3.1.3.1. Descripción de funcionalidades
 - 4.3.1.3.1.1. Formato TBX (Term Base eXchange)
 - 4.3.1.3.1.2. Vinculación entre archivos
 - 4.3.1.3.2. Comparativa entre *Trandix* y sus antecedentes
 - 4.3.1.4. Evaluación de *Trandix*
 - 4.3.1.4.1. Fase preparatoria
 - 4.3.1.4.1.1. Finalidad del experimento
 - 4.3.1.4.1.2. Planificación del experimento
 - 4.3.1.4.1.3. Muestra del experimento
 - 4.3.1.4.1.4. Material para el experimento
 - 4.3.1.4.1.5. Preparación del entorno web
 - 4.3.1.4.2. Fase de ejecución
 - 4.3.1.4.3. Fase de evaluación
 - 4.3.1.4.3.1. Evaluación de traducciones y conceptos relacionados
 - 4.3.1.4.3.2. Métodos de evaluación
 - 4.3.1.4.3.3. Propuesta de plantilla mixta de evaluación
 - 4.3.1.4.3.4. Resultados del experimento
 - 4.3.1.4.3.4.1. Criterio de finalización de la tarea
 - 4.3.1.4.3.4.2. Criterio de tiempo de realización
 - 4.3.1.4.3.4.3. Criterio de calidad
 - 4.3.1.4.3.4.4. Criterio de satisfacción
 - 4.3.1.4.3.4. Conclusiones del experimento y trabajo futuro con *Trandix*

Terminology is not only a matter of terms and term entries that endeavour, no matter how imperfectly, to represent slices of objective reality. In this sense, the representation of a specialized field should be more than a list or even a configuration of objects linguistically translated into either simple or compound nominal forms. It is necessary to situate concepts in a particular setting and within the context of dynamic processes that define and describe the principal event in the specialized field in question (Faber Benítez et ál., 2005).

4.1. INTRODUCCIÓN

Una vez que hemos finalizado la fase de preparación del trabajo ontoterminográfico para el recurso terminológico sobre el turismo de aventura, que presentábamos en el capítulo anterior, seguiremos con las siguientes fases de nuestra propuesta metodológica. Nos parece oportuno destacar que en la fase anterior hemos alcanzado tres objetivos iniciales de nuestra metodología, a saber: en primer lugar, hemos concretado las variables pragmático-lingüísticas del proyecto (destinatarios, lenguas, etc.), así como los recursos tecnológicos y humanos necesarios para llevarlo a cabo; en segundo lugar, hemos realizado un amplio análisis del dominio de trabajo elegido en todos sus niveles (desde el nivel lingüístico y traductológico hasta el estudio de los recursos terminológicos disponibles), con objeto de aproximarnos a él para obtener conocimientos amplios sobre sus características y necesidades propias, así como

para realizar una conceptualización del dominio lo más concreta posible; y, por último, hemos compilado un corpus de trabajo formado por un subcorpus comparable, AVENTURIS 1, y un subcorpus paralelo, AVENTURIS 2, que servirán de base para el trabajo de las siguientes fases.

En las siguientes fases, llevaremos a cabo tanto la gestión ontoterminográfica como la elaboración de la base de datos. Ambas fases engloban el trabajo ontoterminográfico *per se*, ya que en ellas se realiza el verdadero trabajo ontoterminográfico necesario para llevar a cabo un proyecto de estas características, que comienza con la extracción de candidatos a término y finaliza con la introducción de los datos en los campos terminológicos pertinentes. Estas fases dependen directamente de la calidad del trabajo que se haya realizado en las fases previas, puesto que se pone en práctica todo el conocimiento adquirido durante la fase de análisis del dominio así como el corpus de trabajo elaborado según las decisiones iniciales tomadas acerca de las variables pragmático-lingüísticas y los criterios generales y específicos para la compilación de corpus.

Después de haber llevado a cabo estas fases de gestión y elaboración de la base de datos ontoterminográfica, comienza la fase de revisión y validación de la información introducida. Este proceso puede parecer breve, puesto que consiste simplemente en la revisión y validación del trabajo ya realizado, pero es sumamente importante, ya que se trata del paso final antes de comenzar la edición de la base de datos y, por tanto, es la última fase que nos permite detectar cualquier error cometido anteriormente. Cabe recordar aquí que uno de los parámetros que medíamos en la ficha de evaluación de los recursos terminológicos del capítulo dos consistía exactamente en la existencia o no de errores en dichos recursos, es decir, se puntuaba positiva o negativamente un recurso según tuviera o no errores de cualquier naturaleza. Por este motivo, hemos tenido especial cuidado a la hora de llevar a cabo dicha revisión.

Por último, llegamos a la fase de edición de nuestro trabajo. En nuestra opinión, esta última fase se descuida en ocasiones por parte de los terminólogos, que exportan directamente el trabajo realizado en formato de texto (formato PDF, RTF, etc.) o formato electrónico (formato HTML) para su posterior publicación (en papel o formato electrónico) sin prestarle mayor atención al formato, al acceso a la información, a la

ergonomía del recurso, etc. De esta manera, se observa cómo el trabajo tan minucioso realizado hasta ese momento en el que se han utilizado herramientas de extracción, editores de ontologías, etc., es decir, herramientas que agilizan y mejoran el proceso de trabajo, finaliza con una edición tradicional que no incorpora ningún tipo de innovación ni mejora de las búsquedas tradicionales basadas en búsquedas directas de términos o búsquedas alfabéticas. Por este motivo, presentamos en esta última fase una herramienta diseñada dentro de esta tesis doctoral, *Trandix*, que pretende incorporar innovación a esta fase de edición y, por tanto, a las siguientes fases de búsqueda y consulta de información terminológica por parte de sus usuarios. Esta aplicación informática ofrece la posibilidad de importar archivos en formato TBX, el formato estándar para el intercambio de terminología propuesto por los organismos internacionales ISO y LISA, y realizar búsquedas de términos dentro de contexto, es decir, dentro del mismo texto en el que se encuentran. Asimismo, con objeto de ofrecer dinamismo y flexibilidad a los usuarios, *Trandix* dispone de diferentes tipos de búsquedas además de la búsqueda basada en contextos y permite enviar *feedback* al finalizar cada sesión de uso de la aplicación. En definitiva, se trata de una herramienta de búsqueda y consulta muy completa e innovadora que persigue la satisfacción de sus usuarios y de sus necesidades, lo cual queda patente gracias a los resultados obtenidos en el experimento realizado a alumnos de la Licenciatura de Traducción e Interpretación, de la Universidad de Málaga.

El trabajo ontoterminográfico que se presenta en este último capítulo, sumado a las fases explicadas anteriormente, nos permite confirmar la adecuación de la metodología propuesta en esta investigación, ya que se trata de una metodología accesible, completa y coherente que ayuda a alcanzar resultados satisfactorios y acordes con los objetivos iniciales del proyecto en cuestión. Asimismo, la herramienta *Trandix* nos muestra la importancia que tiene el proceso de edición y su posterior proceso de búsqueda y consulta para la satisfacción de los usuarios, es decir, no es suficiente con elaborar una buena base de datos completa y coherente, sino que es necesario trabajar también en formas de búsqueda y consulta innovadoras que se ajusten a las necesidades de los usuarios y que, a la vez, agilice su trabajo. Por último, cabe destacar que esta herramienta innovadora abre una nueva línea de trabajo para seguir trabajando en diferentes formas de acceso y búsqueda de información.

4.2. REALIZACIÓN DEL TRABAJO ONTOTERMINOGRÁFICO

Las fases del trabajo ontoterminográfico realizadas hasta el momento han permitido llegar a este punto en el que, una vez compilada toda la información necesaria tanto textual en forma de corpus como sobre los diferentes aspectos relativos al dominio seleccionado, pasaremos a la fase de gestión ontoterminográfica y a la fase de elaboración de la base de datos.

Como vimos en el segundo capítulo, la fase de gestión de nuestra propuesta metodológica está dividida en varias subfases, a saber: 1. Extracción de candidatos a término y su validación; 2. Búsqueda de equivalentes de traducción; 3. Extracción de relaciones semánticas, y 4. Elaboración de la ontología de dominio (cf. 2.6.1.3.). Por su parte, la fase de elaboración de la base de datos consta de tres subfases diferentes: 1. Selección de campos; 2. Elaboración de definiciones, y 3. Selección de ejemplos contextuales (cf. 2.6.1.4.). A pesar de haber establecido un orden de presentación de estas fases, no es necesario que se siga esta forma cronológica, ya que, dependiendo de las herramientas con las que se cuente a la hora de llevar a cabo el trabajo ontoterminográfico o de las preferencias del terminógrafo, estas fases se podrán realizar en este orden o, por el contrario, de forma intercalada o incluso simultánea. En definitiva, el terminógrafo es quien deberá establecer el orden que sigue para llevar a cabo estas dos fases de trabajo ontoterminográfico.

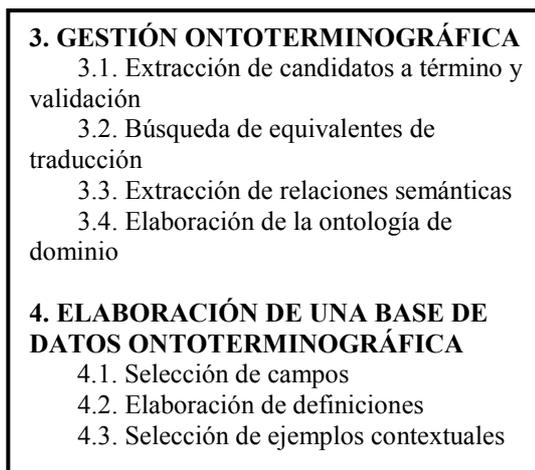


Figura 48. Fases de realización de la base de datos ontoterminográfica.

En nuestro caso, las fases han venido marcadas por la utilización de la herramienta de trabajo *ProTermino*, que presentamos en el capítulo anterior. Esta aplicación informática sigue un proceso de trabajo muy similar al presentado anteriormente, aunque permite la realización de algunas subfases de forma simultánea. Concretamente, las subfases de búsqueda de equivalentes y de elaboración de ontologías de dominio, como veremos más adelante.

4.2.1. LA GESTIÓN ONTOTERMINOGRÁFICA Y LA ELABORACIÓN DE BASE DE DATOS CON PROTERMINO

A continuación, presentaremos paso a paso el trabajo ontoterminográfico realizado con esta herramienta y prestaremos atención a algunos aspectos relevantes relativos al dominio del turismo de aventura.

4.2.1.1. Aspectos previos

Antes de comenzar el trabajo propiamente ontoterminográfico, el usuario deberá llevar a cabo dos tareas previas: por un lado, la creación de un proyecto de trabajo; y, por otro, la subida del corpus textual de trabajo.

Con respecto a la primera tarea indicada, la creación de un proyecto de trabajo, es necesaria puesto que esta herramienta da la posibilidad de crear diferentes proyectos de trabajo de forma simultánea, además de permitir el acceso remoto a múltiples usuarios. De esta manera, al permitir la simultaneidad de trabajos y de usuarios, la herramienta requiere de la identificación de los proyectos en curso para evitar posibles confusiones y proporcionar facilidad en las fases siguientes. La creación de estos proyectos es muy sencilla, y se trata del primer paso que se encuentra en la herramienta al acceder a ella. La ventana que encontramos para crear el nuevo proyecto se muestra en la siguiente figura:

Figura 49. Ventana de creación de proyectos en *ProTermino*.

En este paso, el usuario deberá facilitar información básica sobre el proyecto que va a crear, a saber: el nombre y una descripción breve del proyecto, a fin de identificarlo y evitar posibles confusiones en el futuro. En nuestro caso, se ha creado el proyecto *EcoTerm* y se han introducido una descripción breve de este: *Adventure Tourism Multilingual Resource*.

Figura 50. Creación del proyecto *EcoTerm*.

Una vez que se ha creado el proyecto por el usuario, este se añadirá a la tabla de proyectos que se han creado hasta el momento y para los que el usuario tiene el acceso permitido.

According to our record, you have access to the following projects:

| Project Name | Description | Delete? |
|--------------|---|---------------------------------------|
| test_1 | Elementary chemistry for beginners | <input type="checkbox"/> |
| test 7z | test 7z | <input type="checkbox"/> |
| EcoTerm | Adventure Tourism Multilingual Resource | <input type="checkbox"/> |
| Project Name | Description | <input type="button" value="Delete"/> |

Figura 51. Tabla de proyectos creados en *ProTermino*.

Además de crear proyectos, el usuario también tiene la posibilidad de borrar el proyecto que desee seleccionándolo y pulsando el botón de eliminación *Delete*, situado al principio y al final de la tabla. Para esto, el usuario debe tener en cuenta que, una vez eliminado el proyecto seleccionado, toda la información que contenga (documentos, términos extraídos, entradas terminológicas completas, etc.) también se eliminarán.

El otro paso previo necesario antes de empezar a trabajar con la herramienta en nuestro proyecto ontoterminográfico consistirá en introducir los documentos de trabajo que servirán de base para el trabajo posterior. Para llevar a cabo este paso, es necesario seleccionar el proyecto creado en el paso anterior con el que se vaya a trabajar y acceder a él. Una vez dentro, se debe seleccionar la pestaña de corpus en la pantalla inicial, desde donde la herramienta permitirá la subida de los documentos necesarios.

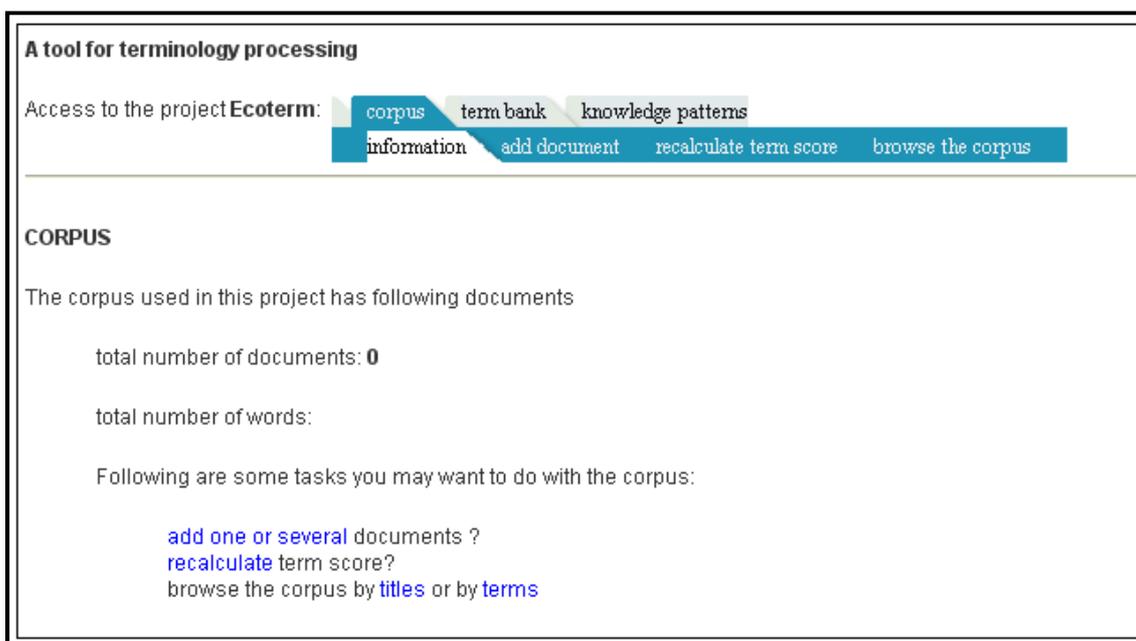


Figura 52. Subida de documentos en *ProTermino*.

En la pantalla del módulo de gestor de corpus (*corpus*), encontramos la opción *add one or several documents*, que permite al usuario subir documentos de forma independiente, es decir, subir un único documento en formato de texto plano (.txt), así como la posibilidad de subir una carpeta comprimida en diferentes formatos (.zip, .7zip) que contenga varios archivos en formato de texto plano. Esta última opción facilita y agiliza la subida de documentos, ya que evita que el usuario tenga que realizar una

subida manual de todos los archivos que componen el corpus de trabajo. Junto con la subida de archivos, es necesario especificar la lengua del texto o de los textos que se van a incorporar a la aplicación. De esta manera, en caso de subir un archivo comprimido compuesto por textos .txt se deberá especificar la lengua del archivo, por lo que todos los textos deberán estar escritos en la misma lengua.

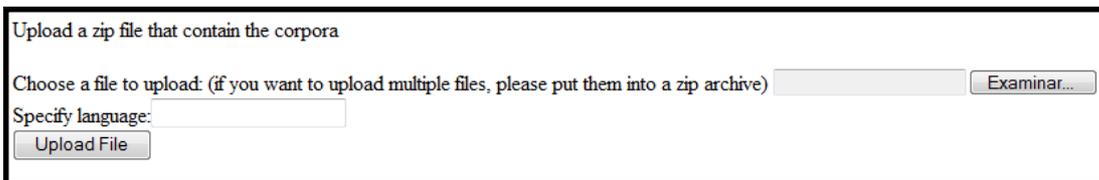


Figura 53. Ventana para añadir uno o más documentos.

En nuestro caso, hemos utilizado la opción de subida de archivos comprimidos y, para ello, hemos realizado tres subidas diferentes correspondientes a las tres lenguas de trabajo. Una vez incorporados estos tres archivos, la aplicación descomprimirá de forma automática los documentos y nos ofrecerá un recuento del total de documentos subidos y del número de palabras (*tokens*) que contienen dichos documentos. Teniendo en cuenta los tres corpus que hemos compilados, nuestro proyecto cuenta con 1.735 documentos y 753.316 palabras en total, que a pesar de no contar con una representación cuantitativa sí que lo hará mediante una representación cualitativa de las características intrínsecas y extrínsecas de los documentos que conforman el corpus de trabajo.

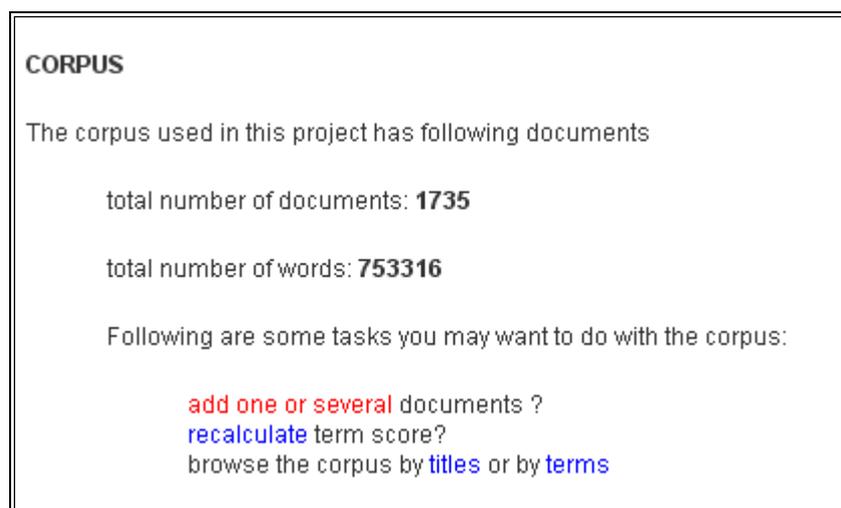


Figura 54. Recuento de documentos y de palabras del proyecto *EcoTerm*.

4.2.1.2. Extracción de candidatos a término

La posibilidad de realizar la extracción terminológica dentro de la misma aplicación facilita el trabajo del terminógrafo en gran medida y agiliza las tareas de las fases posteriores, lo que le confiere gran ventaja frente al resto de aplicaciones similares que existen actualmente y que no ofrecen esta opción.

Esta herramienta extrae de forma totalmente automática candidatos a término monolexemáticos y plurilexemáticos⁵⁹⁵ en las lenguas de trabajo, que recordemos pueden ser español, alemán, francés, inglés e italiano. La extracción que se lleva a cabo parte de corpus comparables, es decir, corpus que contienen únicamente textos originales y no traducciones, mediante el uso de una técnica híbrida, la cual combina una técnica de enfoque estadístico y una técnica de enfoque lingüístico (cf. Mitkov et ál., 2007). Por un lado, la técnica de enfoque estadístico que emplea es la TF.IDF (del inglés, *term frequency-inverse document frequency*),⁵⁹⁶ una técnica que permite evaluar la importancia de un candidato a término mediante su frecuencia con relación a un corpus de trabajo en el que aparece dicha unidad. La frecuencia de un término (TF) se corresponde con el número de apariciones de este término en un documento del corpus de trabajo; y, por su parte, la frecuencia inversa de documento (ITF) filtra y descarta las unidades con mayor número de repeticiones en los documentos, que suelen ser artículos, preposiciones, conjunciones, etc., para que no se tengan en cuenta durante la extracción y otorga mayor peso a las unidades con menos repeticiones. De esta manera, cuanto mayor sea el número de apariciones de una unidad, menor será la puntuación obtenida con esta medida estadística. Por otro lado, la técnica de enfoque lingüístico, que se debe llevar a cabo de forma previa a la anterior, consiste en el etiquetado gramatical de los documentos del corpus de trabajo mediante el uso del etiquetador morfológico de libre distribución *TreeTagger*.⁵⁹⁷ Así pues, las palabras del corpus

⁵⁹⁵ Por términos monolexemáticos y plurilexemáticos entendemos aquí los conceptos que se encuentran formados por una única unidad o por más de una unidad, respectivamente.

⁵⁹⁶ Para profundizar en esta técnica, consúltese Ha (2007: 69 y ss.), donde se recoge información exhaustiva sobre el fundamento teórico y aplicabilidad de esta técnica, junto con otras técnicas estadísticas.

⁵⁹⁷ Como indicamos en el capítulo 1 existen numerosos etiquetadores gramaticales disponibles en la actualidad, pero uno de los más utilizados en los proyectos de Lingüística Computacional es este debido a las ventajas que presenta frente a otros.

quedan etiquetadas según su categoría gramatical, lo que permite su posterior extracción a partir de una serie de patrones sintácticos en combinación con la técnica estadística TF.IDF, descrita anteriormente. Estos patrones sintácticos incluidos en la aplicación tienen como referencia el sustantivo, por lo que todos las unidades extraídas de forma automática corresponderán a un sustantivo, en caso de que se trate de la extracción de 1-grama, o a un sintagma nominal, si se extrajeran 2 o más gramas.⁵⁹⁸ En este segundo caso, encontramos sintagmas nominales del tipo N + prep + N, N +Adj. o Adj. + N, entre otros.⁵⁹⁹

A pesar de la complejidad del proceso de extracción terminológica, el usuario únicamente ha de efectuar una acción muy sencilla para ponerla en marcha. Una vez que se han añadido los documentos que forman el corpus de trabajo, deberá dirigirse a la pestaña *Recalculate term score*, dentro del módulo *Corpus*, y ahí seleccionar una de las dos opciones de extracción disponibles, a saber: 1. el número mínimo de candidatos a término que desea obtener tras la extracción por cada lengua o 2. el umbral de frecuencia mínima de aparición de cada candidato a término.

The image shows a web interface for calculating term scores. It features two input fields: 'Number of terms:' followed by a text box, and 'Frequency threshold:' followed by another text box. The word 'or' is placed between the two fields. To the right of the 'Frequency threshold' field is a blue button labeled 'submit'.

Figura 55. Opción *Recalculate term score* para la extracción de candidatos a término.

En nuestro caso, se ha seleccionado la opción que nos permite obtener un número determinado de candidatos a término y, para ello, hemos introducido la cifra de 2.000 en el campo correspondiente. Como resultado, la aplicación ha efectuado una extracción de forma automática de 6.000 candidatos a término, 2.000 por cada lengua de trabajo (español, inglés y alemán). Una vez extraídos, la aplicación nos permitirá visualizar todos los términos extraídos organizados por lenguas y validar los candidatos a término que sean considerados términos por el usuario, así como eliminar candidatos a

⁵⁹⁸ Esta idea parte del fundamento terminológico de que las dos terceras partes de todo el conjunto terminológico de un trabajo de esta naturaleza está formado por sustantivos o sintagmas nominales (Cabré, 1993).

⁵⁹⁹ Donde N corresponde al nombre o sustantivo; prep, a preposición y Adj., al adjetivo. Para consultar ejemplos de términos con estos sintagmas.

término que no sean pertinentes para el dominio de trabajo o que se consideren ruido, como por ejemplo sintagmas nominales mal contruidos, preposiciones, conjunciones, verbos irrelevantes, entre otros.

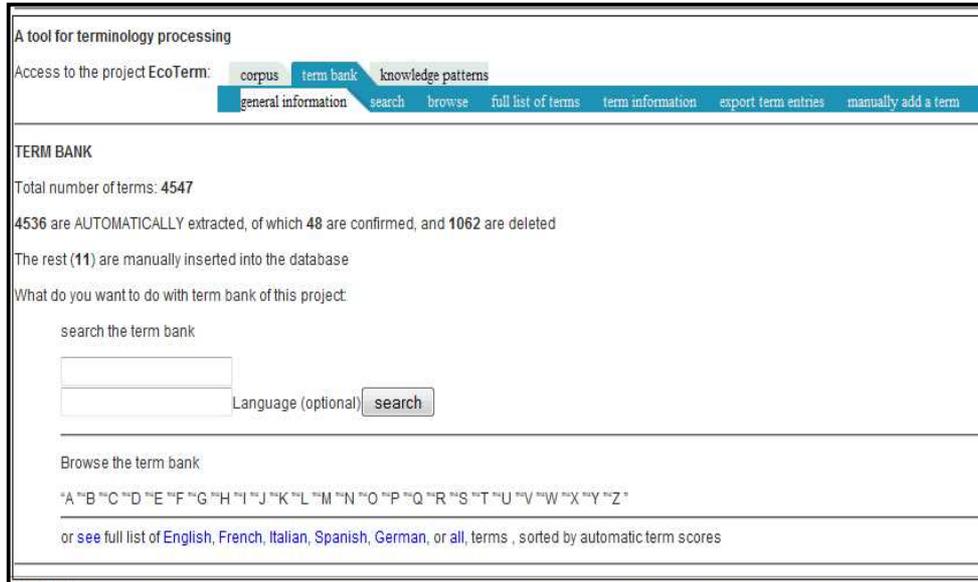


Figura 56. Módulo Base de datos con el número de candidatos a término extraídos, de términos validados, de unidades eliminadas y de términos incluidos de forma manual.

Además de la funcionalidad de extracción automática de candidatos a término, esta aplicación permite la inclusión de unidades terminológicas de forma manual en caso de que el usuario considere necesario incluir alguna unidad después de haber efectuado la extracción. Esta tarea se realiza en el submódulo *Manually add a term* del módulo Base de datos y requiere, por parte del usuario, únicamente la introducción del término que desea incluir y la selección de la lengua de la unidad de las disponibles en la herramienta: español, inglés, alemán, francés e italiano.

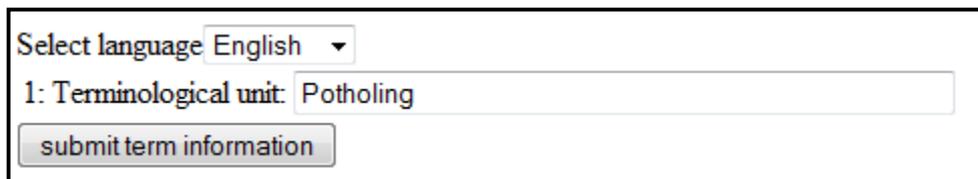


Figura 57. Ejemplo de introducción de término a través de la opción *Manually add a term*.

Una vez que se realice este paso y se guarde en la base de datos a través del botón *Submit term information*, la unidad introducida de forma manual aparecerá en la

lista de las unidades extraídas automáticamente para ser completada posteriormente en la entrada terminológica por el terminógrafo.

4.2.1.2.1. Validación y eliminación de candidatos a término

Como decíamos anteriormente, una vez extraídos los candidatos a término de forma automática por la aplicación es imprescindible que el usuario los revise y valide o elimine según los objetivos del proyecto. En nuestro caso, se ha llevado a cabo una revisión exhaustiva de la extracción a través de la cual hemos encontrado una serie de unidades con un alto índice de repetición que han tenido que ser eliminadas por no considerarse términos dentro de este dominio. Entre estas unidades, destacamos las siguientes:

- Topónimos, como nombre de países, regiones, ciudades y pueblos, así como accidentes geográficos (nombres de cordilleras, sierras, picos elevados, ríos, etc.);
- Las estaciones del año;
- Nombres de animales y de árboles;
- Unidades procedentes de la lengua general relacionados con el dominio del turismo de aventura, como «Wasser», «tierra», «air», «Wochenende», «niño», «family», etc.
- Terminología relacionada con el alojamiento, como «hotel», «recepción», «B&B», etc., y el transporte.

A partir del análisis de estas unidades eliminadas, podemos establecer una distinción de las diferentes unidades que se pueden encontrar en el dominio del turismo de aventura, sin tener en cuenta en este caso su terminología específica, que veremos a continuación: en primer lugar, encontramos los términos procedentes de la lengua común, como las estaciones del año, los topónimos y otras unidades especificadas, debido al hecho de que se trata de un lenguaje especializado dirigido a no especialistas y, por tanto, se encuentra en un nivel de especialización bajo; en segundo lugar, observamos los términos que provienen de otras lenguas de especialidad, como son por ejemplo la

geografía, la biología, la botánica o la agronomía por lo que respecta a los nombres de animales, árboles, etc. frecuentes en las descripciones de rutas a pie; y, por último, distinguimos la terminología procedente del lenguaje turístico en general, como por ejemplo la terminología relativa a los alojamientos, al transporte, a la gastronomía, etc. Asimismo, observamos otra de las principales características de este dominio de especialidad, a saber: su relación con la naturaleza y con el lugar donde se realiza, ya que, como vimos en el capítulo anterior, este segmento turístico se inserta dentro del grupo del turismo de naturaleza al desarrollarse en un entorno natural, por lo que su terminología está directamente relacionada con este ámbito, así como con la protección del medio ambiente, la flora, la fauna, etc., como se desprende de esta terminología. A su vez, se combina con las características propias del dominio del turismo en general, como son el ocio y el disfrute durante un periodo vacacional fuera del entorno habitual del usuario.

Además de estas unidades que no se pueden considerar términos dentro del proyecto que estamos llevando a cabo sobre la terminología del turismo de aventura, la herramienta también ha arrojado cierto ruido en la extracción, relacionado, por una parte, con el funcionamiento limitado del lematizador en la lengua española⁶⁰⁰ y, por otro lado, por cierta imprecisión de la técnica estadística utilizada.

En concreto, el ruido relacionado con la herramienta de lematización hace referencia a ciertas unidades cuyos lemas no se han obtenido correctamente, lo que ha dado lugar a la extracción repetida de sus formas en singular y en plural, como por ejemplo se han incluido en el resultado de la extracción «mosquetón» y «mosquetones», «tobogán» y «toboganes» o «rápel» y «rápeles»; a la inclusión de la forma conjugada de verbos en vez de su forma impersonal en infinitivo, como «descendía» en vez de «descender», entre otros. Esto se debe principalmente al hecho de que el lematizador carece de estas unidades en su vocabulario básico, el cual utiliza para llevar a cabo la lematización, lo que provoca que no reconozca estas unidades y que, por tanto, nos las procese correctamente y las incluya en el resultado final con la misma forma inicial.

⁶⁰⁰ El etiquetador *TreeTagger* no produce estos errores de forma tan frecuente ni en inglés ni en alemán al contar principalmente con un vocabulario básico más amplio en estas lenguas y, además, por no encontrarse con los problemas de flexión propios de la lengua española.

Por su parte, el ruido provocado por la aplicación *ProTermino* se corresponde con la extracción inadecuada de algunos *n*-gramas. Como veíamos anteriormente, la técnica híbrida de esta aplicación consiste en la extracción automática de *n*-gramas mediante la técnica TF.IDF a partir del etiquetado morfológico de las categorías gramaticales de las unidades. En este contexto, encontramos dos tipos de ruido relacionados con la extracción de dos o más gramas, es decir, por la extracción errónea de candidatos a término plurilexemáticos: por un lado, por la extracción de *n*-gramas con una alta repetición pero que no tienen sentido semántico, como por ejemplo «longitud igual», «muchoviaje punto de encuentro» o «panorámicas lleno de frescura», en español; «bathroom into natural gas», «battery srt kit crampons ice axe boots» o «bouldering bouldering», en inglés; o «euro zeitraum», «m charakter» o «wandern wanderweg bundesland», en alemán; y por otro, por la extracción de unidades con sentido semántico pero consideradas candidatos a término erróneos desde un punto de vista gramatical, como por ejemplo «agua bravo», «cuerda suplementario» o «ropa cómodo». Este segundo error se ha debido al hecho de que la extracción de candidatos se realiza a partir de los resultados de la lematización, por lo que la combinación de las unidades da lugar a incongruencias de este tipo, donde, por ejemplo, los adjetivos muestran su lema pero no concuerda con el sustantivo asociado. Algunas de las unidades eliminadas por este motivo se han considerado términos pertenecientes al dominio del turismo de aventura, por lo que nos hemos visto obligados a incluirlas posteriormente de forma manual con su forma correcta, a saber: «aguas bravas», «cuerda suplementaria» o «ropa cómoda», de acuerdo con los tres ejemplos indicados con anterioridad.

Con respecto al proceso de validación, nos hemos basado en el conocimiento adquirido previamente durante el análisis del dominio y de los textos producidos para seleccionar la terminología propia de este dominio, además de haber contado con la colaboración de expertos en este dominio. En general, los términos se han limitado a las categorías semánticas establecidas en la conceptualización del dominio elaborada en forma de marco semántico (cf. 3.3.7.), a saber: Agente, Actividad, Acción, Ubicación e

Instrumento (cf. Tabla 42 abajo).⁶⁰¹ Asimismo, fuera de esta conceptualización se han incluido otros términos que se consideran propios de este segmento pero que cuentan con un elevado índice de aparición, como son los términos relacionados con los tipos de seguros de viajes que se pueden encontrar relacionados con las actividades de aventura, a saber: «seguro de responsabilidad civil», «seguro de asistencia en viaje de aventura» o «seguro de viaje».

⁶⁰¹ En este paso se podrían haber realizado modificaciones de estas categorías si se hubiesen considerado inadecuadas para representar la terminología del dominio de trabajo, pero, en nuestro caso, han resultado ser adecuadas y, por tanto, no se han tenido que efectuar ningún cambio.

| Agente | Actividad | | Acción | Ubicación | Instrumento | | |
|--|--|---|---|---|--|--|---|
| | Actividad acuática | Actividades relacionadas | | | Ropa y calzado | Seguridad | Material complementario |
| Instructor especializado Guía especializado Personal especializado Técnico titulado Barranquista | Actividad acuática Descenso de cañones Barranquismo Descenso de barrancos Descenso de río Descenso de cataratas | Espeleología Rápel Rafting Escalada Alpinismo | Rapel Caminar Deslizarse Asegurar Descender | Garganta Cañón Barranco Acantilado Desfiladero cascada Poza Tobogán Pared escarpada Pared vertical Salto de agua Reserva Mundial de la Biosfera Parque natural Parque nacional | Traje de neopreno Bañador Guantes Escarpines Impermeable | Arnés de seguridad Ocho Casco Casco homologado Chaleco salvavidas Baga de anclaje Punto de anclaje Cuerda Cuerda de seguridad Cabo de anclaje Cabo de seguridad Pata de anclaje Descensor Ascendedor Mosquetón Asegurador Cuerda elástica Cuerda de rápel | Linterna Linterna frontal Mochila impermeable Bidón estanco Botiquín de primeros auxilios |

Tabla 42. Ejemplo de conceptos seleccionados referentes a la terminología de la actividad de barranquismo.

Como se puede observar en esta tabla, entre la terminología validada se encuentran términos monolexemáticos, como por ejemplo «barranquismo» o «asegurador», «mosquetón» o «escarpines», así como términos plurilexemáticos formados por 2 o más gramas: «cuerda de rápel», «parque natural», «traje de neopreno» o «descenso de barrancos». Al igual que en español, encontramos también términos monolexemáticos y plurilexemáticos tanto en inglés como en alemán, como por ejemplo: «gorge walking», «canyoning», «cagoul jacket», «eight», «wetsuit», «Neoprenanzug», «abseiling», «Abseiltaue», «Höhlentauchen», «Abstieg durch Schluchten», entre otros. Además de los *n*-gramas, en este segmento también son frecuentes las abreviaturas y acrónimos del tipo «BTT», «PVC» o «STR», por citar algunos ejemplos.

A parte de estas características destacables desde un punto de vista más general, podemos subrayar que la terminología del segmento del turismo de aventura recibe numerosos términos procedentes de los deportes de aventura, como pueden ser la escalada, el barranquismo, la espeleología, etc., que se practican de forma profesional y, por tanto, en un nivel avanzado. Sin embargo, esta terminología se encuentra limitada y, a veces, incluso adaptada para el público al que va dirigido en el sector turístico, a saber: usuarios con menos conocimientos en la actividad o que la practican con un objetivo más lúdico y de forma más suave. Así pues, en la tabla anterior encontramos términos procedentes de la actividad del barranquismo y de la escalada pero los cuales quedan limitados a la actividad realizada desde un punto de vista turístico y, por ejemplo, el número de términos relativos al equipamiento que debe llevar un barranquista es menor en el lenguaje turístico que en la actividad profesional. Asimismo, otra de las principales características del lenguaje de este segmento son los neologismos y los extranjerismos, especialmente los anglicismos (tanto en español como en alemán), que a menudo se utilizan de la misma manera en los dos idiomas, como por ejemplo «rafting», pero a veces se mantienen en un idioma pero en otro no, como es el caso de «canyoning», que en alemán se mantiene el uso de su forma original mientras que en español se prefiere utilizar el término «barranquismo» o alguna de sus variantes («descenso de cañones», «descenso de barrancos», etc.). Existe un constante crecimiento de términos nuevos, ya sean de reciente creación o evolución, debido a dos aspectos principalmente: por un lado, porque aparecen continuamente nuevas

actividades del turismo de aventura que no existían, como el «zorbing» o que son variantes de las existentes, como el «aquaseiling», para ofrecer nuevas experiencias y oportunidades a los usuarios más intrépidos; y por otro lado, porque se utiliza para atraer a los turistas con ofertas exóticas y novedosas, algo característico en el lenguaje del turismo. Así pues, la terminología de este segmento se encuentra en constante evolución y dinamismo con cambios tanto entre categorías semánticas como con apariciones de nuevos términos, que obligan a realizar una continua actualización de la terminología de trabajo.

Por último, también es característico de este segmento el fenómeno de la multidimensionalidad de los términos, en particular con respecto a las unidades que pueden utilizarse como acciones, es decir, como verbos, y aquellas que pueden utilizarse como actividades, es decir, como sustantivos, así como unidades que pueden considerarse tanto dentro de la categoría <Actividad> así como de la categoría <Instrumento>. Como ejemplo del primer caso, tenemos «montar en bicicleta», «tirarse en paracaídas», «Tauchen», «Radfahren», etc. y, como segundo caso, encontramos «ala delta», «hidrobob», «mountain bike», etc. Estos casos ocurren principalmente en español y alemán, que, al contrario que el inglés, no han formado términos nuevos para distinguir entre las acciones, los instrumentos y las actividades. En inglés, por su parte, se ha optado por la sustantivación de los verbos para denominar estas actividades de aventura, como por ejemplo «mountain biking», «diving», «caving», etc., lo que permite distinguir estas actividades de las acciones y de los instrumentos. Este hecho provoca que el inglés sea más concreto y dé lugar a menos equívocos que las otras lenguas.

4.2.1.3. Extracción de relaciones semánticas

El siguiente paso de nuestro proyecto consistirá en la selección y extracción de las relaciones semánticas pertinentes para nuestro proyecto, así como los principales patrones semánticos que las representan. Para ello, hemos utilizado como base el esquema categorial inicial que se elaboró en la primera fase del trabajo, donde se establecieron las principales categorías semánticas del dominio del turismo de aventura (Agente, Acción, Actividad, Instrumento y Ubicación) relacionadas entre sí por unas

relaciones semánticas básicas. Además de este esquema inicial, seguiremos las indicaciones propuestas por Auger y Barrière (2008: 3-7), que comentamos en el segundo capítulo (cf. 2.6.1.3.3.), que consistían en los siguientes pasos, a saber: a) definición de las relaciones semánticas de interés; b) detección de los patrones que expresan tales relaciones en el texto, y c) la búsqueda de instancias para las relaciones mediante el uso de patrones.

Siguiendo los pasos indicados por estos autores, el primer paso que debemos dar consistirá en establecer las relaciones semánticas para nuestro proyecto. Para ello, nos ha servido de gran ayuda el análisis que se llevó a cabo previamente sobre el dominio del turismo de aventura, así como la conceptualización propuesta al inicio que se realizaron en la primera fase del proyecto. Además, también se ha llevado a cabo un análisis *in vitro* a través del estudio de las definiciones ofrecidas por los recursos terminológicos que obtuvieron mayor puntuación en la fase de evaluación de los recursos disponibles (cf. 3.3.6.), con objeto de obtener más información sobre el dominio y sobre las relaciones que se establecen entre sus conceptos. Todo ello nos ha permitido establecer seis tipos de relaciones semánticas, de las que dos de ellas son jerárquicas, a saber: por un lado, encontramos las relaciones lógicas o de hiperonimia (ES_UN), que nos permiten determinar los conceptos hiperónimos y sus hipónimos; y, por otro, las relaciones ontológicas o partitivas (PARTE_TODO), que posibilitan la organización de conceptos y sus partes. A su vez, como decíamos, hemos establecido otras cuatro relaciones más, que en este caso se trata de relaciones no jerárquicas o asociativas. Estas relaciones permiten mayor dinamismo a la hora de relacionar conceptos, así como más libertad para que cada proyecto ontoterminográfico tenga la posibilidad de crear sus propias relaciones *ad hoc* y, así, pueda representar el dominio de trabajo y adaptarlas, por tanto, a sus necesidades. A continuación, presentamos las seis relaciones semánticas que se han establecido para este proyecto ontoterminográfico y sus respectivas relaciones inversas, vinculadas a las principales categorías semánticas detectadas en el dominio del turismo de aventura:

| Categorías | Relaciones directas | Relaciones inversas |
|--------------------|---|--|
| ACTIVIDAD | Relación ES_UN | Relación INVERSE_ES_UN |
| ACCIÓN | Relación REALIZA Relación CONSISTE_EN | Relación ES_REALIZADO_POR Relación ES_REALIZADO_EN Relación ES_UBICACIÓN_DE |
| INSTRUMENTO | Relación REQUIERE | Relación ES_REQUERIDO_EN Relación ES_REQUERIDO_POR |
| INSTRUMENTO | Relación PARTE_DE | Relación INVERSE_PARTE_DE |
| AGENTE | Relación ORGANIZA Relación GUÍA Relación PRACTICA | Relación ES_ORGANIZADO_POR Relación ES_GUIADO_POR Relación ES_PRACTICADO_POR |
| UBICACIÓN | Relación SE_REALIZA_EN | ES_UBICACIÓN_DE |

Tabla 43. Relaciones semánticas directas e inversas del dominio del turismo de aventura con respecto a las principales categorías semánticas.

Una vez determinadas las relaciones semánticas de este dominio de especialidad, debemos proseguir con la detección de patrones lingüísticos que expresan tales relaciones en los textos seleccionados. Para ello, utilizaremos algunos de los patrones más recurrentes para la búsqueda de las relaciones jerárquicas, como *es un tipo de*, *es un*, *incluye*, etc. (relación de hiperonimia) y *es una parte de*, *constituido por*, *contiene*, *compuesto de*, *consta de*, etc. (relación de metonimia).⁶⁰² Asimismo, llevaremos a cabo un estudio de los contextos en los que aparecen los términos extraídos por la aplicación (y validados por el terminógrafo) con objeto de observar las relaciones que se establecen con otros términos y las expresiones más características que aparecen en dicha relación. Y, por último, utilizaremos las definiciones de los recursos terminológicos seleccionados por su mayor puntuación para observar las expresiones que utilizan a la hora de describir los conceptos de este dominio. Todo esto se llevará a cabo en las tres lenguas de trabajo de nuestro proyecto ontoterminográfico: español, inglés y alemán. Una vez que se haya finalizado la búsqueda de patrones lingüísticos, se cargarán en el módulo *Knowledge patterns* de la aplicación para la posterior detección

⁶⁰² Para seleccionar estos patrones recurrentes nos hemos basado en los estudios realizados por Marshman et ál. (2002) y Feliu Cortés (2004) para la relación de hiperonimia, y en el estudio de Soler Puertes y Alcina Caudet (2008) para la relación de metonimia.

automática de términos relacionados a partir de estos patrones y para la propuesta de una ontología de forma automática.

En primer lugar, presentamos los patrones lingüísticos más recurrentes que hemos utilizado de acuerdo con la primera técnica indicada. Estos patrones hacen referencia a las relaciones jerárquicas (lógicas y ontológicas) y son las expresiones más frecuentes que aparecen en los textos de especialidad en general.

| RELACIÓN SEMÁNTICA | PATRONES EN ESPAÑOL | PATRONES EN INGLÉS | PATRONES EN ALEMÁN |
|--------------------|--|---|---|
| IS_A | <ul style="list-style-type: none"> - act* como - se basa* - está basad* en - es un* - es un tipo de - se ha* convertido en - se trata de - se denomina* - denominad* - por ejemplo - modalidad* de, - conocid* como - llamad* - variante* de - especialidad* de | <ul style="list-style-type: none"> - is a*, - is a kind of, - is a variant of, - is a sort of, - is a type of, - is/are called - is an activity, - such as, - for example, - known as, - similar to - based on - named - a variety of | <ul style="list-style-type: none"> - ist * eine * Variante - ist * eine * Sportart - besteh* in - mit der Bezeichnung - handel* sich um, - ist * ein * - wie z. B. - Zum Beispiel - z. B. - ein Beispiel - bezeichnet * w*rd* - w*rd* * bezeichnet - ist * eine * Modalität, - Im Rahmen de*, |
| PART_OF | <ul style="list-style-type: none"> - compuest* por, - se compone* de, - compuest* de, - consta* de, - pertenece* a, - perteneciente* a, - dispon* de, - clasifica* - se complet* con, - completad* con - se divide* en - dividid* en - se complementa* con, - hech* con, - inclu* - se incl* | <ul style="list-style-type: none"> - contain*, - composed of, - compris*, - is / are comprised of, - a range of, - consist* of, - completed by, - complemented by, - classified, - include*, - including, - is/are included, | <ul style="list-style-type: none"> - verfüg* über - besteh* aus, - bestehend aus, - w*rd* * eingeteilt, - eingeteilt * w*rd*, - eine Reihe von, - gehör* zu, - ist * eingestuft, - sind * eingestuft, - eingestuft * wird, - eingestuft * werden, - sind * aus - ist * aus - inkl. |

Tabla 44. Patrones lingüísticos recurrentes de las relaciones jerárquicas en español, inglés y alemán.

Para detectar los patrones del resto de relaciones creadas *ad hoc* para este proyecto, utilizaremos principalmente la técnica basada en los términos, es decir, aquella que consiste en la búsqueda de contextos en las que aparecen los términos

buscados. Para ello, tendremos que emplear otro programa independiente de nuestra aplicación, puesto que esta, a pesar de contar con un módulo de concordancias integrado, no permite la búsqueda directa de concordancias, sino a través de su plantilla de creación de entradas terminológicas, como veremos más adelante.⁶⁰³ El programa que utilizaremos será *AntCon*, un programa de gestión de corpus de libre distribución que dispone de un módulo de concordancias fácil de usar, además de otras funcionalidades.⁶⁰⁴

Este programa nos permite cargar un corpus textual y realizar búsquedas en él a través de la introducción de los términos deseados en el campo de búsqueda. Como resultado, nos ofrece una lista de líneas en las que aparece el término buscado junto con su contexto más inmediato, tanto el que se encuentra a la derecha del término como a su izquierda. A través de la realización de búsquedas de este tipo con diferentes términos en las tres lenguas de trabajo, se consigue establecer una lista de los patrones lingüísticos más habituales de este dominio, que permitirán realizar búsquedas más amplias.



Figura 58. Resultado de búsqueda de concordancias con el término «barranquismo».

⁶⁰³ Actualmente se está trabajando en la creación de un acceso directo al módulo de concordancias, pero lamentablemente esta funcionalidad no está disponible todavía.

⁶⁰⁴ Para consultar las funcionalidades del programa *AntCon*, consulte la página web de su autor: <<http://www.antlab.sci.waseda.ac.jp/software.html>>.

Por último, indicamos también el estudio *in vitro* llevado a cabo para analizar las definiciones recogidas en los recursos terminológicos fiables del campo del turismo de aventura que, como vimos en el capítulo anterior, son muy escasos desde dos puntos de vista: por un lado, existen en la actualidad un número muy limitado de recursos terminológicos sobre este dominio que contengan definiciones y, por otro, la mayoría de los disponibles son de baja calidad con respecto a su contenido. Debido a esta escasez, hemos incluido en esta búsqueda también las definiciones recogidas en los decretos autonómicos que regulan el segmento del turismo de aventura (cf. 3.3.3.1.2.). A continuación, indicamos algunos de los recursos terminológicos utilizados para el análisis de sus definiciones:

| |
|--|
| Sports and Games - Visual Dictionary Online (<i>Merrien-Webster</i>) |
| Glosario de montaña y geografía |
| Glosario de Aventureros |
| Glossary of Extreme Sports Terms |
| Lexikon der Begriffe des Karstes |

Tabla 45. Ejemplo de recursos terminológicos utilizados para el análisis *in vitro* de patrones lingüísticos.

Como exponíamos anteriormente, a través de este estudio *in vitro* se puede obtener información acerca de las relaciones establecidas entre conceptos de un campo de especialidad concreto. En nuestro caso, el estudio de las definiciones ha servido de gran ayuda a la hora de detectar estos patrones lingüísticos, especialmente las definiciones recogidas en los decretos autonómicos, puesto que ofrecen las descripciones de los términos más habituales en el dominio del turismo de aventura. A nuestro pesar, la legislación referente a este segmento turístico en español es la única que incluye definiciones, por lo que para el resto de lenguas de trabajo (inglés y alemán) nos hemos tenido que basar en las definiciones de los recursos disponibles.

Como ejemplo, vamos a mostrar a continuación algunas definiciones seleccionadas y en **negrita** indicaremos expresiones lingüísticas que representan una relación semántica.

Barranquismo (según el Decreto de Murcia, 2007)

Práctica que **consiste en** el descenso de barrancos, con o sin cauce de agua, **utilizando** técnicas de descenso **empleadas en** montañismo, escalada, espeleología y natación.

En este caso, vemos cómo encontramos en primer lugar una relación hiperonímica (ES_UN) con la unidad «Práctica», es decir, el término «Barranquismo» se considera un hiperónimo de «Práctica». A continuación, observamos otras relaciones, en este caso relaciones asociativas: por un lado, la relación CONSISTE_EN queda expresada a través de la expresión «consiste en», que indica la acción que se realiza en una actividad concreta; y por otro, la relación SE REQUIERE, que se especifica aquí a través de la expresión «utilizando» y «empleadas en» y que hace referencia al tipo de instrumento que se necesita para realizar una actividad.

Otro ejemplo lo encontramos en el Decreto de Andalucía (2002) al ofrecer la definición de «Esquí de río»:

Práctica del esquí que **consiste en descender por** aguas vivas con unos esquís especiales y **con la ayuda de** un remo de doble pala.

De nuevo, encontramos la referencia hiperonímica al inicio de la definición con la expresión «Práctica del esquí», aunque en este caso se especifica más el hiperónimo al incluir que se refiere al esquí y no a una práctica en general. Asimismo, encontramos la misma expresión que en la definición anterior relativa a la relación CONSISTE_EN, a saber: «consiste en», que indica la acción que se realiza en una actividad (al igual que el ejemplo anterior) y otras dos expresiones que hacen referencia respectivamente a la relación SE REALIZA_EN: «descender por», que menciona el lugar donde se lleva a cabo una acción, y a la relación REQUIERE: «con la ayuda de», que se refiere al instrumento necesario para dicha actividad.

Por lo que respecta a las lenguas inglés y alemán, el proceso que se sigue es exactamente el mismo y, como ejemplo, presentaremos aquí dos definiciones en sendas lenguas con sus correspondientes relaciones resaltadas en negrita.

Climbing (Sports and Games - Visual Dictionary Online, Merriem-Webster)

Leisure or competitive sport that consists of climbing up a natural rock face or an artificial climbing structure **using** bare hands and light equipment.

Künstliches Klettern (Lexikon von Fachbegriffen rund um's Klettern)

Klettern mit Einsatz technischer Hilfsmittel wie Seil, Haken, Strickleitern... zur Fortbewegung am Fels.

Al igual que en el caso anterior estas dos definiciones comienzan con la indicación del hiperónimo del término en cuestión, lo que indica una relación de hiperonimia o ES_UN a través de las expresiones «Leisure sport» y «Competitive sport» en la primera y «Klettern» en la segunda. Además, la definición en inglés hace referencia a la relación REQUIERE con la expresión «consists of», así como a la relación SE_REALIZA_EN y REQUIERE con las expresiones siguientes: «climbing up to» y «using», respectivamente. Por su parte, en la definición alemana encontramos una expresión referente a la relación REQUIERE: «mit Einsatz technischer Hilfsmittel wie», así como una expresión que nos presenta la relación SE_REALIZA_EN: «zur Fortbewegung». Gracias al empleo de esta metodología, hemos conseguido detectar gran cantidad de patrones lingüísticos, lo que nos ha permitido ampliar las búsquedas en el corpus y detectar otros términos que no habían sido extraídos por la aplicación *ProTermino*. Asimismo, nos ha permitido llevar a cabo una revisión de las definiciones incluidas en los recursos terminológicos y, así, observar ciertos errores repetitivos relacionados principalmente con la estructura que presenta la definición. Estos errores consisten especialmente en la falta de coherencia a la hora de escribir definiciones, es decir, no se sigue una estructura única para definir los términos que pertenecen a la misma categoría. De esta manera, el resultado es una variedad de estructuras definicionales que rompen la homogeneidad y, por tanto, la calidad del recurso. Como ejemplo, mostramos tres definiciones encontradas en el recurso *Glossary of Rock Climbing and Mountaineering Terms* referentes a términos de la categoría <Instrumento>:

Belay Device

A piece of equipment used by a belayer through which the climbing rope is threaded and which acts a brake on the rope in the event of their partner falling. The device is attached to the harness, and is used to lock off a rope during a fall. Belay devices include the figure eight, the Sticht plate and belay tubes. In some cases, the climbers own body can be used as a belay device.

Ascender

A device for climbing the rope; for example a simple prusik.

Karabiner

An oval or D-shaped device of lightweight aluminium or alloy that serves as a link between protection, slings and ropes. It has a spring-loaded gate through which a climbing rope can be threaded. This may be a locking-gate karabiner (krab), where the gate is locked off with a screw (or some other safety device), or a snap-gate krab, where the gate is held shut with a spring. Karabiners are used variously for such activities as belaying, abseiling, and clipping into safety anchors.

Como se observa, la estructura de estas definiciones es diferente en cada caso a pesar de que todos los términos hacen referencia a tipos de instrumentos que se utilizan en la escalada. Este tipo de errores se evitarían si se utilizaran plantillas definicionales basadas en marcos semánticos, como proponemos en nuestra metodología y que presentaremos más adelante. De lo contrario, ocurre como en estos casos, cada definición sigue un patrón propio y diferente del resto de sus co-hipónimos.

El resultado final del empleo de las técnicas de búsqueda de patrones lingüísticos que hemos comentado hasta el momento ha consistido en la elaboración de una lista de patrones en las tres lenguas de trabajo, la cual presentamos a continuación dividida según la relación semántica correspondiente:

| RELACIONES SEMÁNTICAS | PATRONES EN ESPAÑOL | PATRONES EN INGLÉS | PATRONES EN ALEMÁN |
|----------------------------|---|---|---|
| - REALIZA - CONSISTE_EN | - consist* en - basad* en - se basa* en - requier* - se requier* - con finalidad* de | - involv* - requir* - based on - with the aim of | - betreib* - treib* - besteh* in |
| - requiere | - empleando - mediante el empleo de - emplead* en - el empleo de - se utiliza* - utiliz* - se realiza* con / mediante - usad* para - se usa* - con la ayuda de - ayudándose de / con - ayudad* por - de ayuda - necesit* - se necesita* - necesari* es - necesari* son - es necesari* - se sirve* de - sirv* para - aprovech* - auxiliad* por - equipad* con - propulsad* por - como instrumento* | - required is / are - is / are required - required for - use* - using - the use of - mak* use of - utilis* - practis* with - is / are necessary - need* for - by means of - equipped with - provid* - suppl* - be of help - with the help of - launch* by - deploy* | - braucht * man - man * braucht - w*rd** gebraucht - gebraucht * w*rd* - verwendet * wird - w*rd* * verwendet - verwendet * w*rd* - w*rd* benutzt - benutzt * w*rd* - nutz* - dabei * sein, - ist * dabei - sind * dabei - mit deren Hilfe - h*If* * dabei - Bereitstellung von - Bereitstellung d* - mit Einsatz - Hilfsmittel wie |

Tabla 46. Lista de patrones lingüísticos detectados en las tres lenguas.

| | | | |
|---|---|---|--|
| <p>- ORGANIZA - GUÍA - PRACTICA</p> | <p>- aconsej* - realiz* - organiz* - utiliz* - practic* - ayud* - con la ayuda de - ofrec* - necesit* - proporcion* - acompañ* - enseñ* - impart* - impartid* por - se guía* - guiad* por - dirigid* por - con la* recomendaci* de - servicio* de - amante de*</p> | <p>- taught by - teach* - offer* - organis* - supply* - is / are supplied - guid* - help* - with the help of - need* - lead* - made by - with the support of - the service* of - is / are responsible for - in the safe hands of - practi*ed by - recommend* - advis*</p> | <p>- organisier* - w*rd* * vermittelt - vermittelt w*rd* - behilflich * sind - sind * behilflich - behilflich * ist - ist * behilflich - biet* - geboten - geboten w*rd* - in Begleitung von - mit Hilfe - spezialisiert * auf</p> |
| <p>- SE_REALIZA_EN</p> | <p>- se practica* en - se practica* dentro de - desarrollad* en - descieñd* por - descend* por - deslizz* por - desplaz* por - ascieñd* por - recorr* - escal* en - realiz* en - se encuentra* en - transcurre* por - está* situado* en - se halla* - a bordo de - a través de - un recorrido por - por encima de - apoyado* en - zona de - en pleno corazón de - dónde</p> | <p>- found at - found in - found throughout - in contact with - in the heart of - where - known in - located on - located in - situated in - situated on - tak* place - practi* on - practi* in - climb* up</p> | <p>- organisier* in - im Kontakt mit - an Bord ein* - man findet - findet man - w*rd* gefunden - gefunden * w*rd* - w*rd* * gefunden - gefunden * w*rd* - ist * gefunden - gefunden * ist - sind * gefunden - gefunden * sind - überall - entlang - zur Fortbewegung</p> |

Tabla 46 *Cont.* Lista de patrones lingüísticos detectados en las tres lenguas.

Como hemos indicado con anterioridad, una vez obtenidos estos patrones a través de las búsquedas de contextos con el buscador de concordancias del programa *AntCon* así como mediante la consulta de definiciones en recursos terminológicos, volveremos a la aplicación *ProTermino* y seguiremos con los siguientes pasos del proyecto.

4.2.1.4. Elaboración de la ontología de dominio

Para la elaboración de la ontología del dominio del turismo de aventura utilizaremos el esquema categorial en forma de marco semántico que elaboramos en la primera fase (cf. 3.3.7.) y, a partir de él, realizaremos modificaciones según los conocimientos adquiridos hasta el momento. En este contexto, se podrán incluir, eliminar o mover las relaciones y las categorías principales determinadas al comienzo del proyecto para, posteriormente, vincular los diferentes términos encontrados dentro del corpus de trabajo con los diferentes conceptos de este esquema prototípico. Con todo ello, obtendremos una ontología en forma de marco semántico completa y adecuada para el turismo de aventura, a través de la cual se podrá conectar el nivel conceptual y el nivel lingüístico.

En nuestro caso concreto, consideraremos válido el marco semántico elaborado en la fase inicial del proyecto, puesto que, por un lado, incluye las categorías semánticas más relevantes para la representación del dominio del turismo de aventura y, por otro lado, recoge las relaciones semánticas adecuadas de acuerdo con nuestro análisis y nuestras necesidades. Por lo tanto, el paso de modificación no sería necesario para nuestra investigación y, así, pasaremos directamente al paso de especificar el nivel conceptual. Este segundo paso se llevará a cabo de dos formas diferentes: en primer lugar, concretaremos algunas categorías semánticas para adaptarlas al dominio de trabajo y, en segundo lugar, clasificaremos los conceptos de este dominio según el marco semántico resultante.

Con respecto al primer paso especificado, a saber: completar las categorías semánticas determinadas, presentamos a continuación el marco semántico general donde se especifican estas inclusiones de forma clara y concreta:

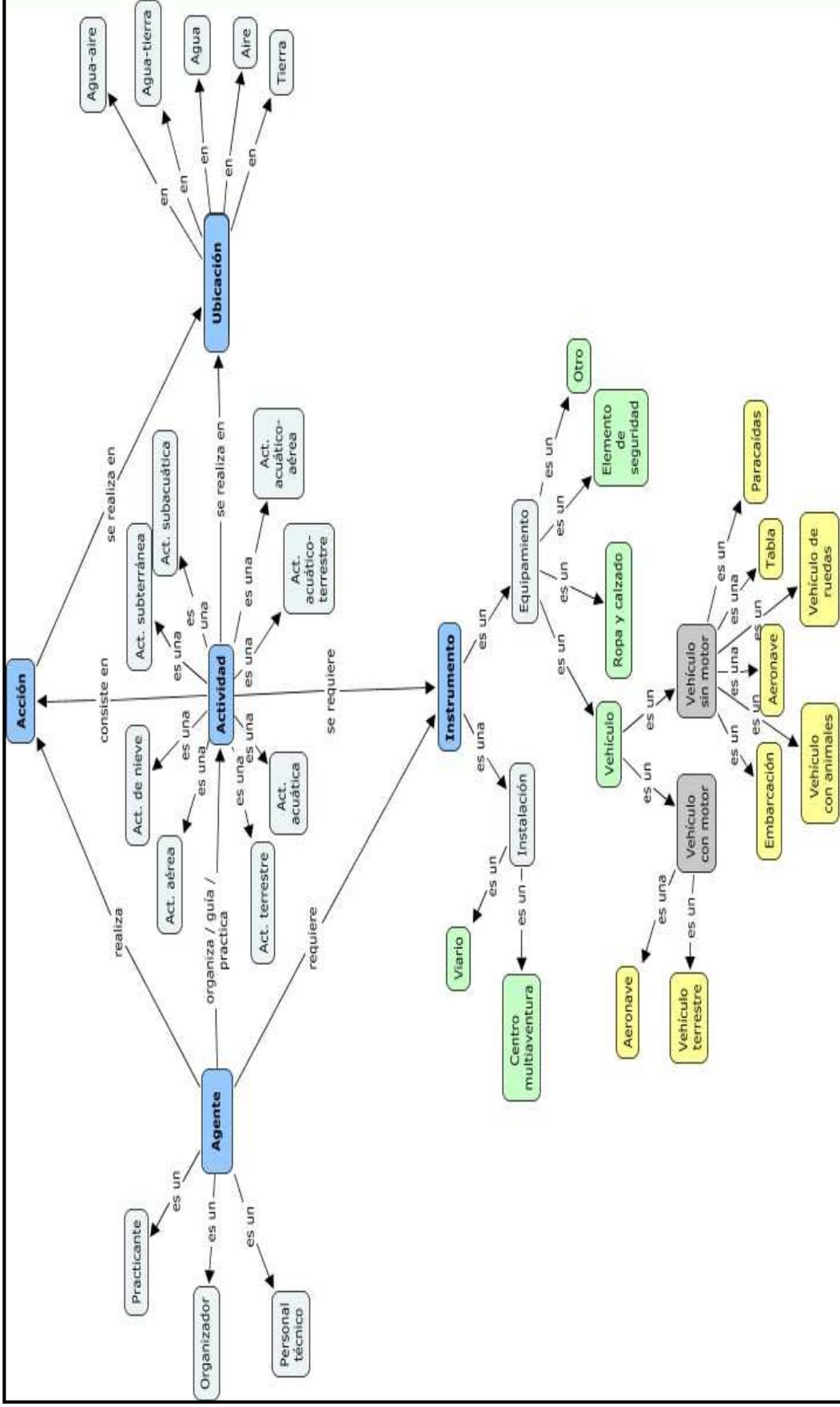


Figura 59. Marco semántico con especificaciones en categorías semánticas principales.

Como se observa en la figura anterior, el marco semántico formado por las categorías semánticas iniciales se han concretado con el fin de completar la representación del dominio del turismo de aventura en el nivel conceptual y, por tanto, independiente de cualquier lengua. Estas nuevas inclusiones, como se puede ver en la figura, se realizan dentro del marco semántico pero de forma jerárquica, es decir, con la relación de hiperonimia, puesto que es necesario a este nivel establecer un orden de hiperónimo-hipónimo de los conceptos que se encuentran en el dominio del turismo de aventura. Con objeto de clarificar el esquema anterior, presentaremos a continuación una explicación relativa a cada categoría y sus diferentes subcategorías:

- En la categoría <Agente> se determinan tres tipos de agentes diferentes, a saber:
 1. <Organizador>, es decir, persona que se encarga de organizar la actividad de aventura;
 2. <Personal técnico>, que se dedica a dirigir, instruir y guiar las actividades;
 3. <Practicante>, que se corresponde con la persona que practica la actividad de aventura.

- En la categoría <Actividad> se diferencian varios tipos de actividad:
 1. <Actividad acuática>, es decir, aquella que se practica en el medio acuático;
 2. <Actividad terrestre>, que hace referencia a la que se lleva a cabo en el medio terrestre;
 3. <Actividad aérea>, que tiene lugar en el aire;
 4. <Actividad acuático-terrestre>, que incluye actividades que combinan el medio acuático y el terrestre;
 5. <Actividad acuático-aérea>, que se corresponde con las actividades que combinan el medio acuático y el aéreo;
 6. <Actividad subacuática>, que hace referencia a las actividades que tienen lugar en el medio subacuático,
 7. <Actividad subterránea>, que se refiere a actividades que tienen lugar en el medio subterráneo, y
 8. <Actividad de nieve>, que se realiza en la nieve.

- En la categoría <Instrumento> se distinguen principalmente dos tipos:
 1. <Instalación> y
 2. <Equipamiento>.Dentro de estos dos tipos, realizamos una distinción más concreta: en cuanto a las instalaciones, establecemos un segundo nivel de <Viario> y de <Centro multiaventura>; y en cuanto al equipamiento, distinguimos
 1. <Vehículo>,
 2. <Calzado y ropa>;
 3. <Elemento de seguridad>;
 4. <Otro>.Además de esta segunda clasificación, consideramos oportuno realizar una mayor concreción en relación con el tipo de vehículo que se puede

encontrar. Así, distinguimos en un tercer nivel entre <Vehículo con motor> y <Vehículo sin motor> y, a su vez, un cuarto nivel donde se especifican los tipos de vehículos con motor y sin motor que se pueden encontrar: en primer lugar, como vehículos con motor, se indican <Aeronave> (para el aire) y <Vehículo terrestre> (para la tierra); y en segundo lugar, encontramos los diferentes tipos de vehículos sin motor: <Embarcación>, <Vehículo con animales>, <Aeronave>, <Vehículos con ruedas>, <Tabla> y <Paracaídas>.

- En la categoría <Ubicación> se clasifican varios tipos de medio natural donde se pueden realizar las diferentes actividades de aventura: <Agua>, <Tierra> y <Aire>, así como la combinación de algunos de ellos (<Aire-tierra> y <Agua-aire>). Además, dentro de estas, se incluyen otras categorías más específicas, con objeto de especificar aún más la representación, a saber: <Roca>, <Nieve>, <Subterráneo> y <Subacuático>.

Una vez que se ha obtenido la ontología completa del turismo de aventura, es decir, que se ha realizado la representación conceptual de dicho dominio, se daría por finalizado el trabajo en el nivel conceptual y, por tanto, independiente de la lengua. De esta manera, se pasaría a trabajar al nivel lingüístico con los términos concretos del turismo de aventura en las diferentes lenguas de trabajo (en nuestro caso, español, inglés y alemán), prestando especial atención a las variaciones denominativas y culturales que podamos encontrar. Para ello, utilizaremos el módulo de ontologías que incluye la aplicación *ProTermino*, que pasaremos a describir a continuación, el cual nos facilitará la adquisición y la organización de la terminología extraída del corpus de trabajo.

El primer paso que debemos llevar a cabo en esta fase consiste en la introducción en la aplicación *ProTermino*, concretamente en el módulo *Knowledge patterns*, de los patrones semánticos detectados según los resultados del paso anterior. Para ello, simplemente necesitamos seleccionar dicho módulo de la pantalla principal de la aplicación y seleccionar un archivo de texto (formato .txt) que contenga los patrones de cada lengua de trabajo. Como resultado, encontramos una tabla formada por tres columnas en la que se indica la lengua del patrón, la relación a la que hace referencia y el patrón en sí, respectivamente.

A tool for terminology processing

Access to the project EcoTerm: corpus term bank knowledge patterns

| Language | Relation | Pattern |
|----------|-----------------|------------------|
| English | ORGANIZA | organise |
| English | GUÍA | teach |
| English | REQUIERE | practice with |
| English | ES_GUIADO_POR | tought by |
| English | REQUIERE | utilise |
| English | REQUIERE | make use of |
| English | ES_REQUERIDO_EN | used for |
| English | ES_UN | is considered |
| English | ES_UN | is considered as |
| English | ES_UN | is a type of |
| English | PARTE_DE | include |
| English | ES_UN | is a sort of |
| English | ES_UN | is a variant of |
| English | ES_UN | is a kind of |
| English | ES_UN | is a |
| English | ES_UN | are some |
| English | ES_UN | such as |

Tabla 47. Ejemplo de patrones semánticos incluidos en *ProTermino* a través del módulo *Knowledge patterns*.

A partir del momento en el que están disponibles dichos patrones en la aplicación, *ProTermino* funciona de forma totalmente automática y propone relaciones semánticas entre términos que encuentra en el corpus de trabajo. Dichas relaciones propuestas de forma automática se encuentran disponibles en la parte inferior de la entrada terminológica de cada unidad, con objeto de que el terminógrafo tenga a su disposición dicha información durante el tiempo en el que esté trabajando con los campos de dicha unidad. En todo momento, el terminógrafo tiene la posibilidad de realizar modificaciones o eliminaciones de estos patrones, así como inclusiones de patrones nuevos, en caso de que lo considere pertinente para el proyecto. Una vez que se han validado las propuestas ofrecidas de forma automática por la aplicación estas pasan a incluirse en la tabla de relaciones confirmadas (*confirmed relations*), desde la que se pueden eliminar también.

| confirmed relations | delete |
|---|--------------------------|
| gorge walking« ES_UN »canyoning | <input type="checkbox"/> |
| gorge walking« ES_ORGANIZADO_POR »qualified staff | <input type="checkbox"/> |
| gorge walking« ES_UN »sport activity | <input type="checkbox"/> |
| gorge walking« SE_REALIZA_EN »canyon | <input type="checkbox"/> |
| gorge walking« SE_REALIZA_EN »gorge | <input type="checkbox"/> |
| gorge walking« REQUIERE »wetsuit | <input type="checkbox"/> |
| other activity ES_UN gorge walking | <input type="checkbox"/> |
| water activity ES_UN gorge walking | <input type="checkbox"/> |
| activity ES_UN gorge walking | <input type="checkbox"/> |

Tabla 48. Ejemplo de relaciones semánticas validadas a partir de las relaciones propuestas automáticamente por *ProTermino*.

A partir de este ejemplo, observamos las diferentes relaciones que hemos establecido para el dominio del turismo de aventura así como los vínculos que existen entre diferentes términos que aparecen en el corpus (en azul). Se observa, por ejemplo, cómo la relación ES_UN es la más frecuente de todas, así como otras relaciones asociativas como la relación REQUIERE, SE_REALIZA_EN o la relación ES_ORGANIZADO_POR. Además de esta opción, la aplicación *ProTermino* permite la visualización de las relaciones detectadas una vez que han sido confirmadas por el terminógrafo, como se puede observar en la siguiente figura:



Figura 60. Visualización de las relaciones semánticas confirmadas.

A pesar de no poder considerarse esta representación gráfica una ontología,⁶⁰⁵ puesto que carece de la organización jerárquica adecuada además de que no representa las relaciones entre todos los conceptos incluidos, consideramos, sin embargo, que sirve de gran ayuda a la hora de buscar instancias (ejemplos reales) para la ontología elaborada con anterioridad, al permitir adquirir mayor conocimiento acerca del dominio del turismo de aventura a través del análisis del corpus de trabajo. Asimismo, permite detectar nuevos conceptos y encontrar instancias en el corpus de trabajo, es decir, ejemplos concretos de los patrones semánticos detectados en las fases previas.

De esta manera, utilizando como base esta representación gráfica que ofrece *ProTermino*, podremos completar la ontología de domino que hemos elaborado anteriormente, que mostraba el nivel conceptual del campo de especialidad, añadiendo los términos correspondientes extraídos del corpus de trabajo. En otras palabras, pasaremos del nivel conceptual al nivel terminológico. Para ello, será necesaria la elaboración de un marco semántico para cada actividad del turismo de aventura que hemos considerado en nuestra investigación.⁶⁰⁶ Como ejemplo de marco semántico elaborado, presentaremos a continuación el marco creado para la actividad de barranquismo con algunos de los términos incluidos en la Tabla 42.

⁶⁰⁵ Actualmente se está estudiando la posibilidad de implementar un editor ontológico dentro de la herramienta que permita la elaboración de ontologías completas, pero se encuentra aún en fase de prueba.

⁶⁰⁶ Cf. Tabla 16.

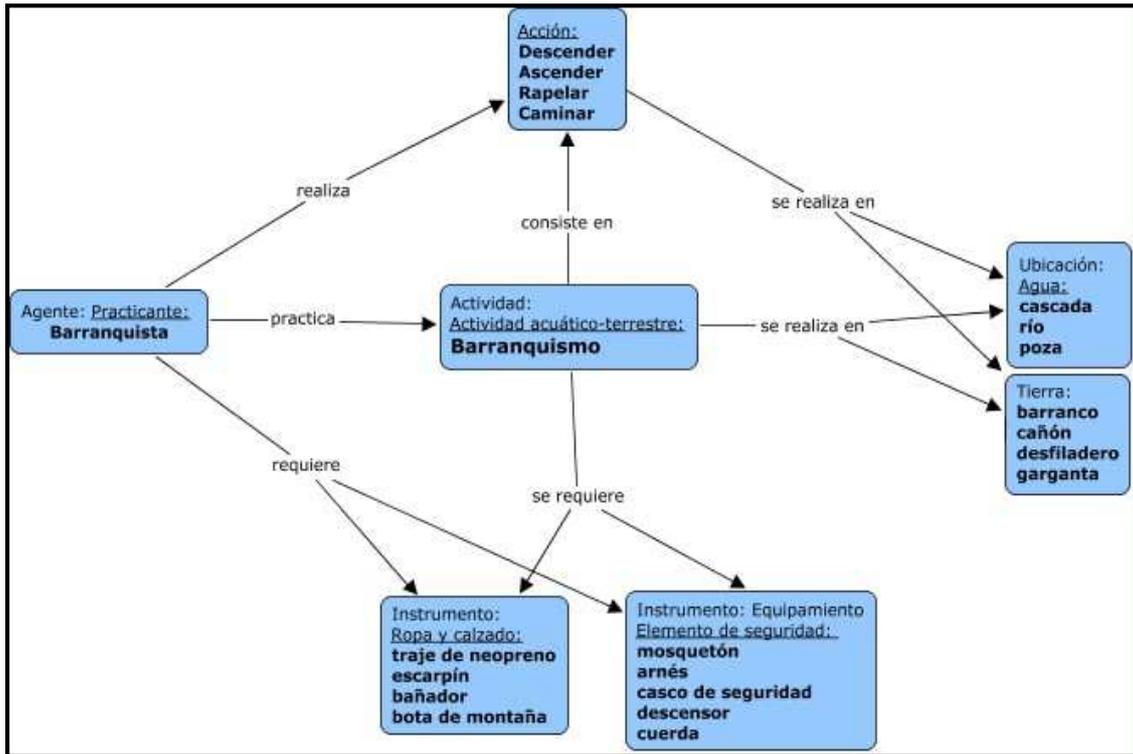


Figura 61. Ejemplo de marco semántico de la actividad de barranquismo.

La utilización de estos marcos semánticos para la terminología validada del dominio del turismo de aventura nos permite llevar a cabo varias tareas, a saber: 1) organizar claramente la terminología de este segmento, 2) observar las posibles carencias y huecos dentro de algunas categorías semánticas, 3) desambiguar términos polisémicos, que adquieren significados diferentes según el marco en el que aparezcan, como por ejemplo «salto» que puede ser «salto de agua» (en la actividad de barranquismo) o «salto en tándem» (en el paracaidismo), y 3) elaborar las definiciones correspondientes basadas en plantillas definatorias, como veremos en la fase posterior. Como conclusión, cabe destacar que la validación de los términos extraídos automáticamente por la aplicación, así como la organización de esta son esenciales para cualquier trabajo terminográfico y que, por tanto, deben ser llevados a cabo con sumo cuidado por el terminógrafo, ya que los resultados obtenidos en esta fase serán determinantes para la calidad del trabajo que se lleve a cabo en las fases siguientes.

4.2.2. ELABORACIÓN DE LA BASE DE DATOS ONTOTERMINOGRÁFICA

Llegados a este punto en el que hemos llevado a cabo los pasos relacionados con la gestión del trabajo ontoterminográfico, como han sido la subida del corpus, la extracción de candidatos a término y su validación, la búsqueda de relaciones semánticas, entre otros, debemos continuar con la elaboración de la base de datos, es decir, con la inclusión de información en las entradas terminológicas de las diferentes unidades validadas. Para ello, presentaremos a continuación los diferentes campos que se han establecido en la base de datos para cada entrada concreta de la terminología de nuestro campo de especialidad y, posteriormente, comentaremos los aspectos más destacados relacionados con alguno de estos campos de forma particular, a saber: la elaboración de definiciones mediante marcos semánticos, la búsqueda de concordancias y contextos, la selección de equivalentes de traducción mediante cognados y corpus paralelos, así como la creación (semi)automática de ontologías en la aplicación.

4.2.2.1. Campos terminológicos para el proyecto *EcoTerm*

Los campos que se encuentran disponibles en la herramienta *ProTermino* están basados en la propuesta de campos de bases de datos terminológicas de la norma ISO 12620 *Computer applications in terminology – Data categories* (1999) (cf. 2.6.1.4.1.), pero se ha llevado a cabo una selección de acuerdo con el grupo de destinatarios al que irá dirigido el producto final, esto es, los traductores. Para ello, hemos tenido en cuenta los resultados obtenidos en la encuesta realizada *ad hoc* para conocer sus necesidades y expectativas, que presentamos en el capítulo segundo (cf. 2.5.3.). De esta manera, se ha establecido la microestructura de cada entrada en la base de datos con los siguientes campos, a saber: 1) *terminological unit* (unidad terminológica); 2) *part-of-speech* (categoría gramatical); 3) *gender* (género), 4) *grammatical number* (número); 5) *term status* (estatus del término); 6) *standard body* (entidad estandarizadora); 7) *geographical usage* (uso geográfico); 8) *subdomain* (subdominio); 9) *definition* (definición); 10) *context* (contexto); 10a) *collocation* (colocación); 11) *nontextual*

illustration (imagen); 12) *equivalence* (equivalencia); 13) *term type* (tipo de término); y 14) *linguistic remark* (notas).⁶⁰⁷

Dentro del proyecto *EcoTerm*, creado concretamente para la elaboración de nuestra base de datos ontoterminográfica sobre la terminología del turismo de aventura, hemos considerado pertinente reducir el uso de estos campos a trece, ya que hemos excluido el campo número 6 (*standard body*), al no existir ninguna entidad estandarizadora concreta para este dominio,⁶⁰⁸ así como el campo número 11 (*nontextual illustration*), al encontrar dificultades a la hora de su exportación en formato TBX.⁶⁰⁹ Por lo demás, todos los campos permanecen disponibles para la introducción de información terminológica.

En primer lugar, encontramos el campo relativo a la unidad terminológica, es decir, el campo donde se debe especificar cuál es la unidad de trabajo en dicha entrada concreta. A continuación, procedemos a especificar la información gramatical de cada entrada mediante los campos 2) *part-of-speech*; 3) *gender* y 4) *grammatical number*, los cuales ofrecen la posibilidad de seleccionar la opción adecuada para cada unidad. Así pues, el campo *part-of-speech* incluye las opciones de *n: Noun*, *v: Verb*, *adj: Adjective*, *adv: Adverb*, *o: Other*, tal y como se observa en la siguiente figura:

⁶⁰⁷ La base de datos presenta estos campos en inglés debido al hecho de que ha sido implementada en la Universidad de Wolverhampton, en el seno del Research Group in Computational Linguistics, concretamente por el Dr. Ha, aunque el diseño ha venido de la mano del Dr. Mitkov, la Dra. Corpas Pastor y, de nuevo, el Dr. Ha, con la colaboración de la Dra. Castillo Rodríguez y la doctoranda Durán Muñoz.

⁶⁰⁸ Los términos incluidos en la base de datos provienen de la extracción de los candidatos a término realizada por la aplicación, y su posterior validación por parte del terminógrafo, así como de algunos textos legislativos españoles. En este sentido, estas fuentes no se pueden considerar como estandarizadoras, lo que obliga al terminógrafo a obviar este campo.

⁶⁰⁹ El material ilustrativo, ya sea imagen, símbolo u otro, es una fuente muy útil desde un punto de vista semántico, ya que ayuda a la comprensión del significado de la unidad terminológica. En nuestro caso, sin embargo, nos hemos visto obligados a no incluirlo debido a que todavía no se encuentra disponible esta funcionalidad para la exportación en formato TBX.

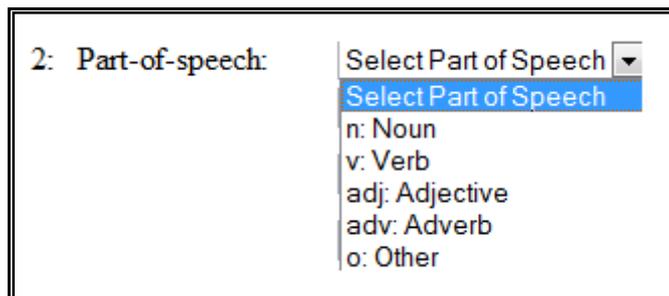


Figura 62. Opciones disponibles para el campo *part-of-speech*.

El campo *gender* contiene cinco opciones diferentes según el género que pueda presentar una unidad, a saber: *m: Masculine*, *f: Feminine*, *n: Neuter*, *o: Other* y *ng: No gender assigned*, como se muestra en la siguiente figura:

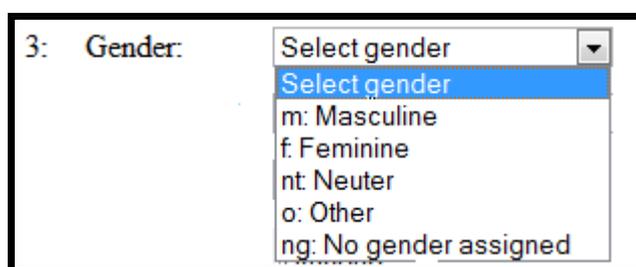


Figura 63. Opciones disponibles para el campo *gender*.

Por su parte, el campo *grammatical number* incluye solamente dos instancias, *sing: Singular* y *pl: Plural*, como se ilustra en la siguiente captura de pantalla con el desplegable para este campo:

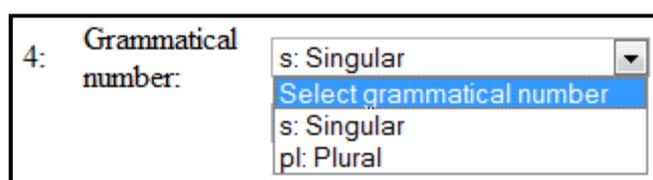


Figura 64. Opciones disponibles para el campo *grammatical number*.

A modo de ejemplo, mostramos la cumplimentación de estos cuatro primeros campos indicados hasta el momento para la unidad terminológica «barranquismo»:

| | | |
|----|---------------------|--------------|
| 1: | Terminological unit | barranquismo |
| 2: | Part-of-speech: | n: Noun |
| 3: | Gender: | m: Masculine |
| 4: | Grammatical number: | s: Singular |

Figura 65. Ejemplo de los cuatro primeros campos rellenos de la entrada terminológica relativa a la unidad «barranquismo».

El siguiente campo de la base de datos, el campo 5) *term status*, hace referencia a la situación en la que se encuentra un término dentro de su dominio particular. Este campo contiene las siguientes instancias, que se han extraído de la norma ISO 12620 (1999): a) *standardized term* (término normalizado), cuando se trata de una unidad que ha sido normalizada por una entidad normalizadora; b) *preferred term*, si un organismo normalizador es el que recomienda el uso de un término; c) *regulated term* (término regulado), si el término en cuestión viene definido por algún tipo de regulación gubernamental;⁶¹⁰ d) *recommended term* (término recomendado), en caso de que se recomiende el uso de un término concreto por parte de un experto en la materia; e) *translator proposal*, si es el traductor el que propone la UT; f) *terminologist proposal*, cuando la propuesta viene determinada por un terminógrafo; y g) *no status assigned*, cuando no se tiene constancia alguna del estatus del término en cuestión.

A continuación, se presenta el menú desplegable con las opciones disponibles permitidas para este campo en concreto:

⁶¹⁰ En el caso del turismo de aventura, muchos términos aparecen recogidos en textos legislativos, pero no se podrían considerar como *regulated terms* (términos regulados), puesto que se tratan de términos acuñados de forma individual por las diferentes comunidades autónomas y no existe unanimidad entre ellas y con respecto al uso (cf. 3.3.3.1.2.).

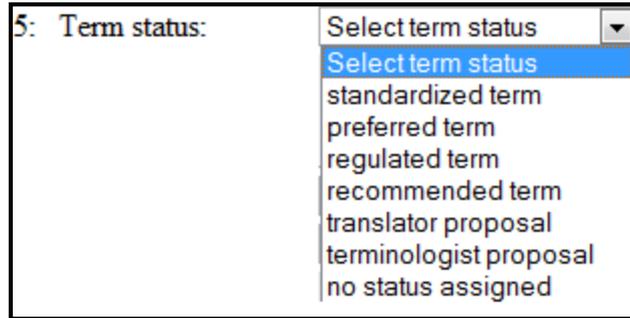


Figura 66. Opciones disponibles para el campo *term status*.

Relacionado con el campo anterior, encontramos a continuación el campo *Standard body*, donde se debe especificar el nombre del organismo que ha recomendado o normalizado una determinada unidad terminológica, siempre y cuando se haya producido dicha normalización. Para esto, se debería haber seleccionado una de las opciones pertinentes en el campo *Term status*, a saber: *standardized term*, *preferred term* o bien *regulated term*. En caso de que hubiese más de un organismo involucrado, se podrían añadir tantas casillas de cumplimentación como fueran necesarias a partir del enlace *click to add more*, que está disponible dentro del campo. Por su parte, si el término no tuviese relación con ninguna entidad estandarizadora, se dejaría este campo sin cumplimentar.⁶¹¹

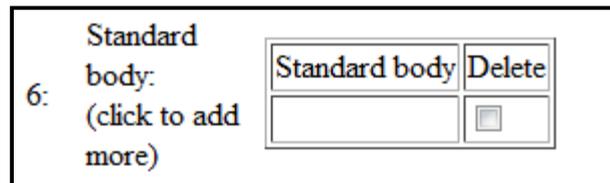


Figura 67. Campo *Standard body*.

El uso geográfico está representado mediante el campo denominado *Geographical usage*, en el que el usuario debe seleccionar el país donde se utiliza la unidad en cuestión. Partiendo de las lenguas de trabajo que permite la aplicación *ProTermino*, se han incluido los siguientes países y sus respectivos códigos según la norma ISO 3166 *Maintenance Agency* (1974), concretamente la norma ISO 3166-1

⁶¹¹ En nuestro caso, hemos considerado que los términos incluidos en nuestra base de datos han sido propuestos por especialistas, es decir, se trata de *recommended terms*, ya que no existe ninguna entidad que se encargue de regular esta terminología, con lo que no cabe ninguna especificación a este respecto.

alfa-2 para códigos de países y áreas independientes de dos letras, a saber: DE, para Alemania; ES, para España; FR, para Francia; GB, para Reino Unido; IR, para Irlanda; IT, para Italia, y US, para los Estados Unidos.

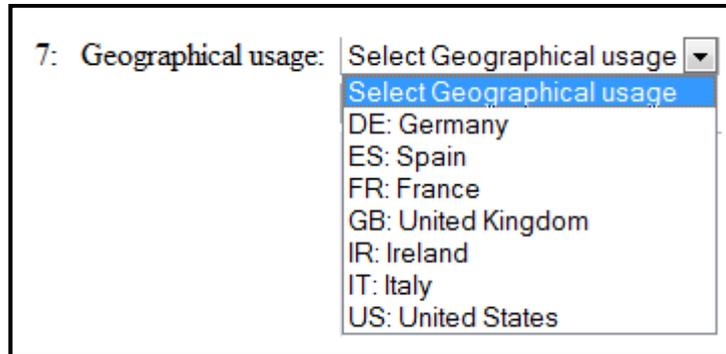


Figura 68. Opciones disponibles para el campo *Geographical Usage*.

A continuación, encontramos el campo *Subdomain*, donde se recogen los diferentes segmentos turísticos que pueden ser objeto de trabajo, tanto en este momento concreto con el segmento del turismo de aventura como en el futuro. Así, encontramos los siguientes: *business tourism*, *cultural tourism*, *language tourism*, *leisure tourism*, *adventure tourism*, *sun and beach* y *wellness and beauty*, como se observa en la siguiente figura:

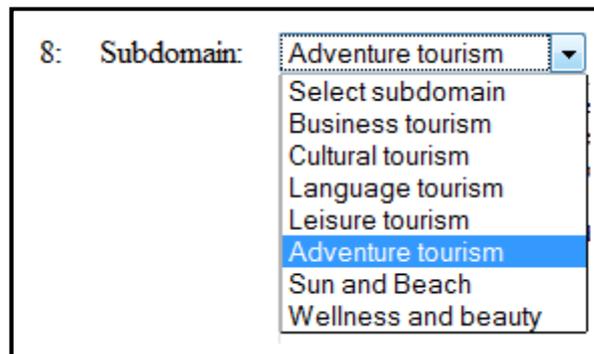
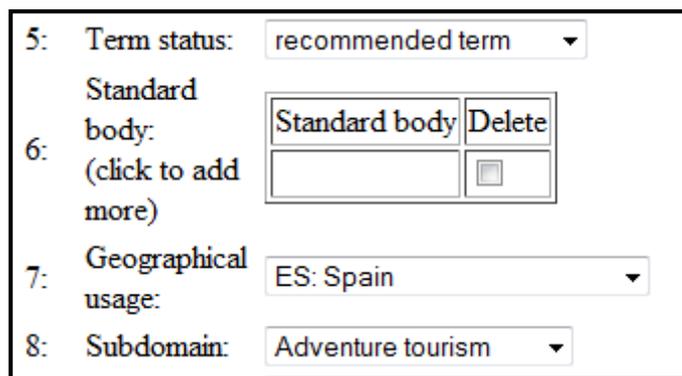


Figura 69. Opciones disponibles para el campo *Subdomain*.

A modo de resumen, presentamos a continuación los siguientes cuatro campos indicados hasta el momento con la información del término «barranquismo», con objeto de seguir completando la ficha de dicha unidad.



5: Term status: recommended term

Standard body: Standard body

6: (click to add more)

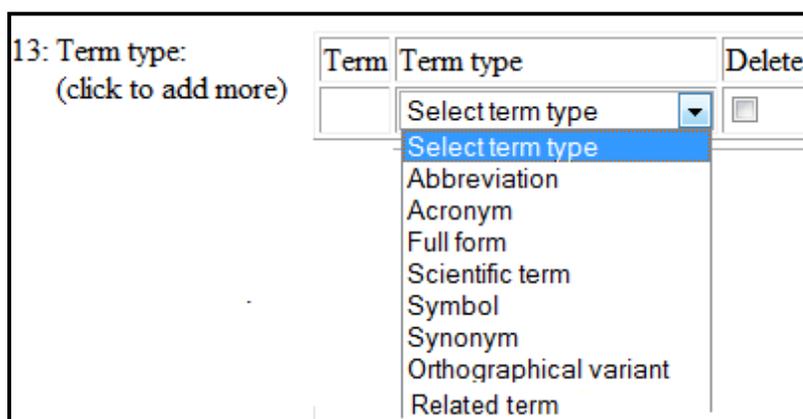
7: Geographical usage: ES: Spain

8: Subdomain: Adventure tourism

Figura 70. Ejemplo de los siguientes cuatro campos rellenos de la entrada terminológica relativa a la unidad «barranquismo».

Los cuatro campos que siguen hacen referencia a la definición, al contexto, a las colocaciones y los equivalentes de traducción de la unidad terminológica. Al considerarse de unos campos complejos y que requieren de una explicación más profunda por nuestra parte, tanto en lo que respecta a la metodología utilizada para completarlos como a la información concreta que los constituyen, los trataremos de forma separada en los siguientes apartados de este capítulo.

A continuación, encontramos el campo *Term type*, que incluye las siguientes opciones de selección: *abbreviation* (abreviatura), *acronym* (acrónimo), *full form* (forma completa), *scientific term* (término científico), *symbol* (símbolo), *synonym* (sinónimo), *orthographical variant* (variante ortográfica) y *related term* (término relacionado), como se observa en la siguiente figura:



| Term | Term type | Delete |
|------|------------------------|--------------------------|
| | Select term type | <input type="checkbox"/> |
| | Abbreviation | |
| | Acronym | |
| | Full form | |
| | Scientific term | |
| | Symbol | |
| | Synonym | |
| | Orthographical variant | |
| | Related term | |

Figura 71. Opciones disponibles para el campo *Term type*.

En este campo, al igual que en otros anteriormente mencionados, se pueden incluir más de un tipo de término a través del enlace *click to add more*. De esta manera, se podrían especificar diferentes tipos de términos, y sus correspondientes unidades, relacionadas con el término de la entrada terminológica, lo que aporta gran variedad y riqueza al resultado y ayuda al usuario final a acercarse al entorno terminológico de la unidad objeto de interés. Así pues, para continuar con la ejemplificación iniciada, presentamos en la siguiente figura los tipos de términos que aparecen en la entrada de la unidad «barranquismo»:

| | | | |
|---------------------------------------|-----------------------|--------------|--------------------------|
| 13: Term type: (click to add more) | Term | Term type | Delete |
| | Descenso de barrancos | related term | <input type="checkbox"/> |
| | Barranquismo | full form | <input type="checkbox"/> |
| | Descenso de cañones | related term | <input type="checkbox"/> |

Figura 72. Ejemplo de tipos de términos en la entrada de «barranquismo».

Por último, llegamos al campo *linguistic remarks* (notas), donde el usuario puede añadir todo tipo de información adicional que atañe a cualquiera de los campos propuestos anteriormente y que considere relevante para mejorar la comprensión de la unidad. Por ejemplo, en algunos casos se puede considerar relevante incluir información acerca de los equivalentes que se incluyen en la entrada; en otros casos, por el contrario, puede ser necesario proporcionar notas acerca de las diferencias culturales entre conceptos, etc. A continuación, mostramos el campo *Linguistic remarks* que hemos elaborado para la entrada «barranquismo», a fin de ofrecer un ejemplo de información que puede aparecer aquí.

| | |
|-------------------------|---|
| 14: Linguistic remarks: | <p>Esta actividad se oferta como actividad terrestre o acuática, dependiendo de si los barrancos o cañones por los que se discurren llevan agua o no.</p> <p>El término "barranquismo" se considera como sinónimo de "descenso de barrancos" pero también como hiperónimo de las diferentes actividades de descenso similares, como son el "descenso de cañones", el "descenso de torrentes", el "descenso de cascadas", etc., aunque a veces</p> |
|-------------------------|---|

Figura 73. Ejemplo de información añadida al campo *Linguistic remarks*.

Como se observa, en esta ocasión se ha utilizado para aclarar dos puntos importantes: por un lado, las diferentes posibilidades que se tiene a la hora de realizar este tipo de actividad; y por otro, los diferentes términos relacionados con la entrada principal y que pueden dar lugar a equívocos.

Este campo se vuelve imprescindible a la hora de elaborar recursos para los traductores, puesto que a menudo son necesarias las explicaciones para concretar algunos aspectos de las unidades terminológicas que no hayan quedado claras con el resto de campos y que puedan causar problemas de traducción como, por ejemplo, ambigüedad, uso incorrecto en un contexto concreto, distinciones entre cuasi sinónimos, etc. De esta manera, en nuestro trabajo defendemos la necesidad de incluir explicaciones de este tipo en este campo relativas a posibles problemas de uso que pueda encontrar el traductor, discrepancias entre lenguas, usos para contextos y registros concretos, entre otros muchos motivos.

4.2.2.1.1. Campo Definition

En el campo correspondiente a la definición del término, campo *definition*, el terminógrafo debe introducir la definición terminológica correspondiente a la unidad de trabajo.



Figura 74. Campo *definition* en la aplicación *ProTermino*.

Para la elaboración de las definiciones, emplearemos la metodología de trabajo expuesta en el capítulo segundo (cf. 2.6.1.4.2.), es decir, una metodología que consiste en la elaboración de plantillas definatorias basadas en marcos semánticos para los conceptos pertenecientes a una misma categoría conceptual. Concretamente, en nuestra investigación elaboraremos diferentes plantillas basadas en el marco semántico elaborado para el dominio del turismo de aventura (cf. Figura 49) para las distintas

categorías que hemos distinguido en dicho campo de especialidad, a saber: Agente, Acción, Actividad, Ubicación e Instrumento, y sus correspondientes clasificaciones internas.

Como veíamos en el capítulo anterior, las plantillas definatorias presentarán la estructura básica de cualquier definición terminológica, a saber: *Definiendum* = *genus* + *differentiae*, donde el *Definiendum* es el término que se debe definir; el *genus* se corresponde con su hiperónimo, que vendrá determinado por la relación jerárquica IS-A; y el *differentiae* hace referencia a las características propias que lo diferencian de otros conceptos, y que será introducido por las relaciones asociativas establecidas para cada dominio.

Siguiendo las pautas establecidas en esta metodología para obtener definiciones sistemáticas y completas que representen el marco semántico al que pertenecen, al igual que los principios básicos que se deben seguir a la hora de elaborar definiciones adecuadas, presentamos a continuación las diferentes plantillas definatorias que hemos elaborado para cada categoría principal.

| CATEGORÍA | PLANTILLA DEFINITORIA |
|---|---|
| ACTIVIDAD | 1. [TIPO DE ACTIVIDAD] que consiste en [ACCIÓN] en/sobre/a través de [UBICACIÓN] con un equipamiento básico formado por [INSTRUMENTO: ROPA Y CALZADO] y [INSTRUMENTO: ELEMENTO DE SEGURIDAD]. |
| | 2. [TIPO DE ACTIVIDAD] que consiste en [ACCIÓN] en/sobre/a través de [UBICACIÓN] con un equipamiento básico formado por [INSTRUMENTO: ROPA Y CALZADO], [INSTRUMENTO: ELEMENTO DE SEGURIDAD] y que se realiza en/sobre [INSTRUMENTO: TIPO DE VEHÍCULO]. |
| AGENTE 1. Organizador 2. Personal técnico 3. Practicante | 1. [TIPO DE AGENTE] que organiza una excursión de [ACTIVIDAD]. 2. [TIPO DE AGENTE] que dirige una excursión de [ACTIVIDAD]. 3. [TIPO DE AGENTE] que realiza una excursión de [ACTIVIDAD]. |
| INSTRUMENTO 1. Instalación 1.1. Viario | 1. [TIPO DE INSTALACIÓN] que se requiere para la realización de [ACTIVIDAD]. 1.1.1. [TIPO DE INSTALACIÓN] de tamaño <i>x</i> que se requiere para la realización de [ACTIVIDAD] y está equipado con [INSTRUMENTO: ELEMENTO DE SEGURIDAD]. 1.1.2. [TIPO DE INSTALACIÓN] de tamaño <i>x</i> que se requiere para la realización de [ACTIVIDAD] y no está equipado con elementos de seguridad. |
| 2. Equipamiento 2.1. Vehículo 2.2. Ropa y calzado, elemento de seguridad y otro. | 2.1. [TIPO DE VEHÍCULO] que se requiere para la realización de [ACTIVIDAD]. 2.2.: [TIPO DE EQUIPAMIENTO] específico/ básico que se emplea en [TIPO DE ACTIVIDAD], como en [ACTIVIDAD] para [ACCIÓN]. |
| UBICACIÓN | [TIPO DE UBICACIÓN] donde se practica [ACTIVIDAD] |
| ACCIÓN | [ACCIÓN] en/por/a través de [UBICACIÓN] en la práctica de [ACTIVIDAD]. |

Tabla 49. Categorías conceptuales y sus correspondientes plantillas definitorias.⁶¹²

Con respecto a cada categoría semántica que hemos tenido en cuenta en esta investigación, nos hemos visto obligados en ocasiones a elaborar diferentes plantillas

⁶¹² Cuando indicamos [TIPO DE...] nos referimos a las subcategorías que se han indicado en la ontología del dominio (cf. Figura 53).

definitorias que presentan pequeñas modificaciones, con objeto de definir correctamente las diferentes subcategorías que establecidos en apartados anteriores. Por este motivo, encontramos, por ejemplo, que la categoría <Instrumento> presenta varias plantillas para definir la subcategoría <Instalación>, <Viario>, así como las diferentes opciones de <Equipamiento>. Por su parte, también se han establecido varias plantillas para casos como, por ejemplo, la categoría <Actividad>, puesto que hay que distinguir entre aquellas actividades que requieren vehículo para llevarse a cabo de aquellas que no lo necesitan.

Para elaborar estas plantillas definitorias ha sido imprescindible el uso de la ontología de dominio creada con anterioridad (cf. Figura 53), ya que, gracias a ella, se ha podido observar el dominio del turismo de aventura organizado de forma clara y concisa y se han podido comprobar las relaciones establecidas entre los diferentes conceptos. Como ejemplo de definición a partir de estas plantillas, presentamos a continuación algunas definiciones de términos pertenecientes a diferentes categorías y subcategorías.

| CATEGORÍA | DEFINICIÓN |
|-------------|--|
| ACTIVIDAD | Barranquismo: Actividad terrestre que consiste en descender caminando, saltando, deslizándose, nadando y manejando cuerdas a través de barrancos, cataratas, ríos, cañones, gargantas, desfiladeros, etc. con un equipamiento básico formado por un traje de neopreno, guantes, escarpines, arnés de seguridad, mosquetón, casco y cuerda. |
| AGENTE | Barranquista: Practicante que realiza una excursión de barranquismo. |
| INSTRUMENTO | Descensor: Elemento de seguridad básico que se emplea en actividades de montaña, como el barranquismo y la escalada, para llevar a cabo un descenso seguro. |
| UBICACIÓN | Barranco: Quebra profunda producida en la tierra principalmente por las corrientes de las aguas donde se practica el barranquismo. |
| ACCIÓN | Descender: Bajar a través de un barranco, garganta, desfiladero, etc. en la práctica del barranquismo. |

Tabla 50. Definiciones basadas en las plantillas definitorias correspondientes.

Además de las definiciones según las plantillas definitorias que hemos expuesto anteriormente, debemos indicar que a menudo nos vemos obligados a incluir en el campo de observaciones de la entrada terminológica información complementaria con objeto de aportar datos relevantes al usuario con respecto a términos concretos. En nuestro caso, hemos intentado homogeneizar también la información que se incluye en

este campo sobre los diferentes términos. Así pues, la información complementaria que hemos añadido de forma general ha estado relacionada con el material complementario que puede requerir una actividad de aventura concreta, con las características propias de las actividades,⁶¹³ entre otros aspectos relevantes.

En definitiva, podemos concluir con respecto a este campo que el hecho de haber empleado la ontología de dominio para elaborar las plantillas definatorias así como el haber elaborado las definiciones a partir de ellas nos ha permitido obtener unas definiciones de la terminología del turismo de aventura homogéneas, completas y concisas.

4.2.2.1.2. Campo Equivalence

En la aplicación, el campo de la equivalencia (*equivalence*) se destina a la introducción de los equivalentes de traducción de la unidad en la que se está trabajando en un determinado momento. Para ello, se dispone de una tabla que permite al usuario introducir el equivalente correspondiente a la unidad principal y a la lengua de trabajo que desee, así como indicar el género que presenta dicha unidad.⁶¹⁴ A continuación, se presenta un ejemplo de esta tabla con los equivalentes en inglés y alemán correspondientes al término «barranquismo» para continuar con el ejemplo iniciado:

| | Language | Equivalent term | Gender | Delete |
|---|----------|-----------------|--------------------------|--------------------------|
| Equivalence: 12: (click to add more) | English | Canyoning | ng: No gender assigned ▼ | <input type="checkbox"/> |
| | German | Canyoning | nt: Neuter ▼ | <input type="checkbox"/> |
| | German | Schluchteln | nt: Neuter ▼ | <input type="checkbox"/> |

Figura 75. Ejemplo de equivalentes de traducción en el campo *equivalence* para el término «barranquismo».

⁶¹³ En el caso del barranquismo, por ejemplo, se ha indicado que puede realizarse en seco, es decir, como actividad terrestre, o en húmedo, como actividad acuático-terrestre, dependiendo de si hay agua o no en el recorrido.

⁶¹⁴ En algunos casos, como ocurre con el inglés, la indicación del género del equivalente es innecesaria, puesto que las unidades carecen de la marca de género y, por tanto, se debe especificar la falta de la asignación de género (*no gender assigned*). Por el contrario, en otros casos, como en la lengua alemana y española, el género del equivalente es necesario para facilitar el uso de una unidad concreta.

La búsqueda y validación de estos equivalentes de traducción se trata de una de las partes más complejas del trabajo ontoterminográfico, ya que es aquí donde entran en contacto las diferentes lenguas de trabajo y donde surgen las discrepancias por cuanto respecta al uso de cada unidad dentro de cada cultura, la equivalencia total o parcial de los diferentes términos, la existencia o no de términos en las diferentes lenguas, entre otros aspectos. Ante esta posible situación, el terminógrafo debe ser capaz de solventar estos problemas y de aportar soluciones, ya sean mediante la creación o adaptación de formas extranjeras (en el caso de neologismos) o mediante la inclusión de información contextual, marcas de uso o cualquier otro dato que ayude al destinatario del producto final a comprender los usos y las posibles diferencias entre unas unidades y otras. En este proceso, la ayuda de especialistas en la material es imprescindible, puesto que son estos los que realmente conocen el discurso especializado y las posibles diferencias de uso que se pueden encontrar entre una unidad y otra. En nuestro caso, la terminología del turismo de aventura no presenta grandes discrepancias entre unas lenguas y otras por dos razones principalmente: por un lado, la mayoría de las unidades, especialmente las actividades de aventura, tienen un origen reciente y común que se encuentra principalmente en el mundo anglófono, ya sean los Estados Unidos, Australia o Inglaterra;⁶¹⁵ y, por otro lado, gran cantidad de terminología, concretamente la que designa el equipo necesario (tanto básico como complementario) de las diferentes actividades, es, por lo general, objetiva y no caben, por tanto, discrepancias, a excepción de posibles variantes denominativas en la misma lengua.⁶¹⁶ No obstante, también hemos encontrado ciertos problemas de equivalencia con algunos términos, especialmente relacionados con los aspectos culturales y legislativos de cada país implicado. Este es el caso, por ejemplo, de los títulos necesarios para guiar, organizar o conducir las actividades de aventura, ya que estas posiciones se encuentran reguladas por una legislación propia en cada país. Por citar un ejemplo, en España la persona que conduzca una actividad de barranquismo debe ser un técnico titulado con el título de Técnico Deportivo en Barrancos, que otorga el Ministerio de Educación y Ciencia de

⁶¹⁵ Algunos términos provienen también del francés, como «rappel», o del alemán, como «Abseil», pero su presencia es menor que los procedentes del inglés.

⁶¹⁶ Téngase en cuenta que, al hablar de equipo utilizado en actividades de aventura, nos referimos únicamente al equipo que se emplea en actividades turísticas y no en profesionales, donde podrían encontrarse diferencias por lo que respecta al equipo básico y complementario que se utiliza en cada cultura.

acuerdo con el REAL DECRETO 318/2000, de 3 de marzo de 2000. Este título acredita, concretamente, que su titular posee las competencias necesarias para conducir a individuos o grupos por barrancos, organizar actividades de descenso de barrancos y programar y ejecutar la enseñanza de la técnica y táctica de esta actividad. Por su parte, el gobierno alemán y el inglés cuentan con otra legislación al respecto y solicitan requisitos diferentes a los recogidos en este Real Decreto español.

Así pues, la búsqueda de equivalencias se vuelve compleja, como indicábamos al comienzo de esta sección, ya que el terminógrafo debe tener en cuenta muchos factores, además de los propiamente necesarios para realizar las búsquedas de estos equivalentes. Actualmente, gracias a los avances informáticos, existen diferentes formas de detectar candidatos a equivalente, como veíamos en el capítulo segundo (cf. 2.6.1.3.2.), que van desde una completa automatización del proceso a una detección prácticamente manual. No obstante, a pesar de la posibilidad de realizar detecciones de forma automática, siempre será necesaria la revisión de los resultados por parte de un terminógrafo para que valide o elimine las unidades recuperadas. Además de tener la posibilidad de realizar búsquedas más o menos automáticas, las técnicas actuales permiten llevar a cabo este trabajo a partir de corpus paralelos, es decir, textos originales y sus traducciones, o corpus comparables, donde todos los textos son originales.

En nuestro caso concreto, la aplicación *ProTermino* nos ofrece la posibilidad de realizar búsquedas de equivalentes a partir de nuestro subcorpus comparable, AVENTURIS 1, mediante la búsqueda de cognados de un candidato a término validado previamente. La técnica utilizada para llevar a cabo esta búsqueda de cognados se denomina *distancia de Levenshtein*, conocida también como *distancia de edición*, cuyo objetivo consiste en calcular la diferencia entre dos secuencias de símbolos.⁶¹⁷ Concretamente, esta técnica calcula el número de modificaciones que se necesitan para transformar una secuencia de símbolos o caracteres en otra, ya sea añadiendo, eliminando o cambiando uno por otro. En este sentido, cuanto menor sea el número de modificaciones necesarias mayor será la similitud entre dichas secuencias.

⁶¹⁷ Esta técnica fue creada e implementada por el científico ruso Vladimir Levenshtein a mediados del siglo XX (cf. Levenshtein, 1966). En Cáceres González (2008), se presenta una revisión de esta técnica, incidiendo en su funcionamiento y en los principales campos en los que se ha aplicado hasta el momento.

Desde el desarrollo de esta técnica a mediados del siglo XX, se ha venido utilizando en diferentes campos, entre los que destacan la bioinformática, principalmente con respecto a las mediciones de similitudes y diferencias entre secuencias de ADN, así como la Lingüística Computacional con relación a los correctores ortográficos, sistemas de reconocimiento de voz, sistemas de alineación de corpus paralelos, entre otros. En el tema que nos ocupa, a saber: la búsqueda de equivalentes, también se han realizado diferentes estudios utilizando esta técnica con el objetivo de reconocer cognados en dos lenguas diferentes (cf. McTait, 2001; Mitkov et ál., 2008) y los resultados obtenidos han sido muy positivos.

En nuestro caso, la implementación de esta técnica en la aplicación *ProTermino* ha permitido localizar candidatos a equivalente de forma rápida y sencilla, lo que ha agilizado en gran medida la labor del terminógrafo en este sentido, a pesar de requerir la revisión y validación posterior por parte de este.⁶¹⁸ Como ejemplo, podemos observar la siguiente figura, donde aparece una tabla en la que se recogen las diferentes opciones de candidatos a equivalente que proporciona la aplicación con respecto a la unidad «*Orientierung*».

| suggested translation | | | | |
|-----------------------|----------|-------------|------------|--------------------------|
| term lemma | language | term status | term score | selected? |
| orienteering | English | AUTOMATIC | 45 | <input type="checkbox"/> |
| briefing | English | AUTOMATIC | 6 | <input type="checkbox"/> |
| writing | English | AUTOMATIC | 7 | <input type="checkbox"/> |
| ski touring | English | AUTOMATIC | 7 | <input type="checkbox"/> |

| term lemma | language | term status | term score | selected? |
|-------------|----------|-------------|------------|--------------------------|
| orientación | Spanish | AUTOMATIC | 22 | <input type="checkbox"/> |
| puenting | Spanish | AUTOMATIC | 11 | <input type="checkbox"/> |

confirm translations

Figura 76. Opciones de candidatos a equivalentes arrojados por la aplicación de forma automática para la unidad alemana «*Orientierung*».

⁶¹⁸ La técnica de *distancia de Levenshtein* se utiliza también en la aplicación en el módulo denominado *term friends*, a través del cual la aplicación sugiere al terminógrafo una serie de posibles variantes terminológicas, particularmente variantes ortográficas. En este caso, esta técnica es muy útil, puesto que calcula las modificaciones de caracteres que, por lo general en estos casos, son escasas. Por ejemplo, en el caso de la unidad inglesa «*dive*», la aplicación ofrece como posibles variantes «*diving*», «*dives*», «*diver*», «*drive*» y «*dyke*», de las cuales el usuario deberá elegir las adecuadas e introducirlas en el campo *term type* si lo considera oportuno.

Una vez obtenidos los diferentes candidatos, el terminógrafo solo debe seleccionar la casilla de la unidad correspondiente y pulsar el botón *confirm translations* para llevar a cabo la validación de los candidatos a equivalente. A continuación, las unidades validadas aparecen de forma automática en sus casillas correspondientes como los equivalentes de la unidad con la que se está trabajando.



Figura 77. Ejemplo de equivalentes validados para la unidad alemana «Orientierung».

A pesar de las ventajas que presenta esta funcionalidad, debemos indicar que no es posible localizar todos los equivalentes de todas las unidades mediante la búsqueda de cognados en corpus comparables (textos originales), ya que a menudo no se obtienen los resultados deseados, especialmente cuando se trabaja con lenguas alejadas como pueden ser el alemán y el español.⁶¹⁹ Por ejemplo, en el caso que se muestra en la siguiente figura se observa que, para la unidad inglesa «helmet», la aplicación muestra una serie de candidatos a equivalentes en español y en alemán, de los cuales solo en alemán se observa un equivalente correcto: «Helm».

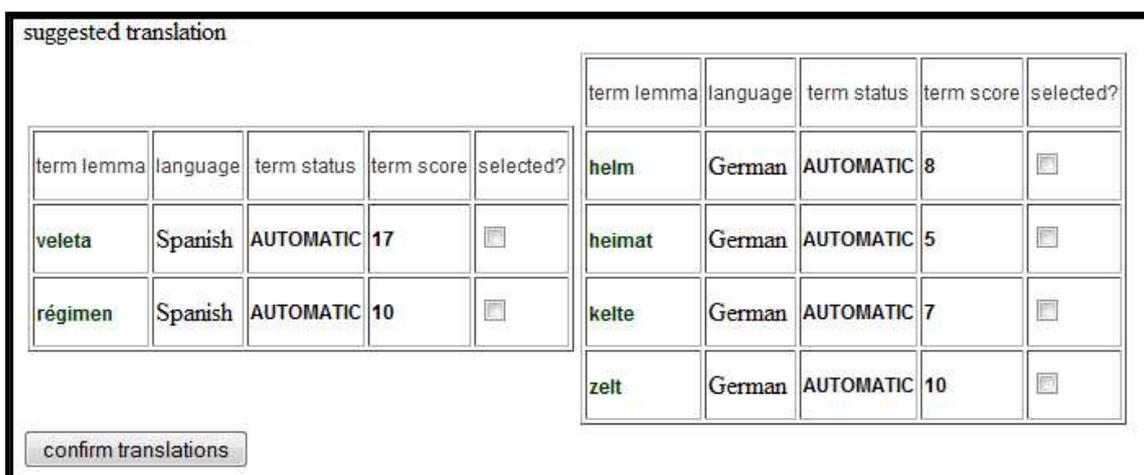


Figura 78. Candidatos a equivalente ofrecidos por la aplicación para la unidad inglesa «helmet».

⁶¹⁹ Con la aplicación *ProTermino*, se han llevado a cabo también trabajos con lenguas romances (español, italiano y francés) y los resultados han sido mucho mejores (cf. Castillo Rodríguez, 2010).

En este sentido, podemos confirmar que, a pesar de que el empleo de corpus comparables, como AVENTURIS 1, nos aporta gran cantidad de información original, tanto en lo que respecta a términos como a relaciones, a veces puede resultar limitado a la hora de extraer equivalentes de componentes en diferentes lenguas, en nuestro caso de los componentes español, inglés y alemán. Por ello, consideramos que la mejor opción para llevar a cabo la extracción de equivalencias consiste en el empleo de un corpus comparable y, a su vez, de un corpus paralelo, con objeto de disfrutar de las ventajas que una y otra opción ofrecen, a saber: por un lado, los corpus paralelos son una fuente de equivalencias de traducción que permiten la extracción de estas de forma rápida y sencilla; y por otro, los corpus comparables ofrecen la posibilidad de observar los términos en su contexto real y original y exentos de ambigüedades y errores semánticos.

Siguiendo esta línea del trabajo, en nuestro caso hemos utilizado, por un lado, nuestro corpus comparable, AVENTURIS 1, para obtener candidatos a equivalente mediante la detección de cognados en la aplicación *ProTermino*, como hemos indicado anteriormente; y por otro lado, hemos utilizado el corpus paralelo AVENTURIS 2 para localizar candidatos a equivalente que no hubieran sido detectados todavía. Para este proceso de detección a partir del corpus paralelo, hemos utilizado de nuevo el módulo de cognados de *ProTermino* y, así, hemos repetido todos los pasos que hemos expuesto hasta el momento pero, en este caso, para el corpus paralelo. De esta manera, hemos obtenido otros candidatos a equivalente de traducción que aparecen en el corpus paralelo y que la aplicación ha buscado de forma (semi)automática y que no aparecían en el corpus comparable.⁶²⁰ Como último paso para finalizar esta tarea de búsqueda de equivalentes, hemos procedido a revisar y validar todos los candidatos a equivalente arrojados por la aplicación.

⁶²⁰ Esta búsqueda de equivalentes se podría haber completado con la utilización de alineadores bilingües a nivel de párrafo u oración y su posterior extracción terminológica (Le Serrec et ál., 2010), pero en nuestro caso no ha sido necesario ya que la búsqueda de equivalentes mediante cognados ha sido satisfactoria. No obstante, no descartamos llevar a cabo en el futuro un estudio siguiendo la metodología propuesta por estos autores.

4.2.2.1.3. *Campo Context y Collocations. Módulo de concordancias*

El campo *Context* se ha destinado a la inclusión de todos aquellos contextos relevantes en los cuales aparecen unidades terminológicas relevantes para el campo de especialidad objeto de estudio, en nuestro caso el turismo de aventura. Como vimos en el capítulo segundo (cf. 2.6.1.4.3.), este campo ocupa un lugar predominante en cualquier recurso terminológico, puesto que desempeña un doble papel (Dubuc y Lauriston, 1997: 81): por un lado, se considera como prueba del uso que se realiza de la unidad en un campo de especialidad y permite, así, asociar un contenido semántico a dicha unidad; y por otro lado, da la posibilidad de determinar las relaciones específicas que existen entre un término y su campo de especialidad a través de la identificación de conceptos. En este sentido, y con referencia a nuestro grupo potencial de destinatarios, los traductores profesionales, podríamos añadir también que permite la visualización del uso que se le da a un término concreto dentro de su campo de especialidad, lo que facilita la comprensión de su significado y de sus relaciones con otros términos y su contexto.

En nuestro proyecto, hemos seguido esta línea de trabajo y, por tanto, hemos completado el campo *Context* para todas las unidades validadas. Al contrario que en otras bases de datos, donde solo se incluye el contexto en la lengua en la que se da el término principal, en nuestro caso hemos considerado oportuno incluir el contexto en las tres lenguas de trabajo, a saber: en español (lengua del término principal), en alemán y en inglés, con el fin de facilitar la comprensión de los usuarios y de mejorar el uso de la unidad correspondiente. Así pues, se diferencian tres contextos diferentes de acuerdo con los tres términos de cada lengua.

Para realizar la selección de los contextos que hemos incluido en cada entrada terminológica, hemos utilizado el módulo de concordancias disponible en la aplicación, al que se accede desde la misma plantilla terminológica donde se encuentran el resto de campos. Dentro de este módulo el usuario tiene acceso a todas las concordancias que aparecen en el corpus de trabajo del término que está analizando y puede consultarlas y ordenarlas a partir de las palabras que aparecen a la derecha o a la izquierda de la

unidad de búsqueda.⁶²¹ Una vez que ha detectado el contexto adecuado⁶²² para que aparezca en el campo *Context* de la entrada terminológica, el terminógrafo solo debe seleccionarlo y pinchar en el botón derecho del ratón para que aparezca, de forma automática, en dicho campo.

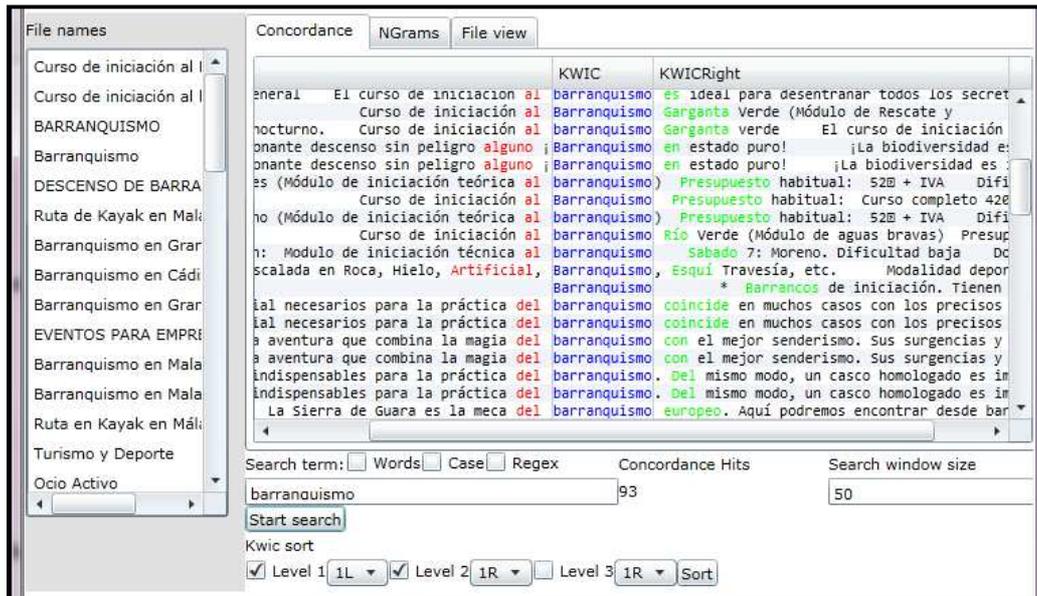


Figura 79. Ejemplo de concordancias para el término «barranquismo» con el módulo *Term concordance*.

⁶²¹ A este módulo de concordancias se tiene acceso únicamente desde la entrada terminológica de cada unidad, lo que agiliza la búsqueda de contextos y de colocaciones durante el proceso de completar la entrada. Sin embargo, se está trabajando actualmente en la herramienta para que este módulo esté disponible también fuera de las entradas terminológicas, es decir, que también tenga acceso desde la pantalla principal de *ProTermino* o a través de una pestaña independiente.

⁶²² Recordemos que el contexto debe cumplir una serie de requisitos: 1. El contexto debe extraerse de una fuente fiable y ser textos originales; 2. El contexto debe ser conciso, pero completo; 3. El contexto debe contener la unidad terminológica que pretenda ilustrar en la misma categoría gramatical de la entrada, y 4. El contexto debe ser representativo del sistema conceptual al que pertenece (cf. 2.6.1.4.3.).

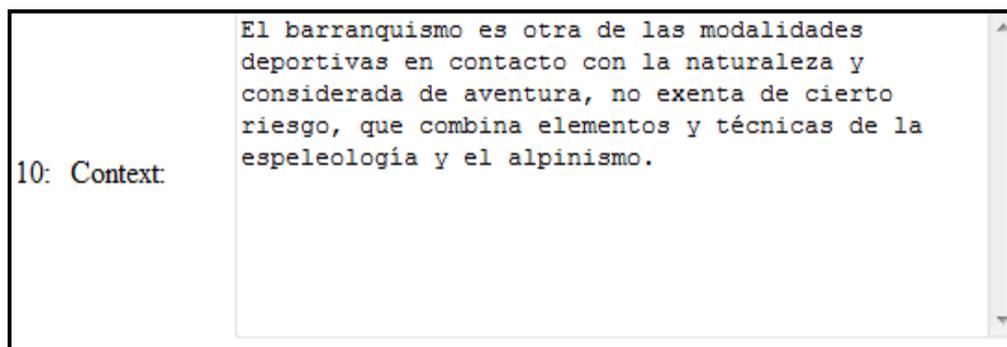


Figura 80. Campo *Context* con la inclusión automática del contexto seleccionado en el módulo *Term concordance*.

Además de utilizar este módulo de concordancias para la búsqueda y selección de contextos, también se ha empleado para detectar las colocaciones más habituales de las unidades terminológicas.⁶²³ Así pues, gracias al uso de este módulo, hemos podido detectar los patrones más comunes de los términos principales, los adjetivos que lo modifican y los verbos que lo acompañan con mayor frecuencia. Como ejemplo, indicaremos algunas colocaciones referentes al término «barranquismo», para continuar con los ejemplos anteriores:

- Curso de barranquismo
- Práctica del **barranquismo**
- Practicar el **barranquismo**
- Barranquismo nocturno
- Técnica de barranquismo

Como se puede observar con estos ejemplos, las colocaciones seleccionadas aportan más información al usuario y le permiten conocer un poco más acerca del uso de la unidad objeto de interés, ya que se trata de combinaciones frecuentes en el campo de especialidad.

El rápido acceso a este módulo de concordancias, junto con la posibilidad de insertar el contexto seleccionado de forma automática en el campo *Context* así como

⁶²³ Utilizamos aquí el término «colocación» en sentido amplio, como sinónimo de co-ocurrencia, esto es, el fenómeno en el que un elemento del discurso aparece combinado a menudo con un determinado término en un campo temático específico.

realizar búsquedas de unidades diferentes a la seleccionada en primer término, facilita y agiliza la consulta de concordancias y la búsqueda de contextos por parte del terminógrafo, así como la selección de colocaciones en las que aparece la unidad de búsqueda.

4.2.2.2. Resultado de entrada terminológica y validación

La aplicación *ProTermino* permite visualizar en el margen derecho de la pantalla de la entrada terminológica la información que se va añadiendo y guardando en la aplicación con el formato que presentará la entrada terminológica al finalizar. Así pues, una vez que se han incluido todos los datos requeridos en los diferentes campos y se han guardado correctamente, aparecerá la entrada terminológica definitiva en el margen derecho de la pantalla de la siguiente manera:

term entry:

Barranquismo (ES) *n m sing (término recomendado)* [Turismo de aventura] Actividad terrestre que consiste en descender a pie a través de barrancos, desfiladeros o cañones que han formado los ríos con un equipamiento específico formado por un traje de neopreno, guantes, escaupines, arnés de seguridad, mosquetón, casco y cuerda. • *El barranquismo es otra de las modalidades deportivas en contacto con la naturaleza y considerada de aventura, no exenta de cierto riesgo, que combina elementos y técnicas de la espeleología y el alpinismo.* □ Descenso de barrancos (*Término relacionado*), Barranquismo (*Forma completa*), Descenso de cañones (*Término relacionado*), ◆ *Practicar barranquismo* ► **Canyoning** *ng* (EN), **Gorge walking** *ng* (EN), **Canyoning** *n* (DE), **Schluchteln** *n* (DE) ▲ *Notas:* Esta actividad se oferta como actividad terrestre o acuática, dependiendo si los barrancos o cañones por los que se discurren llevan agua o no. Como material complementario, es importante llevar bidón estanco, linterna, mochila y chaleco salvavidas, según la cantidad de agua por la que discurra el recorrido. El término "barranquismo" se considera como sinónimo de "descenso de barrancos" pero también como hiperónimo de las diferentes actividades de descenso similares, como son el "descenso de cañones", el "descenso de torrentes", el "descenso de cascadas", etc.

Figura 81. Ejemplo de entrada terminológica para la unidad «barranquismo» según la información guardada.

Esta opción facilita en gran medida la captación de posibles errores relativos a carencias, exceso de información, erratas, etc. y permite, así, una revisión final de la información contenida más rápida y eficiente.

Una vez que se ha llevado a cabo la revisión de todas las entradas terminológicas elaboradas, se procederá a la exportación y edición de esta información con objeto de ponerla a disposición de los destinatarios finales.

4.3. EXPORTACIÓN Y EDICIÓN DEL TRABAJO ONTOTERMINOGRÁFICO

La edición del trabajo ontoterminográfico pasa por la exportación de la información contenida en la base de datos. Para llevar a cabo este paso, *ProTermino* permite varios formatos de exportación según las necesidades de edición posterior de los usuarios, a saber: formato HTML, formato PDF, formato RTF y formato TBX. En este contexto, el terminógrafo deberá determinar cuál de estos cuatro formatos cubre los objetivos de su proyecto. Además, la aplicación le ofrece al terminógrafo la posibilidad de seleccionar los campos terminológicos que desea que aparezcan en las entradas terminológicas, es decir, la entrada terminológica no dispondrá de todos los campos disponibles en la base de datos sino solo de aquellos que el terminógrafo haya considerado pertinentes según las características de su proyecto de trabajo. La selección de estos campos se realiza a través de la tabla de exportación que se encuentra en la pestaña *Export term entry*, del módulo de base de datos

| select fields | |
|---------------------|---|
| terminological unit | <input type="checkbox"/> |
| geographical usage | <input type="checkbox"/> US: United States <input type="checkbox"/> <input type="checkbox"/> GB: United Kingdom <input type="checkbox"/> <input type="checkbox"/> ES: Spain <input type="checkbox"/> <input type="checkbox"/> FR: France <input type="checkbox"/> <input type="checkbox"/> IT: Italy <input type="checkbox"/> <input type="checkbox"/> DE: Germany <input type="checkbox"/> <input type="checkbox"/> IR: Ireland <input type="checkbox"/> |
| part of Speech | <input type="checkbox"/> Noun <input type="checkbox"/> <input type="checkbox"/> Verb <input type="checkbox"/> <input type="checkbox"/> Adjective <input type="checkbox"/> <input type="checkbox"/> Adverb <input type="checkbox"/> <input type="checkbox"/> Other <input type="checkbox"/> |
| gender | <input type="checkbox"/> Masculine <input type="checkbox"/> <input type="checkbox"/> Feminine <input type="checkbox"/> <input type="checkbox"/> Neuter <input type="checkbox"/> <input type="checkbox"/> Other <input type="checkbox"/> <input type="checkbox"/> No gender assigned <input type="checkbox"/> |
| number | <input type="checkbox"/> Singular <input type="checkbox"/> <input type="checkbox"/> Plural <input type="checkbox"/> |
| term status | <input type="checkbox"/> Standardized term <input type="checkbox"/> <input type="checkbox"/> Preferred term <input type="checkbox"/> <input type="checkbox"/> Regulated term <input type="checkbox"/> <input type="checkbox"/> Recommended term <input type="checkbox"/> <input type="checkbox"/> Translator proposal <input type="checkbox"/> |

Figura 82. Parte de la tabla de exportación de campos terminológicos.

Una vez seleccionados los campos adecuados, se debe pulsar el botón de exportación correspondiente al formato elegido y se obtendrá un archivo con la información contenida en la base de datos. A partir de ese momento, se utilizará dicho archivo para la edición del recurso terminológico final, ya sea en formato papel o formato electrónico.

En nuestro proyecto, hemos realizado la exportación de la base de datos con formato TBX (TermBase eXchange),⁶²⁴ con objeto de utilizar dicha información en otra aplicación que nos permita mejorar las posibilidades de búsqueda y consulta en nuestra base de datos. Desde nuestro punto de vista, la exportación en formato HTML, PDF o RTF son poco útiles, puesto que, a pesar de ser fáciles de acceder y de manejar, no facilitan la consulta de la información por parte de los usuarios finales, sino que se basan en una consulta tradicional de búsqueda directa de términos o, incluso, en una búsqueda por orden alfabético. En este sentido, nos parece paradójico que se haya avanzado tanto en la gestión y elaboración de bases de datos, ya sean terminográficas u ontoterminográficas, a la hora de automatizar procesos, extracciones, búsquedas, etc., pero que, una vez que se ha finalizado y exportado la base de datos, se acceda a la información de forma poco eficiente ya sea a través de la búsqueda directa de términos o búsquedas alfabéticas y, por tanto, sin tener en cuenta las necesidades de consulta de los usuarios finales ni los avances tecnológicos que se han realizado hasta el momento en este sentido.

Por este motivo, nos hemos propuesto la elaboración de una herramienta que ofrezca nuevas posibilidades de búsqueda y consulta y, en definitiva, que mejore y agilice las búsquedas realizadas por los usuarios finales, en nuestro caso los traductores. En este sentido, pretendemos elaborar una herramienta que proporcione un avance también en el aspecto de edición y consulta y que, así, nos permita mejorar el acceso a la información desde el punto de vista de rapidez, comodidad y extensión. A continuación, pasaremos a describir esta herramienta nueva, *Trandix*.⁶²⁵

⁶²⁴ Este formato es un estándar de intercambio de terminología y se tratará de forma profunda más adelante (cf. 4.3.1.3.1.1.).

⁶²⁵ La investigación para el desarrollo de esta herramienta y su implementación se ha llevado a cabo conjuntamente entre esta tesis doctoral, que se ha encargado de la parte teórica de la herramienta así como de su evaluación, y el proyecto de fin de carrera de D. A. Fernández Sola, titulado *Aplicación para la*

4.3.1. *Trandix*: herramienta para la búsqueda terminológica en contexto

No dictionary definitions, in fact, provide a totally adequate description of meaning, at least from the point of view of linguistic analysis, because words are dealt with in isolation in dictionaries (Mott, 1996: 122).

A continuación, presentaremos la herramienta *Trandix*, una aplicación informática de ayuda a la traducción que trabaja actualmente con los idiomas español, alemán e inglés y se encuentra actualmente limitada al dominio del turismo de aventura, aunque se podría ampliar a cualquier otro campo de especialidad. Esta herramienta tiene como objetivo facilitar y agilizar la consulta de los usuarios, en concreto los traductores, de la información terminológica que necesitan para realizar sus traducciones. Para ello, sustituye la forma convencional de consulta de recursos terminológicos basada en una búsqueda directa por términos o alfabética por una búsqueda textual, es decir, una búsqueda basada en el texto, con objeto de proporcionar la posibilidad de consultar dicha información en el mismo contexto en el que aparece, el texto original que deben traducir.

Antes de comenzar a describir los objetivos y las funcionalidades de nuestra herramienta, consideramos imprescindible presentar una revisión de las diferentes herramientas anteriores que comparten con *Trandix* algunas características.

4.3.1.1. Antecedentes

A partir de los años 1990 se empezaron a crear aplicaciones denominadas *comprehension assistants* (asistentes para la descodificación) que consideramos aquí como los antecedentes de *Trandix*. Estas herramientas tenían como objetivo facilitar la comprensión de los textos de los usuarios para la descodificación del mensaje en otro idioma, además de agilizar la consulta del significado de las palabras en el mismo texto, es decir, sin necesidad de consultar otras fuentes externas. En principio, aparecieron como una alternativa a la traducción automática, necesaria en ocasiones para lectores

consulta terminológica en el proceso de traducción, dirigida por F. J. Durán Muñoz, del Departamento de Lenguajes y Ciencias de la Computación, de la Universidad de Málaga (fecha prevista de lectura: septiembre de 2011).

sin suficientes conocimientos en otros idiomas, para paliar las carencias y dificultades que presenta a menudo dicho tipo de traducción. Asimismo, como medio de asistencia más flexible y ágil que los diccionarios convencionales.

Al contrario de lo que ocurre con los diccionarios, ya se encuentren estos en formato electrónico o en papel, con los que es necesario realizar la inserción de una palabra en el recurso o seleccionar esta de una lista, estas aplicaciones permitían la consulta de información (equivalentes, principalmente) directamente haciendo clic sobre la palabra seleccionada en el texto donde se encontraba. Principalmente, estas herramientas se han realizado en el seno de proyectos de Lingüística Computacional y del Procesamiento del Lenguaje Natural, por lo que siempre han utilizado recursos ya elaborados como fuentes de información. Como ejemplos de estas herramientas, vamos a citar a continuación las más relevantes, comenzando por el prototipo COMPASS, que fue el primero de la serie de asistentes para la descodificación que se han venido produciendo en la última década.

4.3.1.1.1. Locolex y COMPASS

El sistema *Locolex* ha sido el primer asistente de comprensión que se conoce y fue desarrollado y patentado por el Xerox Research Centre Europe (XRCE),⁶²⁶ en Grenoble (Francia), a mediados de la década de los 1990 (Feldweg y Breidt, 1996). Este sistema es el componente básico del buscador que se incluye en el prototipo *COMPASS* (COMPrehension ASSistant), cuyo principal objetivo es cubrir el hueco que existe entre la traducción automática de un texto completo y la comprensión puntual de un idioma extranjero (Feldweg y Breidt, 1997: 154). Este prototipo se desarrolló en el marco del proyecto *COMPASS*,⁶²⁷ un proyecto financiado por la Comisión de la Unión Europea entre 1994 y 1996 cuyo principal objetivo consistía en la adaptación e integración de diccionarios bilingües en formato papel a sistemas de búsqueda inteligentes con reconocimiento de contextos.

⁶²⁶ <<http://www.xrce.xerox.com/>>

⁶²⁷ <<http://www.sfs.uni-tuebingen.de/Compass/>>

Esta herramienta es un *diccionario de comprensión* bilingüe en línea (*comprehension dictionary*) que ayuda a comprender los documentos electrónicos escritos en una lengua extranjera a través de un acceso *inteligente* a diccionarios electrónicos adaptados especialmente para ser utilizados por esa herramienta. La aplicación *COMPASS* proporciona actualmente ayuda para las combinaciones de inglés a francés y de alemán a inglés a partir de los siguientes diccionarios generales: el *Oxford–Hachette English–French Dictionary* (1994) y el *Collins German–English Dictionary* (1991). Para ser utilizados por *COMPASS*, estos diccionarios deben ser etiquetados conforme al SGML (*Standard Generalized Markup Language*; en español, Lenguaje de Marcado Generalizado) y deben llevarse a cabo una serie de acciones para adaptarlos, revisarlos y reescribir algunas partes, como por ejemplo filtrar información no relevante, reducir el metalenguaje y reforzar el tratamiento de las unidades polilexemáticas (cf. Chanod, 1998: 4-5).

El sistema presupone que un usuario con poco conocimiento de una lengua diferente a la suya pretende leer un texto en dicha lengua. Para ello, podrá utilizar el asistente, que le permitirá realizar una lectura más ágil y contextualizada. El primer paso que necesita realizar el usuario consiste en subir un archivo al lector *COMPASS*, que trabaja con formato HTML, y, una vez hecho esto, podrá seleccionar la palabra del texto que desee para obtener la traducción según el contexto en el que se encuentra dicha palabra. Además, el sistema reconoce si esa palabra forma parte de una unidad polilexemática como compuestos (como «heart attack»), verbos con partícula (como «to nit pick»), locuciones (como «to take the bull by the horns») y paremias (como «Birds of a feather flock together») y seleccionará la traducción apropiada según el contexto sintáctico. De esta manera, en vez de ofrecer como resultado una traducción automática completa del texto en cuestión, que a menudo carece de la calidad suficiente, permite la consulta de las unidades desconocidas de forma rápida y contextualizada.

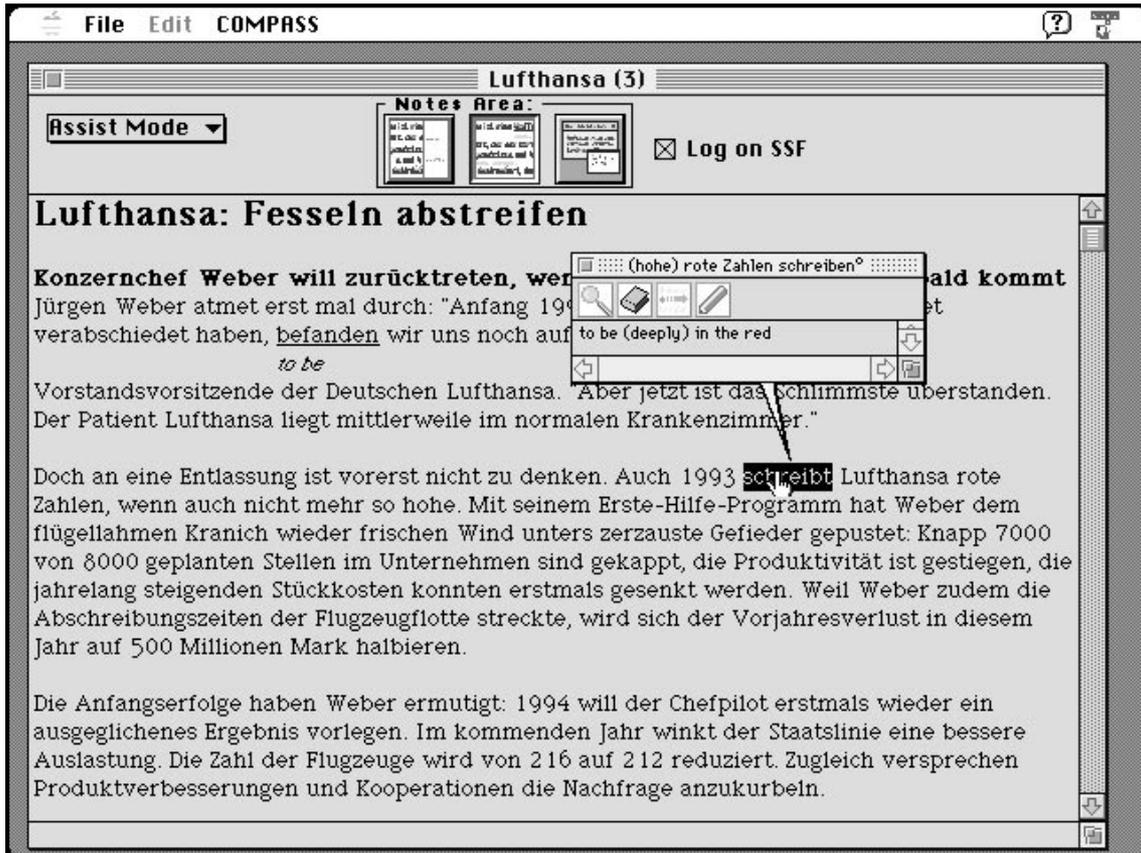


Figura 83. Búsqueda contextual con COMPASS. Fuente: Feldweg y Breidt (1997: 160).

Este sistema utiliza la información sobre los equivalentes recogidos en los diccionarios bilingües generales y toma parte de la acción intelectual de la búsqueda mediante las técnicas de PLN más actuales que, según Feldweg y Breidt (1997: 154), son las siguientes:

- Análisis morfológico de las palabras con inflexión del texto y determinación del lema para la búsqueda,
- Determinación de la categoría gramatical de la palabra seleccionada,
- Acceso a las entradas relevantes del diccionario y almacenamiento de la información completa en una estructura de datos,
- Comprobación de la pertenencia o no de la unidad seleccionada a una unidad polilexemática u otro tipo que requiera de una traducción concreta,
- Desambiguación del sentido de palabras.

En definitiva, este sistema permite a los usuarios la comprensión de un texto en otro idioma de forma superficial cuando no poseen un conocimiento amplio sobre dicho idioma sin necesidad de la traducción automática del texto completo.

A partir de esta aplicación, han ido surgiendo otras herramientas similares, entre las que se encuentran algunas de las que comentamos más brevemente a continuación.

4.3.1.1.2 *MobiMouse Plus*

MobiMouse Plus (Prószéky y Kis, 2002; Prószéky y Földes, 2005) es un asistente de comprensión desarrollado por *Morphologic*⁶²⁸ que se puede adquirir actualmente mediante una licencia comercial. Este asistente es muy similar al prototipo *COMPASS*, que comentábamos anteriormente, tanto en su funcionamiento como en su finalidad. Al igual que el anterior, el objetivo de esta aplicación consiste en facilitar la comprensión de los usuarios con pocos conocimientos en otra lengua a entender un texto cualquiera sin necesidad de recurrir a una traducción automática. En este sentido, esta herramienta será muy útil para cualquier persona que necesite entender el contenido de un mensaje de correo electrónico (o cualquier otro texto electrónico) de forma superficial y de manera rápida, cómoda y exenta de ambigüedades y de información irrelevante.⁶²⁹

De la misma manera que ofrecía *COMPASS*, *MobiMouse Plus* lleva a cabo análisis lingüísticos del contexto en el que se encuentra la palabra seleccionada y utiliza una serie de diccionarios para ofrecer la traducción de dicha palabra con la menor ambigüedad posible. Asimismo, identifica tanto unidades monolexemáticas como polilexemáticas incluso si la palabra seleccionada por el usuario no es la primera de la expresión. La traducción de la palabra seleccionada (o de la expresión a la que pertenece) se muestra, como se puede ver en la Figura 84, en forma de burbuja sobre el contenido que se encuentra en la pantalla en ese momento y desaparece cuando el usuario efectúa un movimiento del ratón.

⁶²⁸ <<http://www.morphologic.hu/en/>>

⁶²⁹ En este sentido, se podría entender como información irrelevante cualquier tipo de información que no ayudara a la decodificación del mensaje sino más bien a su codificación, como por ejemplo información contextual, semántica, notas de uso, etc.

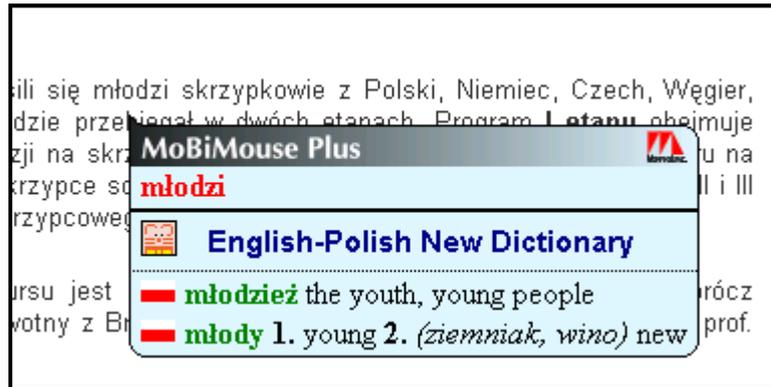


Figura 84. Información que ofrece el asistente, utilizando un diccionario inglés-polaco, al seleccionar una palabra del texto. Fuente: Prószyński y Földes (2005: 2).

A pesar de las similitudes con la aplicación *COMPASS*, *MobiMouse Plus* ofrece algunas pequeñas diferencias que indicamos a continuación:

- No es necesario cargar el texto dentro de la aplicación, sino que esta se instala en el ordenador y permite el reconocimiento de cualquier palabra que aparezca en la pantalla del ordenador.
- No necesita que el usuario realice un clic sobre la palabra, simplemente que este se posicione durante un breve tiempo sobre esta para mostrar la información de la traducción, que desaparece cuando retira el ratón.
- La estructura de los diccionarios que utiliza debe estar en formato XML, ya se encuentren originariamente en formato papel o electrónico.
- La combinación de lenguas con la que trabaja es inglés-polaco.

Además de esta herramienta, *Morphologic* dispone actualmente de otras aplicaciones parecidas a esta, como por ejemplo *MoBiCAT*, una herramienta de traducción automática a través de Internet para la combinación de inglés y polaco. De la misma manera que la anterior, cuando el usuario mueve el cursor sobre el texto en inglés o en polaco, el programa ofrece la traducción automática del texto seleccionado en una burbuja.

4.3.1.1.3. *Smarty*

Al igual que la anterior, el asistente para la descodificación *Smarty* (Arnaudov y Mitkov, 2009) para la combinación inglés-búlgaro está basado en el prototipo *COMPASS* y comparte con este los principales objetivos, a saber: a) facilitar la comprensión del usuario durante la lectura de un texto electrónico en otro idioma y b) reducir al máximo la ambigüedad a través del reconocimiento de contextos; además de algunas de sus funcionalidades, como por ejemplo la necesidad de cargar un texto en la aplicación. Sin embargo, posee ciertas características que lo diferencian, que vamos a comentar a continuación (ibíd.: 3288):

- Se trata de un sistema híbrido con una interfaz más elaborada que sus predecesores que da la posibilidad al usuario de trabajar de forma virtual con dos tipos de diccionarios: uno convencional mejorado, con el que el usuario puede consultar todas las entradas del diccionario y familiarizarse con el significado de la palabra seleccionada, y un asistente de comprensión, a través del cual al usuario se le proporciona la información básica de la palabra para que pueda comprender el significado.
- Utiliza un recurso lexicográfico diferente, *WordNet*, que facilita la desambiguación del sentido de algunas palabras.
- Teniendo en cuenta el tipo de recurso que utiliza y la constante actualización y creación de recursos similares (*EuroWordNet* o *BalkaNet*) se hace más fácil la cobertura del asistente para otras lenguas, al contrario de lo que ocurría con los anteriores asistentes que requerían del etiquetado previo de los diccionarios que empleaban.

Como comentábamos anteriormente, la interfaz de este programa es más compleja que la de los anteriores, puesto que ofrece más cantidad de información (Figura 85).



Figura 85. Pantalla de *Smarty* con información en sus diferentes apartados. Fuente: Arnaudov y Mitkov (2009: 3292).

En esta pantalla se pueden distinguir las tres partes esenciales del programa: en primer lugar, la caja donde se inserta el texto con el que se desea trabajar, que se puede observar en la esquina izquierda; en segundo lugar, se ofrece la traducción (o traducciones) posible según el contexto en el que se encuentra la palabra seleccionada cuando el usuario sitúa el cursor sobre la palabra (se puede ver en la caja que aparece sobre el texto); y, por último, la información adicional que ayuda a ampliar la información sobre la palabra seleccionada (unidades relacionadas, diferentes significados, frases hechas, etc. en los dos idiomas involucrados (inglés y búlgaro) que se recoge en el resto de bloques de la aplicación.

A pesar de estar dirigida también al apoyo de usuarios que tienen problemas a la hora de comprender un texto en otra lengua (búlgaro o inglés), se observa que esta herramienta ofrece más información que las anteriores, por lo que su uso estaría más cerca de las necesidades de los traductores, aunque todavía, como veremos más adelante, dista de completar sus requisitos.

4.3.1.1.4. *Evroterm* y *TERMINATOR*

Por último, comentaremos la aplicación *Terminator*, una herramienta que se ha elaborado recientemente en el marco del proyecto *Evroterm*, financiado por el gobierno

esloveno, y que se dirige, en parte, a traductores, por lo que se aproximaría más que ninguna otra a los objetivos de nuestra herramienta, *Trandix*.

El proyecto *Evroterm*⁶³⁰ tenía como principal objetivo la elaboración de una base de datos terminológica multilingüe de 15 lenguas (aunque la mayoría de los equivalentes se encuentran solo en inglés y esloveno) que sirviera de apoyo a los traductores de esloveno de la Unidad de Traducción de la Oficina del Gobierno para Asuntos Europeos (GOEA) de la República de Eslovenia durante el proceso de traducción de las actas de la Unión Europea, paso previo a su incorporación en la institución. Actualmente, esta base de datos se mantiene desde la Secretaría General del Gobierno de la República de Eslovenia y se actualiza regularmente por medio de la herramienta *Trados MultiTerm*.

Durante el proceso de elaboración de la base de datos, que se denominó también *Evroterm*, se fueron elaborando otras aplicaciones con objeto de elaborar un sistema que fuera eficiente, sencillo y amigable para la consulta terminológica y, así, agilizar y facilitar la labor de los traductores (Željko, 2009). Entre estas aplicaciones, destacamos dos: *Evrokorpus*⁶³¹ y *Terminator*,⁶³² siendo esta última el objeto de nuestro interés. La primera se trata de una colección de textos paralelos extraídos de la base de datos europea Eur-Lex,⁶³³ así como de traducciones de sus propios traductores, que inicialmente estaba compuesto por textos en inglés-esloveno y que, posteriormente, se fue ampliando a otras combinaciones (alemán-esloveno, francés-esloveno, italiano-esloveno y español-esloveno) y que se puede consultar desde la entrada de cualquier término recogido en la base de datos *Evroterm*.

Por su parte, *Terminator* se trata de un analizador terminológico (o lo que se venía considerando hasta ahora como asistente para la descodificación) donde el usuario puede introducir un texto (escrito en inglés, alemán, francés o esloveno) y, tras el análisis que realiza la herramienta, consultar los términos que se encuentran en la base de datos *Evroterm* y que previamente han sido marcados por el programa. Esta

⁶³⁰ <<http://evroterm.gov.si/index.php?jezik=angl>>.

⁶³¹ <<http://evrokorpus.gov.si/k2/index.php?jezik=angl>>

⁶³² <<http://evroterm.gov.si/x/index.php?jezik=angl>>

⁶³³ <<http://eur-lex.europa.eu/es/index.htm>>

aplicación muestra una interfaz muy sencilla e intuitiva, lo que facilita su utilización por parte de los posibles usuarios.

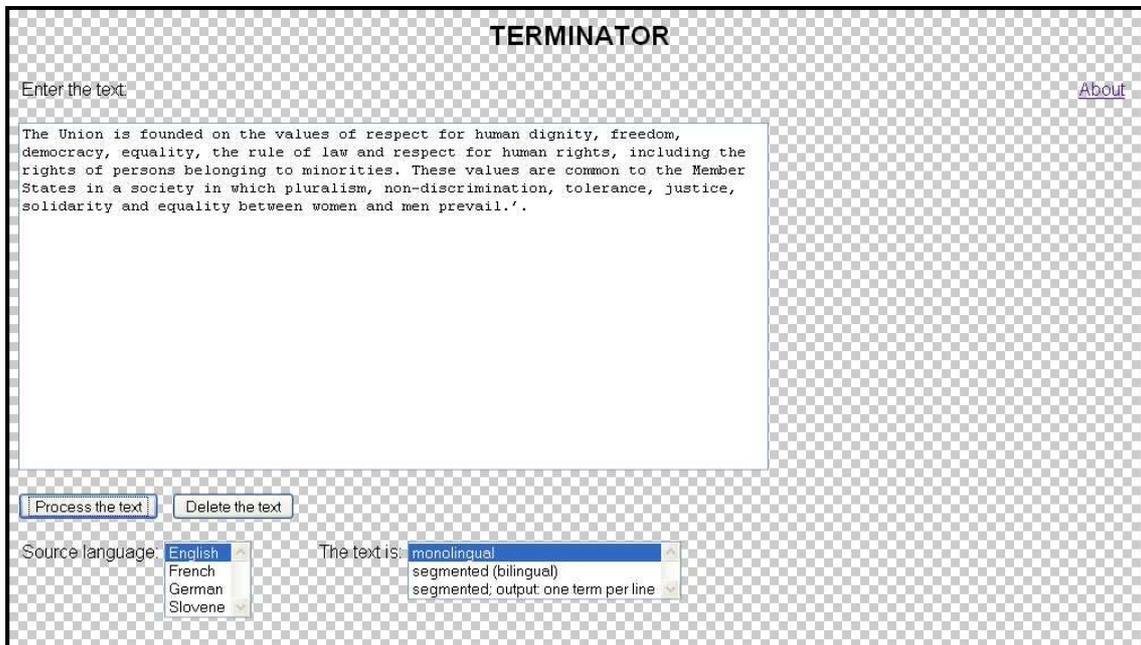


Figura 86. Interfaz de *Terminator*.

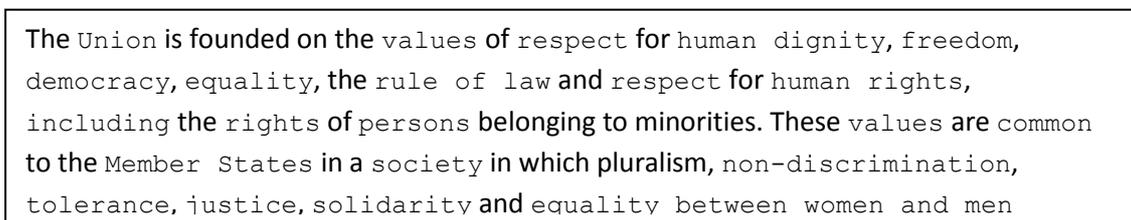


Figura 87. Análisis terminológico de *Terminator*.

Como se puede observar con el resultado obtenido, se marcan gran cantidad de términos en el fragmento del texto introducido, lo que permite al usuario realizar una rápida consulta de estos. Teniendo en cuenta el contexto en el que se ha desarrollado esta herramienta, la traducción de las actas de la Unión Europea para la incorporación de Eslovenia en la institución, cabe destacar que la principal terminología que se recoge en las base de datos está relacionada con la legislación que se ha venido desarrollando en el seno de la Unión Europea y, por este motivo, aparecen tantos términos marcados. No obstante, si se realizan búsquedas de otros dominios (economía, por ejemplo) también el resultado es bastante positivo.

| |
|---|
| <p>CreationDate 03.06.2010 ChangeDate 04.08.2010 Subj informatika Subject informatics AUL Project Microsoft Terminology Collection Entry number 121362</p> <p>EN union  Definition In set theory, the smallest combination of two sets that contains all elements of both sets. See also IATE - PONS - OneLook - Wikipedia - Google define</p> <p>SL unija See also IATE - SSKJ - PONS</p> <p>IT unione See also IATE</p> |
| <p>CreationDate 03.06.2010 ChangeDate 03.06.2010 Subj informatika Subject informatics AUL Project Microsoft Terminology Collection Entry number 118509</p> <p>EN union  Definition In programming, a structure that can be used to store different types of variables (such as integer, character, or Boolean). See also IATE - PONS - OneLook - Wikipedia - Google define</p> <p>IT unione See also IATE</p> |
| <p>EN union  TermRef 32003R2042 Commission Regulation (EC) No 2042/2003 of 20 November 2003 on the continuing airworthiness of aircraft and aeronautical products, parts and appliances on the approval of organisations and personnel involved in these tasks, Appendix I(6.6) See also IATE - PONS - OneLook - Wikipedia - Google define</p> <p>SL cevna spojka TermRef 32003R2042 Uredba Komisije (ES) št. 2042/2003 z dne 20. novembra 2003 o stalni plovnosti zrakoplovov in letalskih proizvodov, delov in naprav ter o potrjevanju organizacij osebj, ki se ukvarjajo s temi nalogami, Dodatek I(6.6) TermSource Ministrstvo za promet See also IATE - PONS</p> |
| <p>CreationDate 09.05.2005 ChangeDate 09.05.2005 Subj pravo Subject law JI Project prevodi slovenskih predpisov Entry number 59655</p> <p>EN union  TermRef Basic Constitutional Charter on the Sovereignty and Independence of the Republic of Slovenia See also IATE - PONS - OneLook - Wikipedia - Google define</p> <p>SL zveza TermRef Temeljna ustavna listina o samostojnosti in neodvisnosti Republike Slovenije See also IATE - SSKJ - PONS</p> |

Figura 88. Resultados de *Terminator* (en *Evroterm*).

En la Figura 88 se observa la consulta realizada para el término «Unión», donde se muestran cuatro posibles dominios y equivalentes al esloveno. Por lo general, la información se recoge únicamente en estos dos idiomas, aunque a veces también se encuentran equivalentes para otros, como son el alemán, el italiano y, en menor medida,

el español. Asimismo, en caso de que el usuario no encuentre la unidad deseada o se requiera más información, la herramienta permite consultar otras fuentes, como la base de datos terminológica *IATE*, diccionarios como *PONS*, recursos como *Wikipedia* o *Google define*, entre otros.

Para terminar, destacamos que este sistema de comprensión es muy útil, especialmente para los textos legislativos en la combinación esloveno-inglés, y que, al contrario que los anteriores asistentes, su principal objetivo consiste en facilitar el trabajo a los traductores y no solo proporcionar una comprensión rápida y superficial del contenido de un texto. Para ello, proporciona más información específica sobre la unidad, como la definición, el contexto en el que se utiliza, así como enlaces a otras fuentes; sin embargo, cabe destacar varios problemas detectados:

En primer lugar, la herramienta no incluye ningún tipo de desambiguación del sentido de la palabra, por lo que, al no especificar la temática sobre la que debe tratar el texto introducido, puede dar lugar a errores de sentido. Por ejemplo, en el caso del término «Unión» ninguna de las cuatro opciones ofrecidas hace referencia a la «Unión Europea», sentido que presenta en el texto.

En segundo lugar, la información que ofrece la base de datos *Evroterm* es muy limitada para los traductores, ya que solo proporciona el contexto, los equivalentes y, en algunos casos, la definición de la unidad, lo que hemos visto a lo largo de este trabajo de investigación que no es suficiente para realizar un proceso de codificación del mensaje original.

Y, por último, muchos de los términos y expresiones polilexemáticas, como «equality between women and men», han sido incluidos en la base de datos por los propios traductores, por lo que no ha habido una revisión previa o posterior por parte de terminólogos, lo que puede dar lugar a equívocos o incoherencias, como por ejemplo en la traducción de «World Heritage Site» que aparece recogida como «sitio incluido en el patrimonio universal».

4.3.1.2. Objetivos

A modo de resumen de todo lo anterior, podemos establecer como objetivos básicos de las herramientas que hemos considerado como antecedentes a *Trandix* los tres siguientes puntos:

- 1) Facilitar la descodificación de un mensaje electrónico en otro idioma por un usuario con conocimientos insuficientes en dicho idioma.
- 2) Agilizar y flexibilizar la consulta del significado de palabras conforme se está leyendo el mensaje de interés.
- 3) Establecer la desambiguación de unidades por el contexto en el que se encuentran.

En definitiva, avanzar hacia la elaboración de unos recursos más accesibles y flexibles para los usuarios de textos electrónicos y evitar la pérdida de tiempo de estos en la consulta de recursos externos al texto durante la lectura.

Partiendo de estos objetivos, y teniendo en cuenta las evaluaciones positivas que han presentado estas herramientas antecesoras, nos planteamos la necesidad de elaborar una aplicación informática similar a las anteriores en el marco de nuestra investigación de recursos terminológicos dirigidos a traductores, ya que presentaban gran cantidad de ventajas de consulta, especialmente la agilidad, flexibilidad y rapidez. Sin embargo, era necesario adaptar algunos de sus aspectos y solventar algunas debilidades que presentaban, como por ejemplo:

- 1) Modificar el grupo potencial de destinatarios de las herramientas.⁶³⁴ En nuestra aplicación, los usuarios finales son traductores; por tanto, esta herramienta debía proporcionar la información requerida por este grupo para satisfacer y cubrir sus necesidades y requisitos de consulta durante la descodificación del mensaje original y la codificación del mensaje meta (cf. 2.5.3.).
- 2) Facilitar la utilización de la información de consulta. La mayoría de las aplicaciones, a excepción de *Smarty* y de *Terminator*, necesitaban la

⁶³⁴ Recordemos que todas las herramientas presentadas, a excepción de *Terminator* iban dirigidas a un público general con necesidades de comprensión en otro idioma.

adaptación de los recursos que empleaban en sus procesos mediante un cambio de formato y etiquetado de las fuentes utilizadas, lo que requería gran esfuerzo por parte de los desarrolladores y evitaba contar con la actualización necesaria del contenido.

- 3) Mejorar la calidad de la información contenida en estas herramientas. En su mayoría, estas herramientas utilizan fuentes fiables, aunque no siempre actualizadas o completas y, a veces, el contenido no es completamente acertado, como es el caso de *Terminator*, lo que disminuye su eficacia. Así pues, es necesario que la información que se consulte sea una información terminológica actual y fiable elaborada según una metodología apropiada y por expertos en terminología.

De esta manera, nos proponemos elaborar una herramienta que siga los objetivos de sus antecesores y que se adapte tanto a las necesidades de sus usuarios como a las técnicas y metodologías actuales. Para ello, se han establecido una serie de objetivos que se pretenden cumplir con el desarrollo de *Trandix*.

- 1) Dar accesibilidad a un recurso terminológico en español, inglés y alemán sobre el dominio del turismo de aventura.
- 2) Facilitar la descodificación del mensaje original y la codificación del mensaje meta, según los requisitos y necesidades de los traductores y con una información terminológica elaborada según la metodología propuesta en el capítulo 2.
- 3) Agilizar la consulta de la información terminológica mediante la consulta en el propio texto, sin necesidad de abrir o buscar otros recursos externos.
- 4) Permitir al usuario seleccionar la información terminológica que desea que se muestre de todos los campos incluidos en la base de datos, según los requisitos indicados por los traductores profesionales (cf. 2.5.3.), en el apartado de *Preferencias* del menú de la herramienta, así como el idioma de la interfaz, entre otras posibilidades.
- 5) Fomentar el intercambio de información terminológica entre diferentes aplicaciones y usuarios con el empleo del formato TBX (TermBase

eXchange), específicamente la variante TBX-Basic (cf. 4.3.1.3.1.1.), recomendada por los organismos internacionales ISO y LISA para el contexto de traducción y localización, que contiene la base de datos terminológica creada previamente en la aplicación informática *ProTermino*.

- 6) Proporcionar la posibilidad de ampliar, actualizar o modificar el contenido terminológico por parte de los terminólogos de forma rápida y sencilla a través de actualizaciones, ampliaciones o modificaciones del archivo TBX.
- 7) Completar o subsanar posibles carencias de la información contenida en el archivo TBX a través de la consulta directa a otros recursos, como son diccionarios o bases de datos especializadas en el dominio de interés (en este caso, el turismo de aventura), buscadores generales (como *Google*, *Bing*, etc.), diccionarios de abreviaturas y acrónimos, entre otros.⁶³⁵ Estos recursos se podrán seleccionar en el apartado de *Preferencias* del menú de la herramienta.
- 8) Permitir la consulta de información terminológica mediante la búsqueda directa de términos en la base de datos sin necesidad de cargar ningún texto, es decir, permitir que la herramienta funcione como una base de datos convencional.
- 9) Proporcionar la posibilidad de *feedback* por parte de los usuarios que han realizado búsquedas externas para ir mejorando y completando la base de datos terminológica.

Con estos objetivos fijados desde un primer momento se comenzó a trabajar en la elaboración de la herramienta *Trandix*, de la que, a continuación, presentamos la descripción de sus funcionalidades y de sus principales características.⁶³⁶

⁶³⁵ Esta parte determina los recursos terminológicos que se pueden consultar a través de la herramienta mediante las denominadas búsquedas externas. De esta manera, es esta parte la que delimita el campo de especialidad al que está dirigido *Trandix*, en este caso el turismo de aventura. A excepción de esta parte, que recoge únicamente recursos relacionados con el campo de especialidad determinado, la herramienta estaría completamente desvinculada de cualquier dominio, ya que solo requeriría de un archivo TBX con la información terminológica en su contenido.

⁶³⁶ Para profundizar en el aspecto más técnico de esta herramienta, véase Fernández Sola (2011), donde se abordan los aspectos relativos a la implementación de *Trandix*.

4.3.1.3. Descripción

Como indicábamos con anterioridad, *Trandix* es una aplicación informática que pretende asistir al traductor durante el proceso de descodificación y codificación del mensaje con el que está trabajando y mejorar la consulta de la información terminológica que necesite a través de una forma rápida y cómoda en el mismo texto que contiene el mensaje en cuestión. Además, esta aplicación permite cargar archivos TBX actualizados y sin límite de tamaño exportados de una base de datos terminológica de cualquier tipo de especialidad. Para ello, el funcionamiento de la aplicación consta de una serie de pasos, que exponemos brevemente a continuación:

- 10) El usuario (en este caso, el traductor) carga el texto original con el que va a trabajar en la aplicación, así como un archivo con una base de datos terminológica elaborado por sí mismo o puesto a su disposición en formato TBX.
- 11) La aplicación procesa automáticamente el texto y el archivo y le ofrece al traductor el texto subido con los términos marcados que aparecen en la base de datos.
- 12) El traductor pincha en el hiperenlace de cualquier término marcado y accede a la entrada terminológica del término seleccionado.

Además de estos tres pasos básicos, el usuario podrá realizar búsquedas externas para consultar información sobre términos que no estén incluidos en la base de datos o para ampliar la información encontrada en otros diccionarios especializados, buscadores generales, entre otras fuentes. Por último, el usuario también podrá realizar consultas a la base de datos sin necesidad de encontrar el término en el texto que está utilizando o sin necesidad de cargar un texto, es decir, tendrá la posibilidad de utilizar la herramienta como una base de datos convencional.

Una vez que hemos descrito de forma general el funcionamiento de la herramienta *Trandix*, continuaremos con la presentación de las diferentes partes y de sus diferentes funcionalidades.

4.3.1.3.1. Descripción de funcionalidades

Se trata de una aplicación multiplataforma, compatible con Windows, Linux y MacOS y desarrollada en el lenguaje de programación Java. Aunque se estudió la posibilidad de hacer de esta aplicación una herramienta con acceso en línea mediante el uso de *applets* o *servlets*, desestimamos dicha posibilidad debido a la gran cantidad de datos que habría que soportar en el servidor, lo que conllevaría una ralentización de la herramienta. Por este motivo, se trata de una aplicación ejecutable en el ordenador de cada usuario.

Uno de los principales objetivos que nos planteamos en un primer momento fue que se tratara de una aplicación intuitiva y fácil de manejar y que, por tanto, no causara problemas al usuario a la hora de aprender a manejarla y de consultar la información. En este contexto, la interfaz gráfica que presenta *Trandix* en todo momento es muy sencilla y amigable.



Figura 89. Ventana inicial de *Trandix*.

Esta figura nos muestra la interfaz gráfica de la pantalla inicial de la aplicación y, como se puede observar, es muy clara y sencilla. En esta pantalla, se nos permite realizar el primer paso requerido, es decir, nos permite seleccionar los dos ficheros con los que se van a trabajar, esto es: por un lado, el archivo de texto en formato plano (.txt) que contiene el texto original sobre el que se va a trabajar, y, por el otro, el archivo TBX, que incluye la información terminológica que se ha exportado previamente de un sistema de gestión terminológico (SGT) u ontoterminológico (SGOT). En este paso inicial también se puede seleccionar el idioma en el que se encuentra escrito el texto

original que, por defecto, será considerado español; es decir, según el texto con el que se desea trabajar se encuentre en español, inglés o alemán, el usuario deberá seleccionar el idioma correspondiente en este paso.⁶³⁷

Además de ser una herramienta fácil de usar, esta aplicación es robusta, ya puede manejar ficheros extensos tanto de textos en formato plano (.txt) como de información ontoterminográfica en formato TBX. Asimismo, gracias al empleo de archivos TBX, el formato estándar recomendado por la organización internacional ISO (cf. 4.3.1.3.1.1.), se consigue la interoperabilidad de la aplicación, ya que puede interactuar con otras aplicaciones similares ya existentes aprovechando información o archivos generados por dichas herramientas.

Una vez que se han seleccionado los archivos correspondientes y el idioma del texto origen, pinchamos en el botón *Aceptar* y pasamos a la siguiente pantalla, donde ya se puede visualizar el texto origen con los términos seleccionados. Asimismo, en caso de que el usuario no disponga de un texto en formato plano (.txt), la aplicación permite el procesamiento de cualquier texto mediante la acción de copiar y pegar en la caja habilitada para ello, por lo que en este primer paso solo sería necesaria la selección del archivo en formato TBX. En la siguiente pantalla, la pantalla principal y de consulta de la aplicación, sigue siendo igual de sencilla y amigable que la anterior (cf. Figura 89) y es donde llevamos a cabo el siguiente paso de funcionamiento de la herramienta.

⁶³⁷ En caso de que, posteriormente, se desee modificar el idioma elegido es posible realizar el cambio en el menú *Archivo* que aparece disponible en la pantalla siguiente.



Figura 91. Opción Buscar en la ventana principal de *Trandix*.

Por su parte, en la parte superior de la ventana, vemos que está disponible un menú *Archivo*, donde se le permite al usuario modificar los archivos cargados inicialmente en la aplicación, seleccionar archivos utilizados en las sesiones anteriores, procesar de nuevo el análisis del texto (en caso de que se hayan realizado modificaciones en el texto cargado), así como cambiar el idioma del texto seleccionado en el primer paso, en caso de que el nuevo texto con el que desee trabajar se encuentre en otro idioma distinto del seleccionado previamente (cf. Figura 89). Por último, este menú también ofrece la opción de *Salir* de la aplicación, que funciona de la misma manera que si cerráramos la ventana con el botón de Cerrar de la esquina derecha de la ventana.

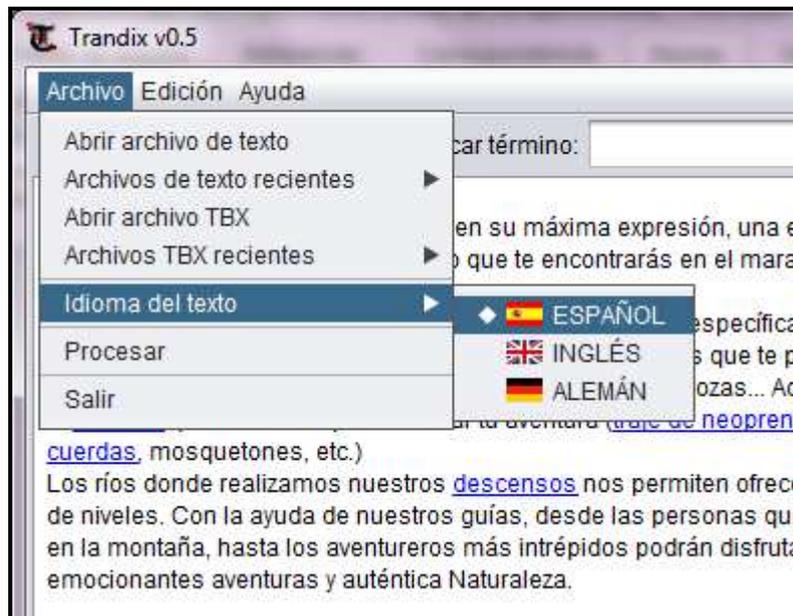


Figura 92. Menú *Archivo* de la ventana principal de *Trandix*.

Junto al menú *Archivo*, se encuentra el menú *Edición* (cf. Figura 93), que permite al usuario establecer preferencias de búsqueda (cf. Figura 95), además de ofrecer al usuario la posibilidad de realizar búsquedas convencionales a través de ese menú (cf. Figura 101) y realizar acciones como copiar, pegar y cortar en la caja de edición del texto.

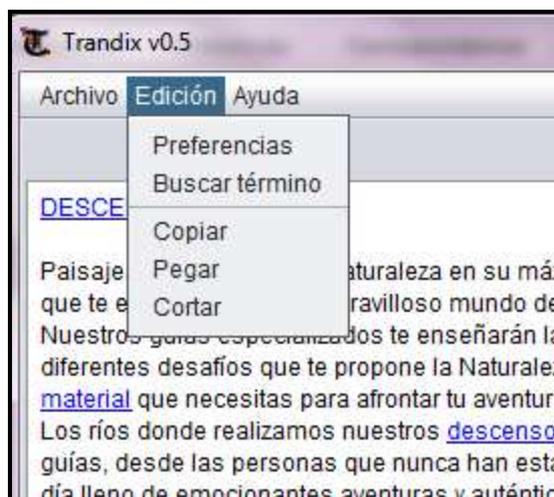


Figura 93. Menú *Edición* de la ventana principal de *Trandix*.

Por último, el menú *Ayuda* (cf. Figura 94) aporta información sobre el origen, la versión y la actualización de la herramienta en la opción *Acerca de*, así como el manual de usuario de la aplicación, con objeto de explicarles a los usuarios todas las funcionalidades, posibilidades que presenta y preguntas frecuentes y optimizar, así, el uso que realizan de la herramienta.

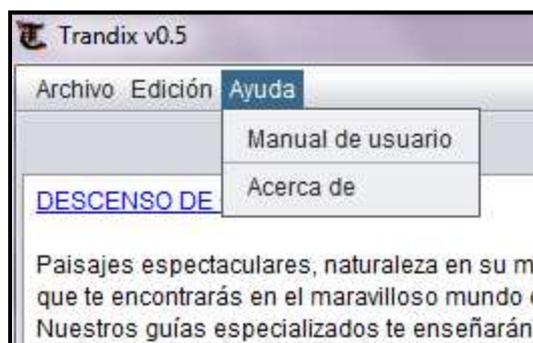


Figura 94. Menú *Ayuda* de la ventana principal de *Trandix*.

Las preferencias de búsqueda, a las que hacíamos referencia anteriormente, permiten al usuario establecer una serie de selecciones relativas a la aplicación y que permiten personalizar algunos aspectos, como son el idioma de la interfaz, los campos que pueden aparecer en la entrada terminológica de los términos que seleccione, así como los recursos externos que desee consultar en caso de que no encuentre la información deseada en nuestra herramienta o desee ampliar dicha información. De esta manera, es recomendable que el usuario establezca sus preferencias antes de realizar cualquier búsqueda en *Trandix*. La ventana de preferencias es, de nuevo, sencilla e intuitiva, como podemos ver a continuación.

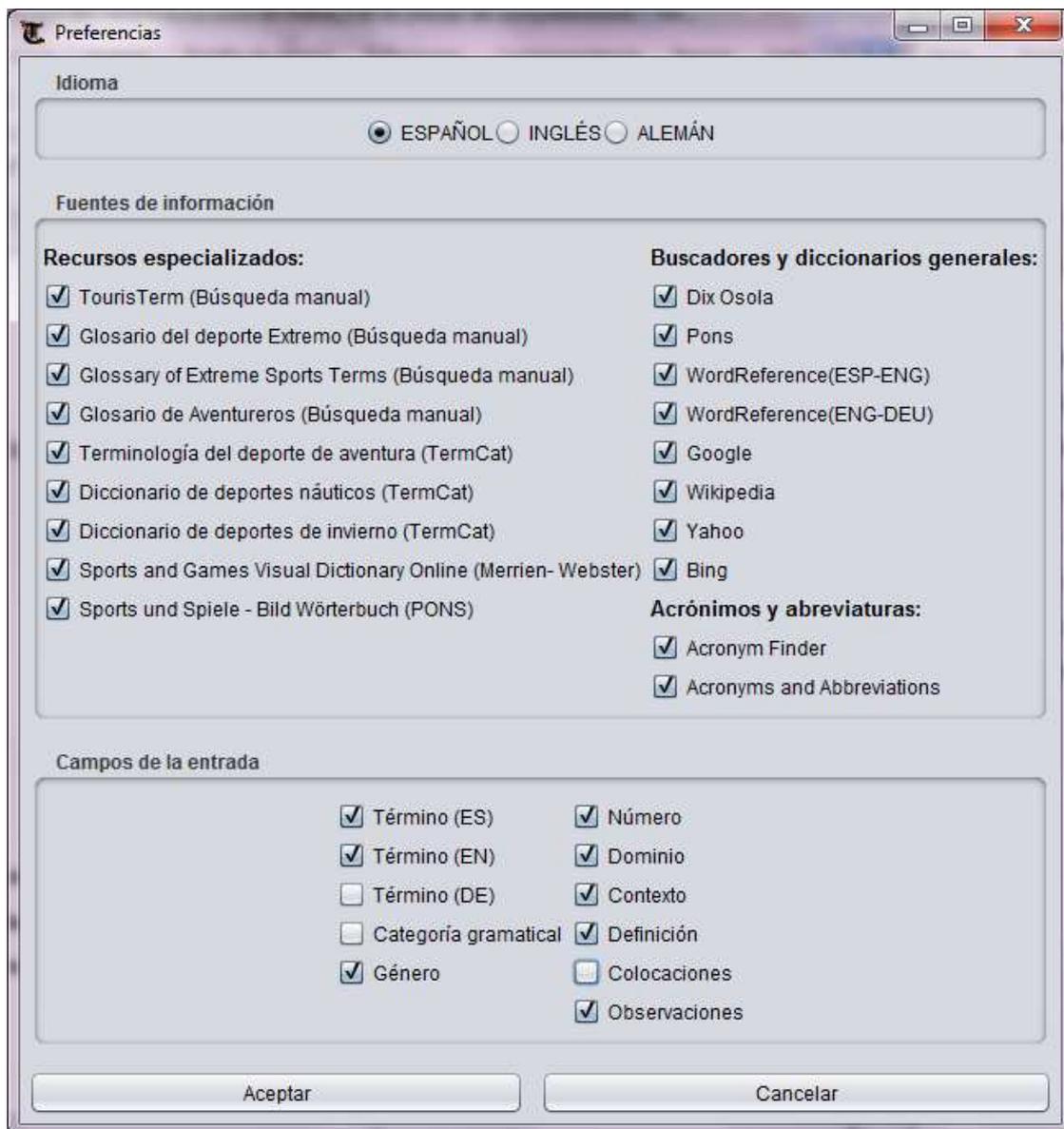


Figura 95. Ventana de *Preferencias*, en el menú *Edición*.

En la primera parte de esta ventana, encontramos una referencia a los tres idiomas con los que trabaja la herramienta, a saber: español, inglés y alemán. En este caso, se le permite al usuario seleccionar el idioma que desea utilizar para la interfaz de toda la herramienta, es decir, el idioma en el que aparecerán las diferentes opciones y campos de la aplicación. De esta manera, el usuario podrá seleccionar, por un lado, el idioma del texto con el que va a trabajar y, por otro, el idioma que desea para la interfaz de la aplicación.

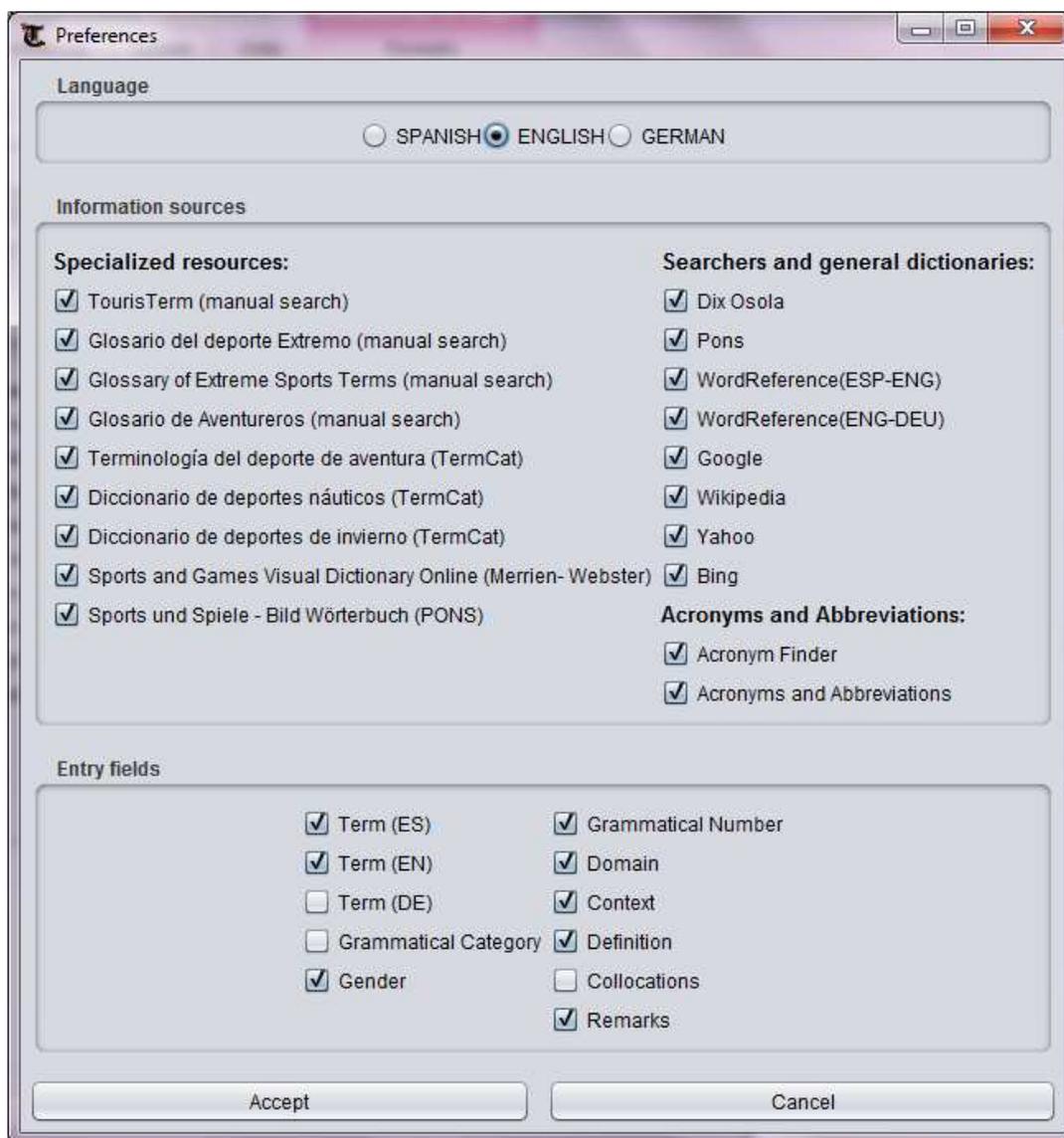


Figura 96. Ventana de *Preferencias* en inglés.

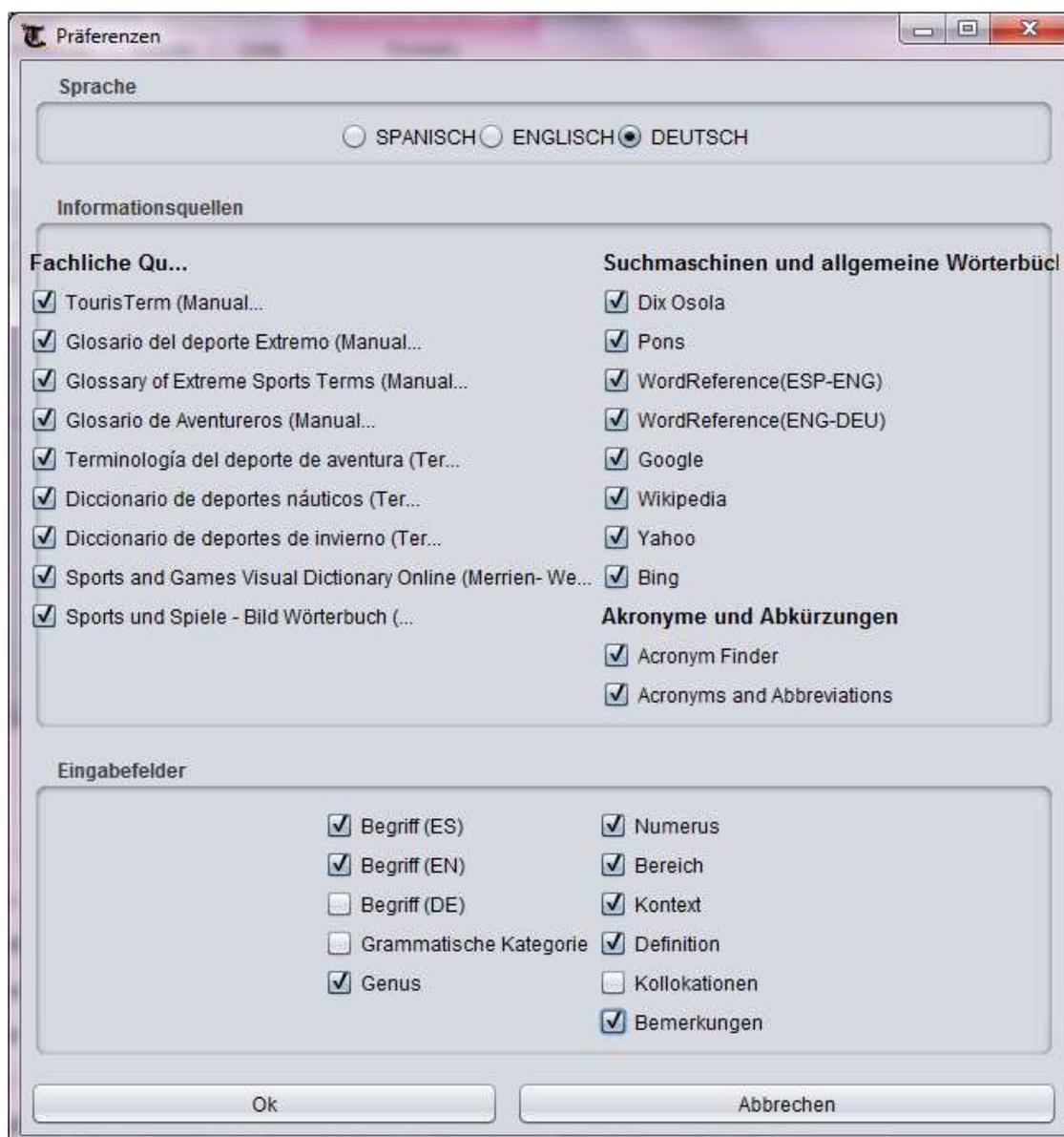


Figura 97. Ventana de *Preferencias* en alemán.

Debajo de la opción para seleccionar el idioma, se sitúan las diferentes fuentes de información que puede consultar con posterioridad el usuario en caso de necesidad. Aquí, hemos establecido una clasificación según la naturaleza de los recursos: en primer lugar, encontramos los recursos especializados, es decir, diccionarios y bases de datos especializadas en el turismo de aventura; en segundo lugar, hemos incluido recursos generales, como buscadores y algunos diccionarios generales; y, por último, hemos incorporado también dos recursos de abreviaturas y acrónimos, que a menudo son necesarios en los textos especializados. De nuevo, el usuario tendrá la posibilidad de

seleccionar los recursos que considere más adecuados para su trabajo y con los que realice posteriormente las búsquedas externas.

Por último, en esta misma ventana, se observan todos los campos que conformarían una entrada terminológica adecuada según las necesidades y requisitos establecidos por los traductores profesionales (cf. 2.5.3.). No obstante, consideramos que la flexibilidad es uno de los aspectos más relevantes a la hora de realizar consultas en recursos terminológicos, ya que no se pueden generalizar las necesidades de todos los traductores ni de todos los campos de especialidad o momentos en los que se encuentra la traducción. Por este motivo, proporcionamos la oportunidad de que sea el propio usuario el que seleccione la información que desea visualizar posteriormente en la consulta de los términos. Así pues, un usuario podría seleccionar los campos definición, equivalentes y contexto, mientras que otro podría indicar otros campos diferentes, como definición y observaciones, o simplemente equivalentes de traducción. A modo de ejemplo, se puede observar en la figura anterior referente a la pantalla de preferencia cómo se han seleccionado una serie de campos determinados, a saber: término en inglés y en español, género, número, dominio, definición, contexto y observaciones. De esta manera, serán estos campos los que aparezcan en la entrada terminológica de los términos cuando se realicen las búsquedas.

Una vez que se han determinado todas las posibilidades que proporciona la aplicación, el usuario ya puede comenzar a realizar consultas. Así pues, cuando el usuario pinche dos veces sobre un término marcado con hipervínculo en el texto, le aparecerá en la parte derecha de la ventana (que permanecía vacía hasta este momento) la entrada terminológica del término que haya seleccionado conforme a los campos que haya indicado previamente en la ventana de *Preferencias*.

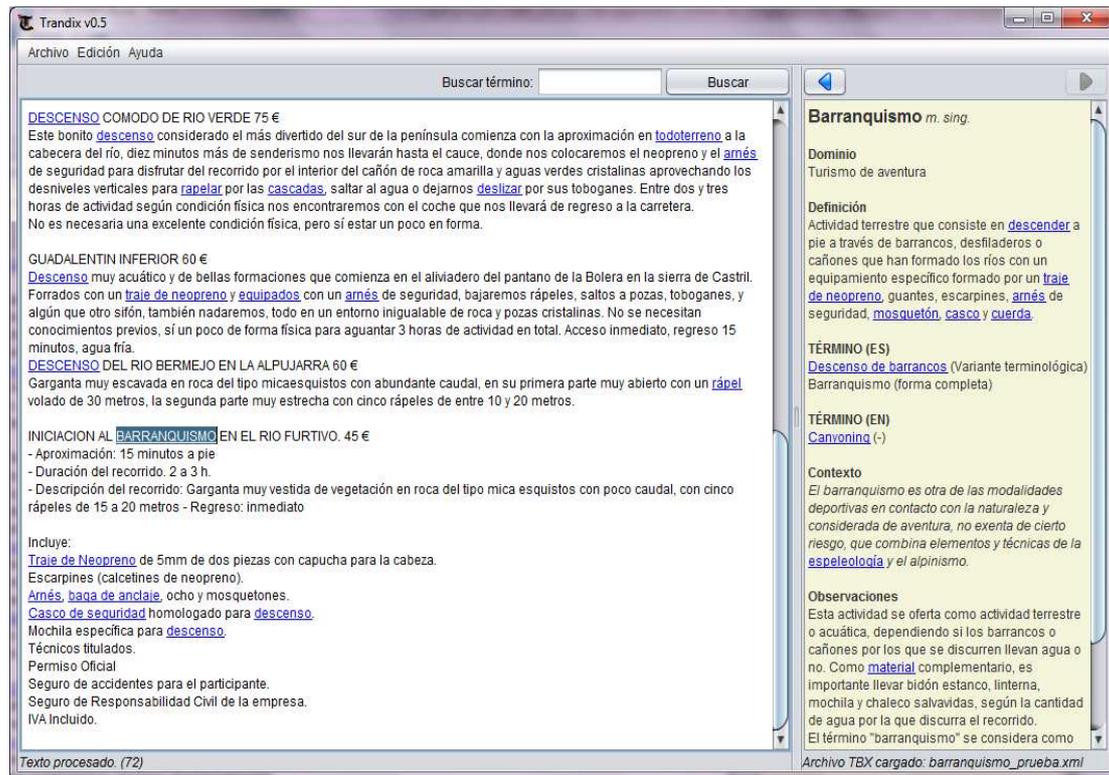


Figura 98. Ventana principal de *Trandix* con el texto cargado (a la izquierda) y la entrada terminológica del término seleccionado (a la derecha).

En nuestro caso, al haber seleccionado el término «Barranquismo» hemos obtenido en la ventana de la derecha la entrada terminológica referente a dicho término y según los campos que habíamos seleccionado previamente, a saber: en primer lugar, el término; en segundo lugar, la información gramatical (género y número); en tercer lugar, el dominio; en cuarto lugar, la definición; en quinto lugar, el contexto; y, en sexto lugar, las observaciones sobre la entrada.

Llegados a este punto, cabe destacar que el orden establecido para mostrar la información terminológica no es arbitrario, sino que sigue las indicaciones de la norma ISO 16642 (2003). Esta norma establece una división de la entrada terminológica en tres niveles básicos y con un orden determinado, a saber: en primer lugar, un nivel independiente de la lengua, que contiene información relacionada con el concepto, como por ejemplo el dominio al que pertenece el concepto; en segundo lugar, un nivel lingüístico, formado por campos como definición y observaciones; y, por último, el nivel referente al término, que incluye información como sinónimos, abreviaturas, términos relacionados, así como usos regionales, contextos, etc.

También es importante destacar los hipervínculos entre los diferentes términos recopilados en la base de datos a través de las entradas terminológicas. Se trata de otra forma más de flexibilizar la consulta de información y de facilitar la búsqueda de información a los usuarios de la aplicación. Asimismo, otra facilidad que aporta esta herramienta consiste en la posibilidad de consultar rápidamente términos polilexemáticos, como «descenso de barrancos», sin necesidad de consultar los términos de forma separada o de encontrar dificultades y dudas sobre la forma de buscar este tipo de términos. Es decir, la aplicación *Trandix* da acceso a los conceptos del dominio del turismo de aventura, ya sean estos monolexemáticos o polilexemáticos, de la misma manera, ya que recoge la información terminológica como si fueran una unidad indisoluble.

Como decíamos con anterioridad, otra opción para mejorar y facilitar las búsquedas de los usuarios consiste en las búsquedas externas. Estas búsquedas permiten al usuario realizar consultas en recursos externos a la aplicación, ya sea porque no encuentran la información que buscan o porque desean ampliar la ya existente. Para ello, primero deben seleccionar los recursos que desean en la opción *Preferencias*, del menú *Edición*, como explicamos arriba, y realizar dos pasos muy sencillos: primero, seleccionar en el texto de la derecha el término que desean consultar; y segundo, pinchar en el botón derecho del ratón y seleccionar el recurso en el que desean realizar una búsqueda externa. En el momento en que se pinche en el recurso elegido, se abrirá automáticamente una ventana del navegador predefinido con el resultado de la búsqueda.

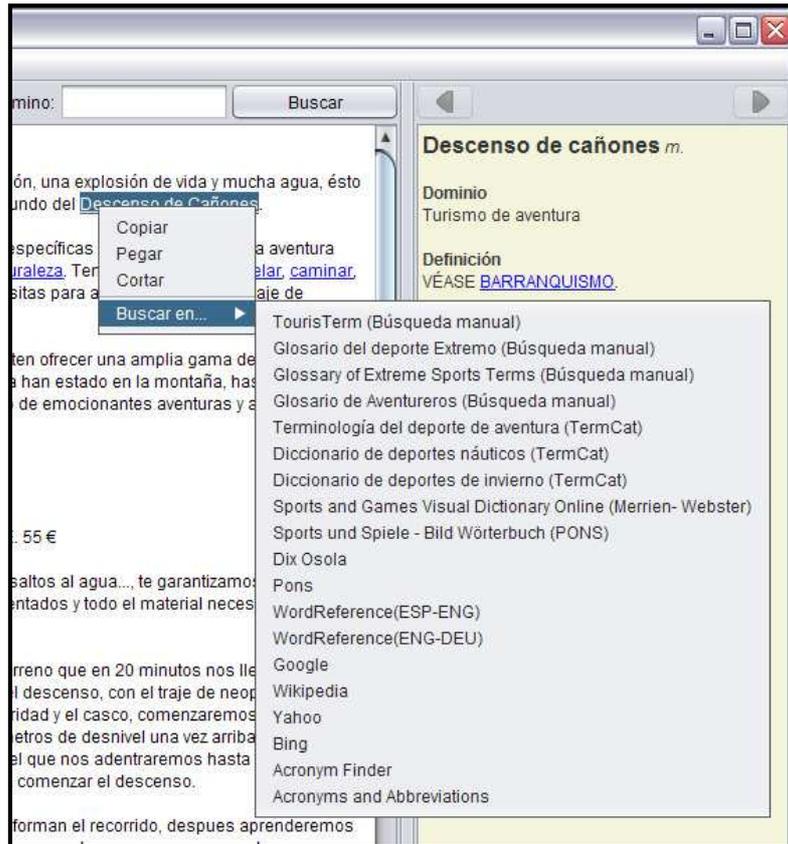


Figura 99. Proceso para realizar búsquedas externas.

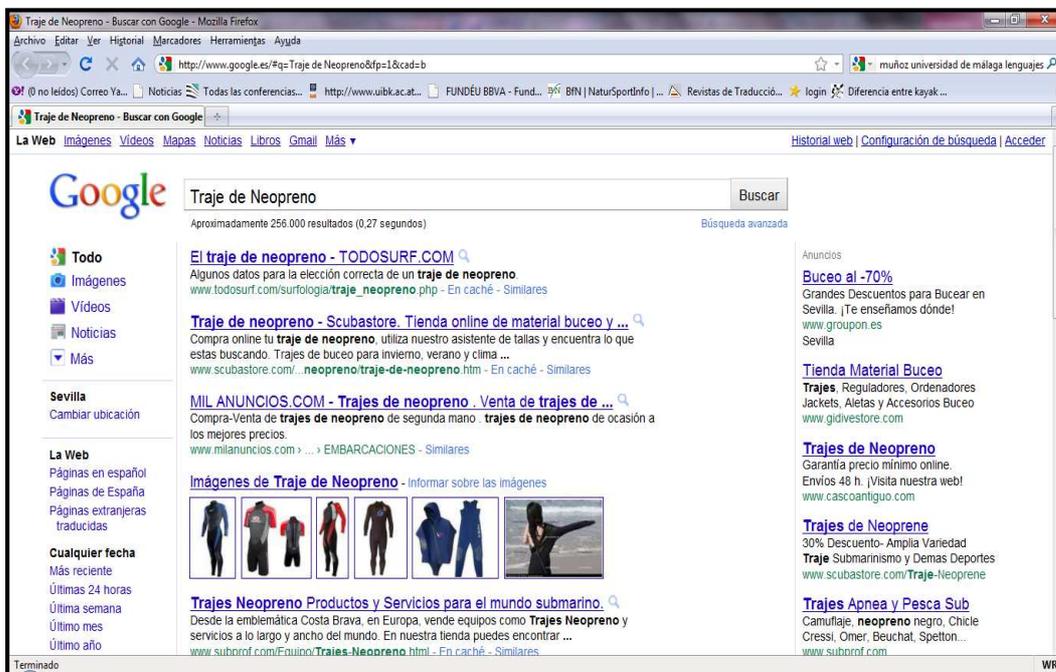


Figura 100. Resultado obtenido tras una búsqueda externa en el buscador *Google* con el término «traje de neopreno».

Siguiendo con los tipos de búsquedas que se pueden realizar, ya hemos visto anteriormente que la aplicación permite la búsqueda convencional, es decir, realizar búsquedas mediante la introducción directa de términos en un buscador interno de la base de datos cargada mediante un archivo TBX en la misma pantalla principal. No obstante, esta opción también está disponible a través del menú *Edición*, seleccionando la opción *Buscar término*. Una vez que el usuario ha seleccionado esta opción, aparece una ventana que permite la introducción del término que se desea buscar, así como los términos buscados recientemente.



Figura 101. Ventana para búsquedas convencionales.

La funcionalidad de búsqueda que ofrece esta opción y la que se encuentra en la pantalla principal es exactamente la misma, ya que las dos permiten la búsqueda directa de términos, pero la primera ofrece además una lista desplegable con las búsquedas más recientes realizadas. De esta manera, se le da la opción al usuario, de nuevo, de poder elegir si desea un acceso rápido a través de la pantalla principal o un acceso al registro histórico de las búsquedas realizadas anteriormente a través de la opción de buscar en el menú *Edición*.

Por último, es importante destacar no solo el interés de la aplicación informática por satisfacer las necesidades de los usuarios (en este caso, traductores) a la hora de realizar sus consultas, tanto en lo que respecta al contenido como en la forma de consultar, sino que también requiere de *feedback* para llevar a cabo una constante mejora y actualización de la información que ofrece. Para ello, establece un sistema de comunicación con el usuario al finalizar el uso de la aplicación. Este sistema consiste en lo siguiente: al cerrar el programa aparece un mensaje donde se le pregunta al usuario si desea enviar información sobre las búsquedas externas que ha realizado con objeto de mejorar la información que aparece en la base de datos.

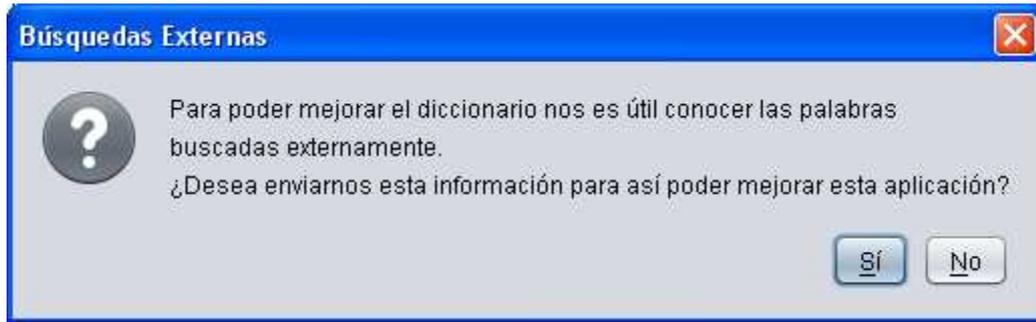


Figura 102. Ventana con el mensaje de petición de colaboración

En caso de que el usuario seleccione que sí desea colaborar con el envío de dicha información, se abriría automáticamente el programa de envío de correo electrónico del usuario con la dirección del envío, el asunto y los términos que ha buscado de forma externa.

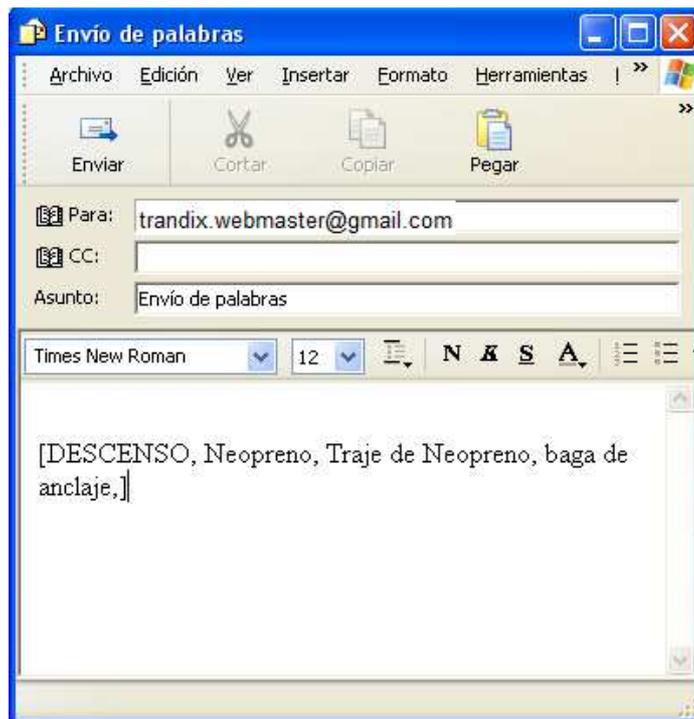


Figura 103. Ejemplo de mensaje de correo electrónico con la información buscada de forma externa.

Una vez que hemos descrito detalladamente las funcionalidades de la herramienta, pasaremos a comentar algunos de sus aspectos técnicos más relevantes y que han sido esenciales para su correcto funcionamiento, como son las razones por las

que hemos utilizado un archivo en formato TBX como archivo de exportación-importación para nuestra aplicación, así como su estructuración básica, y la vinculación de dicho archivo con el archivo de texto.

4.3.1.3.1.1. Formato TBX (Term Base eXchange)

Uno de los aspectos más importantes a los que se le está prestando especial atención la terminología moderna y, con ella, los principales organismos de estandarización internacionales (ISO, TEI, LISA⁶³⁸) consiste en la necesidad de intercambiar y reutilizar información terminológica. Existe tal volumen de trabajos terminológicos, de sistemas de gestión terminológica con formatos de exportación propios, etc. que se está produciendo un crecimiento exponencial de módulos de conversión entre diferentes formatos que, a menudo, resultan insuficientes y provocan un aumento del trabajo por parte del usuario, ya sea éste un traductor, un terminólogo o una empresa.⁶³⁹

Otro concepto relevante que ha de tenerse en cuenta al hablar del intercambio de datos terminológicos es, sin lugar a dudas, la pérdida de información. En este sentido, debemos considerar un concepto fundamental en las tecnologías de la información y, concretamente, en las bases de datos: el *modelo de datos* (o esquema conceptual), que hace referencia al tipo de datos que hay en la base y la forma en que estos se relacionan entre sí. Como mencionamos en el capítulo anterior, la mayoría de los sistemas de gestión de bases de datos terminológicos (SGT) modernos permiten a los usuarios la estructuración de un modelo de datos individualizado para cada una de sus bases de datos (por ejemplo, *Multiterm 5*, *TschwaneLex*, entre otros), lo que provoca gran disparidad de modelos de datos y, por tanto, dificulta la transmisión de datos de unos

⁶³⁸ Recientemente, se ha publicado que la asociación LISA se ha disuelto por problemas financieros. Este es el anuncio que podemos leer en su página web: «As of 2011 February 28, the Localization Industry Standards Association (LISA) was declared insolvent. The results of the Extraordinary General Assembly meeting and ballot of 2011 April 21 LISA has formalized the dissolution of the association and the transfer of the LISA Standards to the public domain». URL: <<http://www.lisa.org/>>.

⁶³⁹ Como ejemplo, consúltese esta página web: <http://producthelp.sdl.com/SDL%20MultiTerm%202009/client_en/SDL_MultiTerm_Convert/About_MultiTerm_Convert.htm>, donde se presenta el conversor *SDL MultiTerm Convert*.

sistemas a otros y de unas aplicaciones a otras (cf. 3.2.1.). De esta manera, a la hora de intercambiar información terminológica almacenada en dos bases de datos distintas pueden darse tres situaciones bien diferenciadas, como nos indica Pérez Hernández (2002):

- Tanto el sistema gestor de bases terminológicas A como el B se basan en un modelo de datos único, aunque distintos el uno del otro. En este caso, es factible aunque costoso que los fabricantes A y B creen sus propios algoritmos de conversión desde sus respectivos productos al otro. La pérdida de información se producirá únicamente en aquellos casos en los que el modelo de datos difiera en gran medida.
- El sistema A se basa en un modelo de datos único, mientras que el sistema B se basa en un modelo de datos definible por el usuario. En este caso, el fabricante del sistema A puede construir un modelo de datos sobre la marcha, que genere durante la conversión un esquema conceptual en formato legible por B idéntico al original, que, posteriormente, el usuario del sistema B podrá modificar. Así, la pérdida de información será nula o despreciable, dando por supuesta una flexibilidad real en cuanto a la definición del modelo de datos, así como la posibilidad de la generación de modelos de datos *off-line* por parte del sistema B. En la otra dirección, no obstante, la conversión de datos no es posible, ya que el fabricante del sistema B, aunque conoce el modelo de datos de A, no puede prever cuál va a ser el definido por el usuario de su sistema. Únicamente sería factible realizar un análisis del modelo creado por el usuario y mostrar un índice de pérdida de información al convertir los datos al formato de B, que el usuario podría aceptar o rechazar.
- Ambos sistemas se basan en un modelo de datos definible por el usuario. En este caso cualquier intento de conversión entre formatos sería prácticamente imposible, debido al elevado volumen de pérdida de información.

Estas posibilidades se producen únicamente entre dos sistemas de gestión terminológica, pero el problema crece exponencialmente al aumentar el número de sistemas de gestión disponibles. Ante esta situación se pueden dar dos posibilidades para garantizar en cierta medida el intercambio de información entre dos sistemas:

- Suponiendo que antes del comienzo del proyecto se tenga la intención de compartir la información resultante, las dos partes (A y B) deberán determinar de antemano la utilización de un mismo modelo de datos, es decir, deberán establecer los tipos de campos, las relaciones, etc. que se va a incluir en la base de datos. Sin embargo, esto no sería suficiente, ya que, al emplear los distintos SGT diferentes formatos de exportación y de estructuración interna, se hace imprescindible que ambas partes empleen el mismo SGT. A esta posibilidad se le denomina «intercambio negociado» (ibíd.).
- En caso de no acordar los puntos iniciales básicos (modelo de datos y SGT), sería necesario que ambas partes utilizaran un formato de intercambio *ciego* de datos terminológicos. El intercambio es «ciego» (Hardman, 2000; Melby, 1998: 4), puesto que ninguna de las dos partes puede conocer el proceso de trabajo de la otra, es decir, ninguna de las partes conoce el SGT al que serán importados los datos ni del que se exportan.

La creación de un formato que permita este último tipo de intercambio de información terminológica, es decir, el establecimiento de una *interlingua* (Pérez Hernández, 2002), entre los diferentes formatos, a fin de garantizar la reutilización y el intercambio de datos terminológicos en formato electrónico ha sido durante algunos años el objetivo del Comité Técnico n.º 37 de la ISO, así como de otras organizaciones como LISA (Localization Industry Standards Association). Este trabajo ha dado como fruto la norma ISO 30042: *Systems to manage terminology, knowledge and content. TermBase eXchange (TBX)* (2008), donde se especifica la *interlingua* que se debería utilizar para el intercambio ciego de información terminológica: el formato TBX (*TermBase eXchange*).⁶⁴⁰

El formato TBX (*Term Base eXchange*) es un formato estándar basado en XML que fue desarrollado por LISA (*Localization Industry Standards Association*) en 2002 y que está basado en las categorías de datos y restricciones especificadas en la norma ISO 12620 (1999). Desde su creación en 2002, esta organización, junto con la ISO, ha venido trabajando en la concreción de su estructura, etiquetas, especificaciones, etc. hasta la publicación de la norma ISO 30042 (2008). Como decíamos anteriormente, la

⁶⁴⁰ <<http://www.lisa.org/Term-Base-eXchange.32.0.html>>

utilización de TBX se basa en la idea de que, debido a la variedad de colecciones de datos terminológicos y escenarios de usos, ningún lenguaje de marcado terminológico (TML) podría satisfacer todos los requisitos de los usuarios ni la interoperabilidad necesaria entre los diferentes SGT y posibles aplicaciones. De esta manera, el principal objetivo de este formato estándar es permitir el intercambio de datos terminológicos, además de facilitar la difusión, el análisis y la representación de estos. A continuación, presentamos las diferentes aplicaciones relativas a sus objetivos anteriores (LISA, 2008: 5):

- Intercambio, para permitir:
 - la transferencia de datos terminológicos entre tecnologías y sistemas,
 - la integración de datos terminológicos desde diversas fuentes, y
 - la conversión de datos debidos a un cambio en las aplicaciones o tecnologías.
- Divulgación, incluyendo:
 - la consulta de múltiples bases de datos terminológicos mediante un único interfaz de usuario,
 - el permiso de descarga de datos terminológicos de un sitio web para los usuarios interesados,
 - el permiso para la consulta de entradas en proceso de revisión para recibir *feedback*, y
 - la puesta a disposición de los usuarios de forma dinámica terminología en aplicaciones en red.
- Análisis y representación, incluyendo:
 - la comparación de contenidos de diferentes bases de datos terminológicos,
 - el estudio de la pérdida de información entre diferentes SGT durante el proceso de conversión, y
 - el diseño de una nueva base de datos terminológicos para minimizar la pérdida durante las conversiones.

Con respecto a su estructura, el TBX, al igual que cualquier lenguaje de marcado, tiene un número de etiquetas posibles para su correcta utilización. A continuación, presentamos las principales (ibíd.: 10):

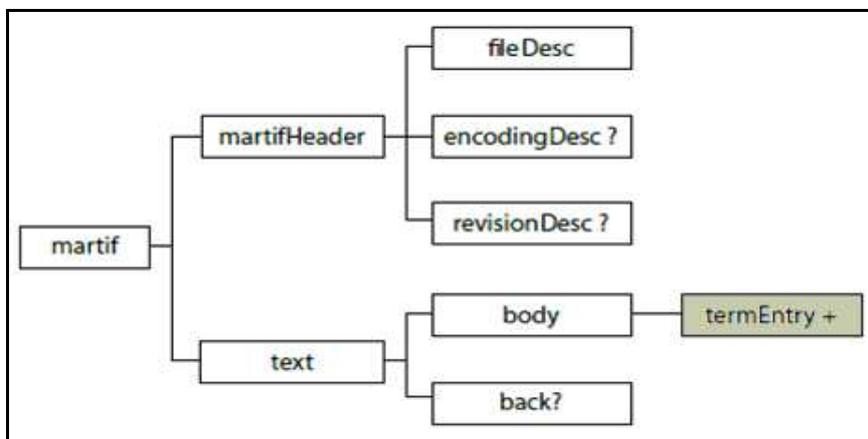


Figura 104. Jerarquía de etiquetas en TBX.

Donde la etiqueta <martif> es el elemento de nivel más alto de un archivo TBX, que está formado por la etiqueta <martifHeader> y la etiqueta <text>, los cuales están tomados de la norma ISO 12200 (1994) y tienen sus orígenes en las normas de la *Text Encoding Initiative*.⁶⁴¹ Por su parte, la etiqueta <text> consiste en las entradas terminológicas, que se encuentran delimitadas por la etiqueta <body> y por información complementaria, que se encuentra en la etiqueta <back>. Y, por último, la etiqueta <martifHeader> corresponde a información general del metamodelo del TMF (Terminology Markup Framework) y está formada por una descripción de la colección completa de datos terminológicos (en la etiqueta <fileDesc>), por información sobre el archivo XCS aplicable (en la etiqueta <encodingDesc>) y una historia de las principales revisiones de la colección (en la etiqueta <revisionDesc>).

De estas etiquetas, la más interesante desde nuestro punto de vista es la correspondiente a la información contenida en <termEntry>, que es la que contiene la información relativa a las diferentes entradas terminológicas:

⁶⁴¹ <<http://www.tei-c.org/index.xml>>

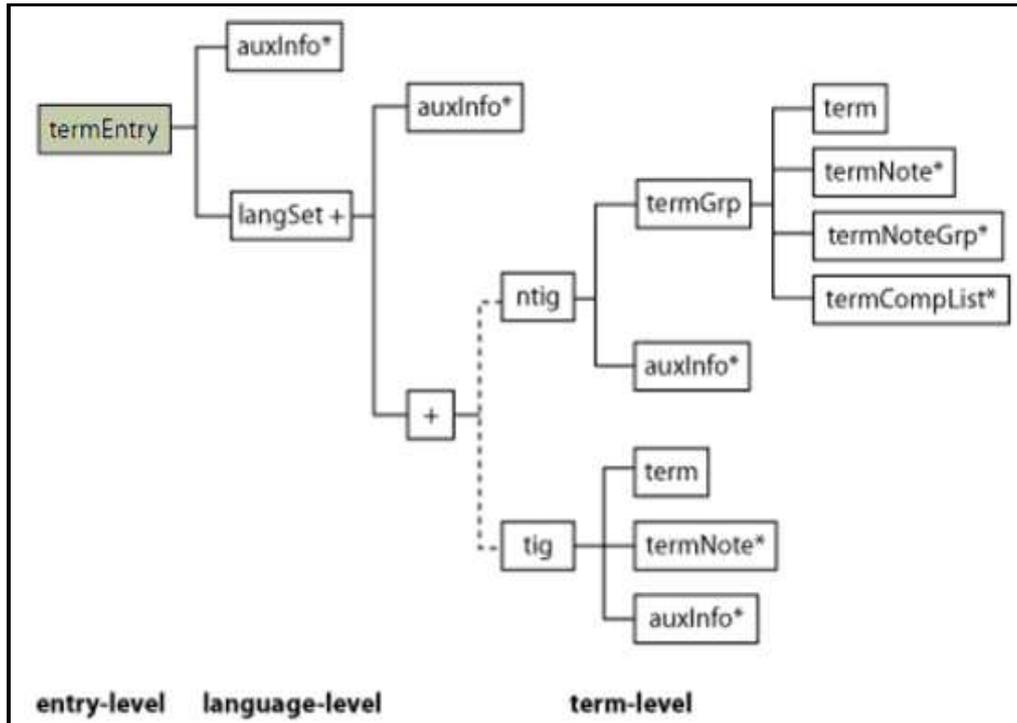


Figura 104. Jerarquía de etiquetas en una entrada terminológica de TBX.

Esta estructura está formada por los conceptos incluidos dentro de la etiqueta <body> (y </body>) y cada uno se denomina <termEntry>. Todos los TBX comparten una misma estructura central, en la que se representa cada uno de los tres niveles estructurales: nivel conceptual (<termEntry>), el nivel de idioma (<langSet>) y el nivel de término (<ntig>, o su versión simplificada <tig>). Como se puede observar en la Figura 104, se trata de una estructura jerárquica donde las entradas de conceptos contienen entradas de lenguas, que, a su vez, contienen entradas para los términos individuales, es decir, va pasando de un nivel a otro de forma jerárquica. La estructura central del archivo también proporciona un conjunto de elementos genéricos para añadir información descriptiva y administrativa de dichas entradas, como son la etiqueta <auxInfo>, que incluye cualquier información que pueda relacionarse con cualquiera de los tres niveles; las etiquetas <termNote> y <termNoteGrp>, que pueden aparecer únicamente en el nivel de término e inferior, y la etiqueta <termCompList>, que está formada por cualquiera de las etiquetas que pueden aparecer en diferentes niveles de la

entrada seguidas por una o más etiquetas del tipo <termComp> o <termCompGrp>. ⁶⁴² Estos elementos genéricos se pueden utilizar en los diferentes lenguajes de marcado terminológico (TML, Terminological Markup Language).

```

<?xml version="1.0" encoding="UTF-8" ?>
- <martif type="TBX" xml:lang="en-US">
- <martifHeader>
- <fileDesc>
- <titleStmt>
    <title>Microsoft Terminology Collection
Export</title>
    </titleStmt>
- <sourceDesc>
    <p>Microsoft Terminology Collection</p>
    </sourceDesc>
  </fileDesc>
</martifHeader>
- <text>
- <body>
  - <termEntry id="2_76365">
    - <langSet xml:lang="en-US">
      - <descripGrp>
        <descrip type="definition">To terminate a session with a
        computer accessed through a communications line usually a computer
        that is both distant and open to many users.</descrip>
      </descripGrp>
      - <ntig>
        - <termGrp>
          <term id="76365">log off</term>
        </termGrp>
      </ntig>
    </langSet>
  - <langSet xml:lang="es-es">
    - <ntig>
      - <termGrp>
        <term id="76367">cerrar sesión</term>
      </termGrp>
    </ntig>
  </langSet>
</termEntry>
</body>
</text>
</martif>

```

Figura 105. Ejemplo de etiquetas TBX. Fuente: *Colección de terminología de Microsoft*. URL: <<http://www.microsoft.com/Language/es-es/Terminology.aspx>>

⁶⁴² Para profundizar en el significado y utilización de estas y otras etiquetas del formato TBX, así como de sus restricciones, véase LISA (2008).

A pesar de los requisitos y restricciones a las que se debe prestar atención para la creación correcta de cualquier archivo en formato TBX, cabe destacar que se trata de un formato flexible y modular, es decir, es un formato que permite la elección de campos (concretamente, categorías de datos terminológicos, según la norma ISO 12620:1999) para que el usuario adapte el resultado a sus necesidades. De esta manera, posibilita la “personalización” del archivo final dentro de unos estándares establecidos. No obstante, para maximizar la interoperabilidad de los datos terminológicos reales, el TBX proporciona también un conjunto por defecto de las categorías de datos más utilizadas en las bases de datos terminológicas, que se denomina *TBX-default TML* según el estándar internacional. Asimismo, existe otro tipo de TBX más reducido que incluye las categorías de datos más comunes utilizadas en el trabajo de traducción y localización y para cualquier aplicación que requiera un manejo más ligero de la terminología, que se denomina *TBX-Basic* (LISA, 2008; Melby, 2008).⁶⁴³ En otras palabras, el principal objetivo de este formato TBX-Basic consiste en formalizar las necesidades de la industria de la traducción y de la localización para el mercado terminológico en un estándar XML, a fin de aumentar las posibilidades de intercambio de recursos terminológicos entre usuarios, además de permitir el uso de estos recursos en diferentes entornos informáticos. Según nuestras propias necesidades, y de acuerdo a lo expuesto con anterioridad, hemos considerado que esta variante del formato XML es la más apropiada para nuestra aplicación informática, *Trandix*.

Como ejemplo de este formato, presentamos a continuación un fragmento del archivo TBX-Basic que hemos confeccionado con nuestro trabajo ontoterminográfico sobre el turismo de aventura, donde se pueden observar las etiquetas de apertura y cierre del archivo, así como un ejemplo de los campos seleccionados para la información terminológica monolingüe de la entrada «barranquismo».

⁶⁴³ <<http://www.lisa.org/TBX-Basic.926.0.html>>

```

<?xml version="1.0" encoding="ISO-8859-1" ?>
- <martif type="TBX-Basic" xml:lang="en-US">
- <martifHeader>
- <fileDesc>
- <sourceDesc>
  <p>Exported from terminology workbench</p>
</sourceDesc>
</fileDesc>
- <encodingDesc>
  <p
type="XCSURI">http://www.lisa.org/fileadmin/standards/tbx_basic/TBxB
asicXCSV02.xcs</p>
  </encodingDesc>
</martifHeader>
- <text>
- <body>
- <termEntry id="214644">
- <descripGrp>
  <descrip type="definition">Actividad terrestre que consiste en
descender a pie a través de barrancos, desfiladeros o cañones que
han formado los ríos con un equipamiento específico formado por un
traje de neopreno, guantes, escarpines, arnés de seguridad,
mosquetón, casco y cuerda.</descrip>
  </descripGrp>
  <descrip type="context">El barranquismo es otra de las
modalidades deportivas en contacto con la naturaleza y considerada
de aventura, no exenta de cierto riesgo, que combina elementos y
técnicas de la espeleología y el alpinismo.</descrip>
  <descrip type="subjectField">Turismo de aventura</descrip>
- <langSet xml:lang="ES">
- <tig>
  <term>Barranquismo</term>
  <termNote type="termType">forma completa</termNote>
  <termNote type="grammaticalnumber">sing.</termNote>
  <termNote type="grammaticalgender">m.</termNote>
  <termNote type="administrativeStatus">término
preferente</termNote>
  <termNote type="linguisticremark">Esta actividad se oferta como
actividad terrestre o acuática, dependiendo si los barrancos o
cañones por los que se discurren llevan agua o no. Como material
complementario, es importante llevar bidón estanco, linterna,
mochila y chaleco salvavidas, según la cantidad de agua por la que
discurra el recorrido. El término "barranquismo" se considera como
sinónimo de "descenso de barrancos" pero también como hiperónimo de
las diferentes actividades de descenso similares, como son el
"descenso de cañones", el "descenso de torrentes", el "descenso de
cascadas", etc.</termNote>
  <termNote type="geographicalUsage">ES</termNote>
  </tig>
</langSet>
</termEntry>
</body>
</text>
</martif>

```

.Figura 106. Ejemplo de etiquetas TBX-Basic exportadas de *ProTermino*.

4.3.1.3.1.2. Vinculación entre archivos

La vinculación entre los archivos de texto con los que trabaja el usuario y los archivos en formato TBX, donde se carga la base de datos terminológica, requiere de un procesamiento previo de lematización, ya que la forma en la que se encuentra la información que contienen uno y otro archivo no concuerda siempre. Como ejemplo, podemos indicar que es necesario lematizar formas plurales («rápeles» por «rápel»), femeninos («escaladora» por la forma neutra «escalador»), formas personales de verbos («equipados» por el infinitivo «equipar»), entre otros casos. De esta manera, conseguimos que las formas lematizadas que aparecen como entradas en la base de datos terminológica (en formato TBX) se vinculen a las formas en uso del texto de trabajo.

Para llevar a cabo este preprocesamiento, se emplea el lematizador *TreeTagger*,⁶⁴⁴ una aplicación informática de libre distribución desarrollada por Helmut Schmid en el proyecto TC (*Textual Corpora and tools for their exploration*), en el Instituto para la Lingüística Computacional de la Universidad de Stuttgart. Esta herramienta ha sido utilizada con éxito en numerosos proyectos de PLN y los resultados obtenidos han sido muy satisfactorios. Además, permite etiquetar textos en gran variedad de lenguas (alemán, inglés, francés, italiano, holandés, español, búlgaro, ruso, griego, portugués, chino, swahili y francés antiguo), así como adaptar su uso a otros idiomas si se dispone de un léxico y un corpus de entrenamiento etiquetado manualmente. En nuestro caso, no será necesario este último paso, puesto que nuestras lenguas de trabajo (español, alemán e inglés) están incluidas en el paquete de la aplicación. Por todo lo anterior, decidimos desde un primer momento utilizar esta herramienta y no otras que se encuentran disponibles actualmente, tanto de libre distribución como de carácter comercial.

En nuestra aplicación informática, el procesamiento realizado por este lematizador comienza de forma automática en el momento en el que el usuario selecciona los dos archivos de trabajo (el archivo en texto plano y en formato TBX) y pincha en el botón *Aceptar* de la pantalla inicial.

⁶⁴⁴ <<http://www.ims.uni-stuttgart.de/projekte/complex/TreeTagger/DecisionTreeTagger.html>>

| | | |
|------------------|-------|-----------|
| Río | NP | río |
| Verde | NP | verde |
| es | VSfin | ser |
| conocido | VLadj | conocer |
| en | PREP | en |
| toda | QU | todo |
| Andalucía | NP | Andalucía |
| por | PREP | por |
| la | ART | el |
| espectacularidad | NC | <unknown> |
| de | PREP | de |
| sus | PPO | suyo |
| formas | NC | forma |
| saltos | NC | salto |
| zonas | NC | zona |
| de | PREP | de |
| nado | VLfin | nadar |
| toboganes | NC | <unknown> |
| y | CC | y |
| rápeles | ADJ | <unknown> |
| acuáticos | ADJ | acuático |
| que | CQUE | que |
| se | SE | se |
| sucedan | VLfin | suceder |
| sin | PREP | sin |
| descanso | NC | descanso |

Tabla 51. Ejemplo de texto etiquetado con *TreeTagger*.

Como se observa en la tabla anterior, el resultado de la aplicación consta de tres columnas bien diferenciadas: en la columna de la izquierda aparece la ocurrencia del texto; en la columna central se encuentra la categoría gramatical que presenta dicha ocurrencia (PREP = preposición, ADJ = adjetivo, ART = artículo, etc.); y, por último, en la columna de la derecha se muestra el lema de la palabra inicial, es decir, la forma neutra de la ocurrencia del texto. De esta manera, observamos, por ejemplo, cómo el verbo «es» se convierte en su forma impersonal en infinitivo «ser». Cabe destacar aquí dos puntos importantes: por un lado, el procesamiento realizado por el lematizador no se hace visible al usuario en ningún momento, puesto que no es pertinente para el uso de la herramienta; y, por otro lado, dicho procesamiento se realiza de forma rápida y eficiente y, por tanto, no retrasa la obtención de los resultados de búsqueda.

Una vez finalizado este preprocesamiento, nuestra aplicación informática, *Trandix*, utilizará únicamente la forma lematizada de las ocurrencias del texto, es decir, las formas que se recogen en la columna de la derecha. Este paso nos permite vincular perfectamente y de forma eficiente los dos archivos de trabajo que, de otra manera, carecerían de una coherencia formal.

A pesar de las ventajas obtenidas con esta fase de lematización a la hora de establecer vínculos entre las diferentes informaciones contenidas en los dos archivos de trabajo, debemos admitir que hemos encontrados algunos problemas derivados del empleo de este lematizador y que consideramos relevante recogerlos en este apartado. Principalmente, estos problemas vienen derivados de la falta de reconocimiento que presenta el lematizador *TreeTagger* en algunas ocasiones, especialmente con el idioma español.⁶⁴⁵ Por ejemplo, en la tabla anterior (Tabla 51) se observa que algunos términos no han sido reconocidos por el lematizador y este da como resultado <unknown>, lo que dificulta la conexión con el contenido del archivo TBX y obliga a la aplicación a realizar una serie de modificaciones internas.

Hemos procedido a realizar una clasificación de los problemas encontrados en tres categorías diferentes, a saber: en primer lugar, problemas ortográficos; en segundo lugar, problemas morfológicos; y, en tercer lugar, problemas terminológicos, que presentamos a continuación.

En el primer grupo, los problemas ortográficos, están relacionados principalmente con el reconocimiento de «ñ» y de tildes. Así pues, cuando un término está en mayúsculas y contiene una «ñ» no se considera como el mismo término en minúscula, por lo que la vinculación con la base de datos terminológica fracasa en este caso. Con respecto a las tildes, ocurre lo siguiente: si una unidad aparece en el texto de trabajo sin tilde (por algún error ajeno a la aplicación), la aplicación no lo lematiza correctamente y, por tanto, da como resultado el mismo término o, como anteriormente, <unknown>. De nuevo, provoca una desvinculación con la base de datos, que incluso puede incluir el término correspondiente con su forma correcta. En nuestro caso, esto ha

⁶⁴⁵ Principalmente se encuentran problemas en *TreeTagger* con el idioma español, puesto que el léxico y las reglas gramaticales con las que cuenta son menos específicas y extensas que otras lenguas como, por ejemplo, el inglés y el alemán, que son lenguas más trabajadas en este lematizador.

ocurrido con el término «rápel», que aparecía en el texto de trabajo sin tilde («rapel») por un error ortográfico del texto original pero en la base de datos está incluido con ella.

En segundo lugar, incluimos los problemas morfológicos causados por la lematización errónea de la herramienta ante formas conjugada de verbos. En nuestro caso, por ejemplo, encontramos que formas como «rapelamos», «caminé» o «descendía» no se lematizan con sus infinitivos («rapelar», «caminar» o «descender», respectivamente) incluidos en la base de datos, lo que causa, de nuevo, un problema de vinculación. No obstante, en la tabla anterior se observa que este proceso sí funciona correctamente para verbos más comunes como son «ser», «suceder» o «conocer».

Por último, encontramos los problemas terminológicos, que se encuentran a menudo relacionados con el problema anterior, ya que a veces los anteriores se dan por causas que encuentran una explicación en este grupo. Estos problemas se deben a la falta de inclusión de algunos términos en el léxico básico con el que trabaja la herramienta. En otras palabras, *TreeTagger* utiliza de base un léxico formado por un número x de palabras, por lo que, cuando aparece una palabra que no está incluida en dicho léxico, la herramienta no la reconoce y, por tanto, no la lematiza. Esto ocurre con los términos que obtienen el resultado <unknown> en la Tabla 51, como por ejemplo .

A pesar de las dificultades encontradas por el empleo de este lematizador, *TreeTagger*, nos decidimos en su momento por él después de realizar pruebas con otros lematizadores, como *Lematizador*⁶⁴⁶ y *Freeling*,⁶⁴⁷ que hubieran evitado los errores vistos anteriormente con el idioma español pero que no se ajustaban a nuestras necesidades en otros aspectos, como por ejemplo el hecho de que el programa *Lematizador* tiene que ser utilizado en línea y no dispone de opción para utilizarlo dentro de otra aplicación como hace *TreeTagger*. Sin embargo, y como decimos, a pesar de los problemas encontrados, elegimos *TreeTagger* por ser el lematizador que mejor se adaptaba a nuestras necesidades y que cumplía, por tanto, nuestros requisitos previos, a saber: multilingüidad, accesibilidad desde todos los sistemas operativos (Mac, Windows, Linux, etc.), integración en la aplicación *Trandix*, entre otros. Asimismo, no debemos olvidar que es uno de los lematizadores más utilizados en el campo del PLN.

⁶⁴⁶ <<http://www.gedlc.ulpgc.es/investigacion/scogeme02/lematiza.htm>>

⁶⁴⁷ <<http://nlp.lsi.upc.edu/freeling/>>

4.3.1.3.2. Comparativa entre Trandix y sus antecedentes

Una vez que hemos descrito la herramienta *Trandix*, así como sus características más relevantes, pasaremos a comparar esta aplicación con sus antecedentes, es decir, los sistemas de comprensión que presentábamos al comienzo de este apartado (cf. 4.3.1.1.). Para llevar a cabo esta comparativa, hemos diseñado una tabla formada por las principales características de estos sistemas, con objeto de ordenar la información recogida de estos sistemas y, así, ser capaces de establecer una comparación entre ellos y determinar las ventajas e inconvenientes de cada uno con respecto a nuestra herramienta.

De esta manera, hemos considerado relevante los siguientes criterios:

- Fecha, que hace referencia a la fecha de creación o a su actualización (en su caso). En los casos en los que se sigue trabajando en esta herramienta y produciendo mejoras o ampliaciones, hemos optado por indicar «Actual».
- Lenguas, que recoge la combinación de lenguas con las que trabaja el sistema. Para indicar estas lenguas, hemos optado por seguir la norma ISO 3166 (1974).
- Objetivo, que se refiere a la finalidad de la aplicación correspondiente. En este contexto, podemos encontrar que algunos de estos sistemas han sido diseñados para apoyar la comprensión lectora de una lengua extranjera y, en otros casos, se observa cómo estos sistemas han trasvasado ese objetivo y pretenden asistir a los traductores a realizar su trabajo en mejores condiciones.
- Contenido original/procedente de diccionarios ya elaborados, que distingue entre los recursos que utilizan contenido procedentes de recursos anteriores ya elaborados (ya sea en papel o en formato electrónico) de aquellos que cuentan con contenido elaborado *ad hoc*, es decir, concretamente para el recurso en cuestión y que, por tanto, forman parte del trabajo terminológico.
- Importación y etiquetado, que requiere información acerca de las posibilidades de importación que ofrece el sistema, además de la necesidad o no de etiquetar la información original para su uso en la aplicación.

- Disponibilidad y acceso, que hace referencia a la posibilidad de acceder a la aplicación y al modo de acceso, es decir, si requiere de licencia comercial, el acceso es gratuito, etc.
- Campos de información que presenta, que se refiere a los campos que se muestran en la entrada de la unidad, es decir, si la entrada consta únicamente de equivalentes o, por el contrario, incluye también definiciones, categoría gramatical, contexto, etc.
- Independencia o imbricación en el sistema, que solicita información acerca del funcionamiento de la aplicación y, por tanto, de si se trata de un programa independiente al que hay que subir datos o textos concretos o si, por el contrario, se trata de un programa que se instala en el sistema y que puede utilizarse en cualquier entorno.
- Otros comentarios, que permite introducir aquellos datos que nos parezcan relevantes y que no se hayan indicado en los apartados anteriores por tratarse de características propias de los sistemas.

Teniendo en cuenta estos criterios, presentaremos a continuación la tabla confeccionada con los datos correspondientes a los sistemas analizados:

| Nombre: | Fecha: | Lenguas: | Objetivo: | Fuente de la información: | Importación y etiquetado: | Disponibilidad y acceso: | Campos de información que presenta: | Independiente / Imbricación en el sistema: | Otros comentarios: |
|-----------------------|--------|-------------------|---|--|---|---|---|---|---|
| COMPASS | 1996 | EN- DE / FR-EN | Lectura de texto en lengua extranjera con escasos conocimientos de la lengua. | De diccionarios generales elaborados en papel. | Etiquetado en SGML y acciones de adaptación y revisión. | No disponible | Definición y equivalente. | Plataforma independiente | Se deben efectuar acciones para adaptar, revisar y reescribir las fuentes utilizadas para poder ser incluidas en el sistema. |
| MobiMouse Plus | Actual | EN-PL | Lectura de texto en lengua extranjera con escasos conocimientos de la lengua. | De diccionarios generales elaborados en papel. | Formato XML | Licencia comercial | Categoría gramatical, definición y equivalente. | Imbricada (se instala directamente en el ordenador) | Constantemente están trabajando en la elaboración de diccionarios que se pueden cargar a este sistema, pero siempre son diccionarios generales y con polaco. |
| Smarty | 2009 | EN-BG | Lectura de texto en lengua extranjera con escasos conocimientos de la lengua. | Del recurso general <i>WordNet</i> . | Sin etiquetar (utiliza directamente el formato del recurso original). | Disponible para uso interno del grupo de investigación. | Categoría gramatical, género, número, definición, fraseología, equivalentes. | Plataforma independiente | Sistema híbrido que permite la consulta directa sin necesidad de cargar un texto, además de ser un asistente para la descodificación similar a los demás. |
| Terminator | Actual | EN, FR, DE, SL | Ayuda a la traducción. | De la base de datos sobre terminología de la UE, <i>Eyvroterm</i> . | Sin etiquetar (utiliza directamente el formato del recurso original). | Acceso libre | Definición, equivalentes en las lenguas disponibles, dominio, fecha de creación. | Plataforma independiente | Forma parte de la base de datos <i>Eyvroterm</i> , una base de datos elaborada a partir de textos de la UE y principalmente para la lengua eslovena. Dispone de búsqueda en recursos externos. |
| Trandix | Actual | EN, DE, ES | Ayuda a la traducción. | De recurso especializado previamente sobre terminología del turismo de aventura para la herramienta. | Formato TBX | Acceso libre | Por defecto, los campos de la ISO requeridos por traductores profesionales, aunque los usuarios pueden elegir los campos que desean visualizar para cada entrada. | Plataforma independiente | Ofrece diferentes formas de búsqueda (búsqueda textual, en recursos externos y por términos). Permite la posibilidad de introducir otros recursos terminológicos, además del disponible actualmente. Permite editar el texto en la misma pantalla de la aplicación. |

Tabla S2: Tabla comparativa entre *Trandix* y sus antecedentes.

Gracias a esta tabla, podemos observar las diferencias y similitudes de los diferentes sistemas y, además, establecer las ventajas y desventajas que presenta cada uno de ellos.

Para empezar, comentaremos las características que comparten la mayoría de estos sistemas, a saber: fecha de creación y su independencia con el sistema. Con respecto al primer aspecto, vemos que todos los sistemas, a excepción del sistema *COMPASS*, presentan una fecha de creación o de actualización muy reciente, lo que permite detectar el reciente interés que se ha despertado sobre estos sistemas y la actualidad del trabajo llevado a cabo. Con relación al segundo aspecto, observamos también que la mayoría de los sistemas, excepto *MobiMouse Plus*, se tratan de aplicaciones independientes, es decir, de programas que tienen una interfaz y necesitan una interacción por parte del usuario. Por el contrario, *MobiMouse Plus* se integra dentro del sistema y permite su acceso desde cualquier entorno del ordenador, lo que presenta ventajas en un sentido, por ejemplo, la posibilidad de acceder al léxico en cualquier momento; pero, a la vez, puede dar lugar a desventajas como, por ejemplo, la posible ralentización del sistema tanto al inicio como durante su uso por culpa de este programa.

Otra característica de los sistemas que resulta importante a la hora de utilizarlos y probarlos consiste en el acceso a la herramienta. De los cinco sistemas que hemos evaluado, solo dos tienen acceso libre, mientras que uno tiene acceso bajo licencia comercial, otro solo es accesible dentro del grupo de investigación en el que se ha llevado a cabo y, por último, otro ofrece ningún tipo de acceso porque, según nuestro conocimiento, ha dejado de funcionar. Con esta realidad pocos usuarios pueden acceder a este tipo de sistemas, por lo que las ventajas que puedan ofrecer se limitan a la duración de la investigación correspondiente.

Seguidamente, trataremos las características más relevantes para nuestra investigación, que no debemos olvidar que está dirigida a traductores. En este sentido, solo dos de los cinco sistemas, *Terminator* y *Trandix*, tienen como objetivo la asistencia a los traductores durante su trabajo, es decir, la asistencia tanto a la comprensión del texto origen como a la codificación del texto meta. En la misma línea, encontramos que solo estos dos sistemas contienen información relativa a campos de especialidad

(terminología de la Unión Europea y turismo de aventura, respectivamente) y que, de nuevo, solo estos dos sistemas incluyen información original y elaborada con unos objetivos definidos y unos destinatarios concretos. Por el contrario, los demás sistemas presentan la finalidad de asistir durante la lectura de textos en lengua extranjera y, por tanto, se limitan a la fase de descodificación del mensaje origen; todos ofrecen parte de recursos generales, es decir, no ofrecen información especializada sobre ningún dominio concreto; y, por último, cuentan con información procedentes de recursos ya elaborados (en formato papel o electrónico) que únicamente han tenido que adaptar a un formato diferente del original. Por último, es importante reseñar también que solo estos dos sistemas, *Terminator* y *Trandix*, cuentan con la posibilidad de intercambiar las lenguas de trabajo, es decir, disponen de cuatro y tres lenguas respectivamente con las que se puede trabajar en cualquier combinación, lo que amplía el uso de dichos sistemas en gran medida, frente al resto de sistemas que cuentan solo con combinaciones bidireccionales.⁶⁴⁸

Por último, indicaremos las ventajas que presenta *Trandix* con respecto a estos sistemas y que, por tanto, la convierte en una aplicación innovadora, completa y eficiente para el proceso traductor:

- En primer lugar, cabe destacar el tipo de formato de importación con que trabaja, formato TBX, el formato estándar para el intercambio de terminología propuesto por las organizaciones internacionales ISO y LISA. En este sentido, esta aplicación se ajusta a los requisitos actuales de normalización y, además, favorece el intercambio de terminología y la reusabilidad de trabajos terminológicos.
- En segundo lugar, es relevante la diversidad de campos terminológicos que ofrece para la entrada de unidades. Asimismo, estos campos proceden de la norma ISO 12620 (1999) y han sido seleccionados según las indicaciones extraídas de la encuesta llevada a cabo a traductores profesionales durante el comienzo de esta investigación (cf. 2.5.3.), lo que persigue una adecuación a las necesidades y requisitos del grupo de destinatarios del recurso. Además de la

⁶⁴⁸ Hay que tener en cuenta, además, que *Trandix* es la única aplicación que ofrece el español como lengua de trabajo.

selección justificada de campos y de la variedad que se proporciona, *Trandix* ofrece la posibilidad al propio usuario de seleccionar los campos que desea visualizar en las entradas de los términos, ya que la casuística de usuarios y de conocimientos puede dar lugar a necesidades contrarias dentro de un mismo colectivo.

- En tercer lugar, nos parece interesante subrayar la posibilidad de importar diferentes archivos TBX con bases de datos terminológicas que versen sobre diferentes dominios de especialidad, es decir, actualmente *Trandix* está concebido para el dominio del turismo de aventura, puesto que el archivo TBX que está cargado en la herramienta está formado por la base de datos terminológica de este sector, pero la aplicación no está vinculada a él y, por tanto, permitiría que se añadieran otros archivos en formato TBX con información de otros campos de especialidad y eliminar, así, la información sobre el turismo de aventura que contiene.
- Por último, la aplicación ofrece como novedad las diferentes opciones de búsquedas que presenta, a saber: búsqueda textual, como el resto de aplicaciones; búsqueda externa, como *Terminator*, y búsqueda directa de términos, como *Smarty*. En otras palabras, *Trandix* es la única aplicación que aúna los tres tipos de búsquedas que ofrecen por separado las otras herramientas, las cuales solo tienen en común la búsqueda textual. En este sentido, nuestra aplicación ofrece la ventaja de la flexibilidad a la hora de realizar búsquedas, lo que permite una mayor eficiencia a la hora de trabajar con esta herramienta frente a las demás.

Finalmente, con objeto de comprobar la validez de nuestra herramienta, llevaremos a cabo un experimento con semiprofesionales⁶⁴⁹ de la traducción que pasamos a presentar en el siguiente apartado.

⁶⁴⁹ En este caso, seguimos las indicaciones de Corpas Pastor (2008:118) al considerar traductores *semiprofesionales* a alumnos que se encuentran cursando el segundo cuatrimestre del último curso de licenciatura.

4.3.1.4. Evaluación de *Trandix*

Llegados a este punto en el que hemos presentado la aplicación *Trandix* y sus diferentes funcionalidades y características más relevantes, además de haber realizado una comparación entre *Trandix* y sus antecedentes, podemos afirmar objetivamente que se trata de una herramienta completa, flexible y robusta, ya que ofrece diferentes tipos de búsqueda, permite editar el texto en la misma pantalla de la aplicación, da la posibilidad de seleccionar los campos terminológicos que se desean visualizar en la entrada, funciona con todos los sistemas operativos, entre otras ventajas. Asimismo, consideramos que esta aplicación se adecúa a las necesidades de sus usuarios finales de acuerdo con los resultados obtenidos en la encuesta realizada *ad hoc* durante esta investigación, en la que se dejaba patente la necesidad de ofrecer un recurso dinámico que facilitara la accesibilidad a la información sin la necesidad de consultar numerosas fuentes, que incluyera información pragmática sobre el uso de las unidades, posibles diferencias entre lenguas y el dominio concreto de utilización, que las definiciones ofrecidas fueran concretas y correctas, entre otros. En este contexto, *Trandix* también se puede considerar como satisfactoria en este sentido, puesto que permite diferentes tipos de búsquedas para acceder a la información, incluido una opción para buscar directamente en recursos externos, ofrece información pragmática en el apartado de *Observaciones* y de *Contexto*, ofrece definiciones concretas basadas en el marco semántico de la unidad de búsqueda en cuestión así como hipervínculos internos para acceder a otras unidades de la base de datos, permite la edición de los textos en la misma pantalla principal, entre otras opciones importantes. Por todo ello, consideramos que la herramienta *Trandix* cumple las expectativas y los requisitos más comunes tanto desde un punto de vista técnico como humano, referente a los usuarios finales.

No obstante, y a pesar de nuestras conclusiones presentadas con anterioridad, hemos estimado oportuno evaluarla de forma externa y conocer, así, las opiniones positivas o negativas de personas ajenas a la investigación. De esta manera, hemos llevado a cabo un experimento empírico con alumnos oficiales⁶⁵⁰ matriculados en el último curso de la licenciatura de Traducción e Interpretación de la Universidad de

⁶⁵⁰ Agradecemos a los alumnos su participación en este experimento empírico, ya que sin ellos este trabajo no habría podido realizarse.

Málaga, es decir, traductores semiprofesionales de acuerdo con la terminología de Corpas Pastor (2008: 118).

El objetivo de nuestro estudio consiste en comparar el proceso de traducción de textos sobre el turismo de aventura con la herramienta *Trandix* y sin ella, es decir, estudiaremos las posibles mejoras que proporciona el uso de la herramienta en un trabajo de traducción frente a un trabajo de traducción sin esta herramienta. De esta manera, podremos, por un lado, observar si realmente existen dichas mejoras a la hora de traducir con respecto a varios aspectos establecidos previamente (tiempo, facilidad para la búsqueda de información, gestión de la herramienta, etc.) y, por otro, comprobar que efectivamente se adecúa a las necesidades de los traductores (en este caso, semiprofesionales)⁶⁵¹ según nuestras propias conclusiones. En definitiva, la finalidad de este experimento consistirá en comprobar si la herramienta reporta los beneficios que consideramos desde nuestro punto de vista con respecto a sus funcionalidades y forma de uso frente a otras formas de trabajo (uso de diccionarios, bases de datos, corpus, etc.) más tradicionales y, así pues, si realmente es eficiente para aplicarse en los procesos de traducción, inicialmente de textos de turismo de aventura, pero posteriormente en otros campos de especialidad.⁶⁵²

A continuación, presentaremos el experimento empírico realizado dividido en las diferentes fases de las que se compone, a saber: 1. Fase preparatoria, 2. Fase de ejecución y 3. Fase de evaluación de los resultados. Finalmente, cerraremos este apartado con la exposición de nuestras conclusiones, extraídas a partir de los resultados obtenidos en este experimento.

⁶⁵¹ Para el objetivo de nuestro estudio quisimos incluir también a traductores profesionales, pero lamentablemente no ha sido posible llevarlo a cabo por lo que este estudio se incluirá como línea de investigación futura.

⁶⁵² Recordemos que actualmente *Trandix* está dirigida únicamente al turismo de aventura, ya que cuenta con una base de datos en formato TBX con su terminología y, además, los recursos especializados ofrecidos están relacionados con dicha temática. No obstante, se pretende continuar trabajando en la ampliación de la herramienta para posibilitar el trabajo con otros dominios y otras lenguas.

4.3.1.4.1. Fase preparatoria

La primera fase de nuestro experimento consiste en determinar todos los aspectos relevantes para llevarlo a cabo, a saber: la finalidad que se persigue con el estudio, la forma en la que se va a llevar a cabo, la selección de la muestra y la preparación y búsqueda del material necesario para llevarlo a cabo. En este sentido, esta fase es sumamente importante, puesto que de ella depende que todo el experimento esté bien diseñado y que, por tanto, se puedan alcanzar los objetivos planteados.

En el orden establecido en la enumeración anterior, pasaremos a comentar estos aspectos y a especificar las decisiones tomadas a este respecto.

4.3.1.4.1.1. Finalidad del experimento

Como se ha indicado con anterioridad, la finalidad que se persigue con este estudio consiste en el estudio de las ventajas y los beneficios que puede aportar la herramienta *Trandix* durante la traducción de textos de turismo de aventura, concretamente durante la búsqueda y consulta de información terminológica sobre dicho dominio, frente a otras formas de búsqueda y consulta más tradicionales como pueden ser la consulta de diccionarios, bases de datos, etc. De esta manera, podremos, por un lado, confirmar si realmente existen dichas mejoras a la hora de traducir con respecto a varios aspectos establecidos previamente, como son el tiempo de ejecución de la traducción, la facilidad para la búsqueda de información, las dificultades a la hora de utilizar la herramienta *Trandix*, etc.; y, por otro, comprobar que efectivamente se adecúa a las necesidades de los traductores semiprofesionales objetos de estudio según nuestras propias expectativas. En definitiva, la finalidad de este experimento consiste en comprobar si la herramienta reporta los beneficios que consideramos desde nuestro punto de vista con respecto a sus funcionalidades y forma de uso frente a otras formas de trabajo (uso de diccionarios, bases de datos, etc.) más tradicionales y, así pues, si realmente es eficiente para aplicarse en los procesos de traducción de textos de turismo de aventura. En segunda instancia, se podrá observar la validez de la base de datos elaborada con la terminología del turismo de aventura según la metodología propuesta en esta investigación con respecto a otras fuentes terminológicas.

Una vez especificada la finalidad de nuestro estudio, pasaremos a indicar cómo se ha organizado la ejecución de dicho experimento.

4.3.1.4.1.2. Planificación del experimento

El objetivo propuesto para nuestro estudio se alcanzó mediante un seminario teórico-práctico que tuvo lugar durante el segundo cuatrimestre del curso académico 2010/2011 en la Universidad de Málaga, concretamente en los grupos A y B de la asignatura de *Traducción especializada inglés/español-español/inglés (b): textos científico-técnicos* del último curso de la licenciatura de Traducción e Interpretación.⁶⁵³ Este seminario estuvo dividido en tres partes: durante la primera parte, los dos grupos A y B de la asignatura asistieron a una sesión teórica sobre las nuevas tecnologías relacionadas con la competencia documental del traductor, donde asistieron a una presentación de los diferentes tipos de fuentes de información que pueden emplear durante sus traducciones (tanto en un ámbito académico como profesional), haciendo especial hincapié en las fuentes de información terminológica y sus formas más innovadoras. Dentro de esta parte, se les dio a conocer la herramienta *Trandix* y se procedió a explicar sus funcionalidades y características principales. En la segunda parte de este seminario, se les solicitó a la muestra completa de sujetos que realizasen una traducción de un texto de turismo de aventura con los medios que conocían y que tenían a su disposición (tanto en formato electrónico como en formato papel), es decir, los sujetos tuvieron acceso a Internet y a todo el material del que disponen en una situación normal de traducción, a fin de que el proceso de traducción se diera en un entorno lo más parecido a la realidad. En este sentido, pudieron utilizar todas las herramientas presentadas en la fase inicial del seminario, a excepción de la herramienta *Trandix*. Asimismo, dicho texto vino acompañado de una serie de indicaciones con respecto al encargo de traducción para contextualizar el trabajo e indicar el escopo.

En la tercera parte de este seminario, se les solicitó a todo el grupo de sujetos que realizasen otra traducción de las mismas características que la anterior pero, en esta ocasión, con la herramienta *Trandix*. Para ello, antes de comenzar con la traducción, se

⁶⁵³ Agradecemos a las docentes de cada grupo de esta asignatura, a la Dra. D.^a Gloria Corpas Pastor y a la Dra. D.^a Míriam Seghiri Domínguez, el habernos dado la oportunidad de organizar este seminario como parte de su docencia.

les facilitó a todo el grupo acceso a la herramienta y se les recordó los pasos que debían seguir para utilizarla, con objeto de que los alumnos pudieran utilizarla de la forma más eficiente posible. Seguidamente, se les entregó el texto para traducir junto con un encargo de traducción, igual que en la primera parte, y se les tuvo en cuenta el tiempo de ejecución al igual que en la fase anterior.

Con la utilización del mismo grupo de sujetos en ambas partes del seminario práctico, es decir, los mismos sujetos realizaron en primer lugar la traducción sin herramienta y, posteriormente, la traducción con la herramienta, hemos pretendido fijar las variables relativas a la formación, competencia traductora, nivel de conocimientos y habilidades de los sujetos, que surgirían al trabajar estas partes con grupos de sujetos diferentes, y, así, tener como única variable el cambio de situación comunicativa al incluir o no la herramienta *Trandix*. De esta manera, se pueden comparar objetivamente los resultados obtenidos en la primera parte del taller (sin herramienta) con los de la segunda parte (con herramienta). De lo contrario, podrían surgir diferencias entre los grupos que *a priori* no podrían controlarse.

4.3.1.4.1.3. Muestra del experimento

Como ya hemos indicado, la muestra de este experimento empírico está compuesta por alumnos matriculados en el último curso de la licenciatura de Traducción e Interpretación de la Universidad de Málaga que se encuentran cursando la asignatura de *Traducción especializada inglés/español-español/inglés (b): textos científico-técnicos* durante el curso académico 2010/2011. De esta manera, al no haber podido acceder a traductores profesionales para llevar a cabo el experimento debido a la falta de disponibilidad y de tiempo por su parte, hemos seleccionado esta muestra por una serie de motivos que exponemos a continuación:

- Estos alumnos se encuentran a punto de finalizar su formación en traducción, por lo que podemos confirmar que los estudiantes cuentan con las competencias traductológicas suficientes para enfrentarse a una traducción directa y ofrecer buenos resultados.

- La gran mayoría de estos alumnos accederá al mercado laboral a la finalización de este curso académico, por lo que su interés por las nuevas tecnologías de apoyo a la documentación se presupone alto.
- El hecho de encontrarse en el último año de licenciatura nos asegura que están familiarizados con la traducción especializada y con las características generales que presenta, así como el protocolo que deben seguir a la hora de realizar traducciones de este tipo.
- A su vez, la traducción de textos turísticos les resulta familiar al haber sido tratada como parte de la asignatura de *Traducción general inglés/español-español/inglés (b)* en el segundo curso de licenciatura y en las asignaturas de traducción general de las lenguas C (alemán, italiano, francés, griego, árabe) durante el tercer curso. La formación adquirida durante estas asignaturas les proporciona suficientes conocimientos terminológicos, fraseológicos, documentales y textuales para abordar las traducciones de textos promocionales del sector turístico con las características que presentan los seleccionados en este experimento.

Tras describir a los sujetos del experimento y haber motivado nuestra elección, nos centraremos a continuación en el material seleccionado para llevar a cabo el seminario, que consta de los textos originales seleccionados para realizar las traducciones, dos cuestionarios para las diferentes partes del seminario práctico y un documento con las indicaciones de la traducción.

4.3.1.4.1.4. Material para el experimento

Llegamos a la preparación del material necesario para llevar a cabo el experimento. Como ya hemos avanzado, para el experimento se realizó una selección de dos textos originales en inglés para llevar a cabo las traducciones programadas del seminario y, además, se elaboró un cuestionario dividido en dos partes correspondientes a sendos seminarios prácticos, que se entregó a la finalización de la última traducción, y un documento con las indicaciones de traducción que acompañaba al texto origen y donde se especificaba la situación comunicativa en la que se insertaba dicha traducción,

así como la finalidad que se perseguía con ella, es decir, el escopo.⁶⁵⁴ A continuación, especificamos las características y el contenido de cada uno de estos documentos.

Con respecto a la selección de textos que hemos realizado para el seminario práctico, nos propusimos que fueran textos de las mismas características a fin de poder realizar comparaciones y valoraciones objetivas y libres de sesgos o incongruencias. En definitiva, se ofreció la misma situación comunicativa con y sin la herramienta *Trandix* a los sujetos del experimento. De lo contrario, la comparación entre una y otra situación no habría sido posible. Los textos seleccionados han seguido los mismos criterios que establecimos para la compilación de nuestro corpus, puesto que han sido extraídos de él, a saber:

- Criterio temporal: los textos seleccionados son actuales, por lo que comparten el mismo periodo temporal.
- Criterio geográfico: los textos seleccionados están limitados geográficamente a Reino Unido.
- Criterio formal: los textos seleccionados pertenecen se basan en una situación comunicativa de especialista - no especialista, es decir, de especialista a turista que busca una actividad de este tipo o a *amateur* que desea realizar estas actividades durante sus vacaciones. Además, los textos son escritos, originales y completos, es decir, no son fragmentos de textos de mayores dimensiones.
- Tipología textual: los textos seleccionados son textos promocionales publicados en Internet por empresas registradas que ofrecen actividades del turismo de aventura.
- Número de páginas: ambos textos ofrecen aproximadamente el mismo número de páginas (170 palabras el primer texto y 165 palabras el segundo texto).
- Número de términos: los dos textos presentan un número similar de términos, a fin de evitar que un texto pueda resultar más complicado que otro por la densidad léxica que presente.

⁶⁵⁴ Para consultar el material elaborado para llevar a cabo este experimento, remitimos al Anexo VI.

A continuación, presentamos los dos textos que se han seleccionados para realizar el experimento: en primer lugar, se presenta el texto escogido para realizar la parte del seminario sin la herramienta *Trandix* (Figura 107) y, en segundo lugar, se indica el que se ha utilizado para la parte con la herramienta (Figura 108).

Canyoning/ Gorge scrambling

Probably one of the best extreme adrenaline activities around. If you haven't tried canyoning before, you're in for a real adrenaline adventure.

Canyoning is a mixture of sliding, swimming, climbing, jumping down waterfalls, jumping off cliffs and high banks into deep plunge pools. It's a real adrenalin rush.

The session starts with a short gorge scramble up to the canyoning venues and then the fun really starts. As with all activities all specialist equipment: buoyancy aids, cagoul jackets, helmets, ropes, etc. All you need to bring are an extra fleece with the wetsuit for warmth, a full change of clothing including shoes or boots and a reasonable level of fitness and a touch of nerve.

Canyoning is a great outdoor adventure activity for a half day session and can be combined with one of the following activities to make a full day, choose from either rock climbing, abseiling, open Canadian canoeing, kayaking, mountain biking, orienteering, hill walking, and archery.

Figura 107. Texto promocional seleccionado para la traducción 1 (sin Trandix).

Aquaseiling

What is Aquaseiling?

As an outdoors adventure adrenalin activity Aquaseiling is one of the best. Abseiling down waterfalls is one of the fastest growing adventure activities. We use some of the best aquaseiling venues in Lancashire and the Lake District.

As with abseiling, you will be kitted out with a harness and a helmet for the aquaseil and then attached firstly to a safety rope and then to static abseil rope using a figure of eight and a karabiner. Using the figure of eight you will make your way down the aquaseil at your own pace. With an instructor holding the safety rope this ensures your safety at all times.

Don't feel nervous our team of highly trained staff will talk you through the aquaseil activity and make you feel at ease.

All specialist equipment is supplied for the aquaseil, along with a wetsuit, waterproof trousers and a waterproof jacket all you need to bring is spare boots or shoes and a swimming costume!

Figura 108. Texto promocional seleccionado para la traducción 2 (con *Trandix*).

Una vez presentado los textos seleccionados, pasaremos a describir el resto de documentación elaborada para este experimento, a saber: el cuestionario y el documento indicativo del encargo de traducción.

Para llevar a cabo la finalidad que nos hemos propuesto en este estudio, se elaboró un cuestionario dividido en dos partes, una para cada taller práctico, en el que se les incluyeron diferentes preguntas con respecto a la traducción solicitada en cada fase del experimento. Las preguntas que se incluyeron son de tres tipos:

- Preguntas cerradas: En ellas, los alumnos deben seleccionar las respuestas de las opciones indicadas en el cuestionario. Dentro de este grupo, encontramos varios tipos de preguntas:
 - o Preguntas de calificación, en la que los alumnos deben otorgar una puntuación del 1 al 5 de acuerdo con una escala en la que 1 es muy fácil y 5 es muy difícil. Este tipo de preguntas se ha empleado, por ejemplo, para valorar la dificultad de los textos para traducir.

- Preguntas de alternativa simple, donde las posibles respuestas eran Sí o No. Este tipo de preguntas se ha utilizado, por ejemplo, para preguntar si han tenido suficiente tiempo a la hora de realizar la traducción.
- Preguntas de opción múltiple, en las que los alumnos deben seleccionar o bien una única respuesta de varias opciones dadas o bien se les permite seleccionar más de una opción. En este caso, encontramos preguntas como, por ejemplo, relativas a su opinión acerca de la ayuda que les ha ofrecido *Trandix* a la hora de traducir, donde las posibles respuestas eran 1. Totalmente, 2. Mucho, 3. Bastante, 4. Poco y 5. Nada; y, por el contrario, preguntas en las que los alumnos podían elegir varias opciones, como por ejemplo, cuando se les pregunta por los aspectos más complicados a la hora de realizar la traducción.
- Preguntas abiertas: Los alumnos indican su opinión sobre la pregunta realizada. Estas preguntas se utilizan, por ejemplo, para conocer las ventajas y desventajas que encuentran en la herramienta *Trandix* los aspectos que le han resultado más complicados de la traducción o los cinco términos que les ha supuesto una gran dificultad.
- Preguntas mixtas: Los alumnos tienen la oportunidad de razonar, especificar o ampliar su respuesta a través de la opción «otros» o de la opción «Especifique:» en preguntas cerradas. Este tipo de pregunta ha sido el más utilizado en estos cuestionarios, puesto que permite conocer un poco más el protocolo de trabajo de los alumnos así como sus opiniones y posibles críticas.

En general, las preguntas de este cuestionario se han planificado con la idea de poder recopilar información sobre las dos partes del experimento (es decir, con la herramienta *Trandix* y sin ella) acerca de los puntos que consideramos más relevantes y que indicamos a continuación:

- El tiempo de ejecución de la traducción, es decir, el tiempo que tardan los sujetos desde que se les entrega la traducción hasta que la terminan y se disponen a entregarla.⁶⁵⁵
- La eficiencia en la búsqueda de información, es decir, la combinación del tiempo y los resultados obtenidos. En este caso, se mide la rapidez con la que encuentran la información deseada.
- El protocolo de búsqueda seguido para realizar la traducción con una metodología y con otra.

Además de las preguntas relativas a estos tres puntos más importantes, se incluyeron también una serie de preguntas generales para conocer a la muestra que hacen referencia a la asignatura en la que realiza el experimento: la edad del sujeto y su lengua materna. Por último, cabe destacar un conjunto de preguntas de la segunda parte del cuestionario referentes a la herramienta *Trandix*, puesto que, recordemos, uno de los principales objetivos es conocer exactamente la opinión de los sujetos con respecto a su funcionamiento, características principales, así como las ventajas o inconvenientes que encuentran en su uso. Por todo ello, el cuestionario de la segunda parte resulta más completo que el primero, ya que, además de incluir preguntas relativas a los cuatro puntos anteriores, añade cuestiones sobre los beneficios (en su caso) que considera que aporta la herramienta, los posibles cambios que propondría, la metodología que prefiere y por qué, la facilidad de uso, entre otros. De esta manera, consideramos que se puede obtener información adecuada y pertinente para extraer conclusiones en torno a la valoración de estos sujetos sobre *Trandix* e, incluso, considerar para un futuro la realización de cambios de aquellas propuestas que puedan surgir.

Una vez comentado el cuestionario, seguiremos con las indicaciones del encargo. Desde nuestro punto de vista, consideramos imprescindible que toda traducción que se encargue, ya sea en el ámbito académico o profesional, se acompañe de un documento en el que se detallen los datos relativos a la situación comunicativa en torno a la cual surge esa necesidad de traducción y, muy especialmente, su finalidad o escopo. De esta manera, el traductor conocerá desde un primer momento los

⁶⁵⁵ Como veremos más adelante, la entrega de la traducción se realizará a través de las asignaturas virtuales albergadas en el Campus Virtual de la Universidad de Málaga que disponen las asignaturas en las que hemos llevado a cabo el experimento.

condicionantes extratextuales que influyen en la traducción y adaptará su traducción a ella para alcanzar los mejores resultados de su trabajo. En esta misma línea, elaboramos como material de apoyo al texto para traducir un documento con esta información.

El documento elaborado en el seno de esta investigación se divide en tres apartados diferenciados, a saber: 1. Finalidad o escopo, 2. Situación comunicativa y 3. Entrega. En el primer apartado, se incluye información acerca de la finalidad que persigue la traducción que se encarga, que en este caso consiste en la traducción del texto origen para su publicación en una página web de una empresa de turismo de aventura como parte de su información promocional y publicitaria sobre las actividades de aventura en Reino Unido. El fin último que se persigue es promocionar las actividades de aventura entre el mercado turístico español en este país. Con esta información, el traductor debe entender que se trata de una traducción turística, según las características que presenta este tipo de traducción especializada en cuanto a su léxico, sintaxis, restricciones, formato, etc.,⁶⁵⁶ así como el tipo de comunicación de que se trata, que en este caso es de especialista - no especialista. En definitiva, el traductor recibe información para poder adecuar su traducción a dicha finalidad o escopo.

El segundo apartado del documento hace referencia a la situación comunicativa en la que se da la traducción solicitada. Como en cualquier situación de este tipo, encontramos una serie de elementos que intervienen, a saber: los participantes, el medio, el código y el contexto (cf. 1.2.2.3.), que aparecen especificados en el documento junto con un elemento añadido referente a la temporalidad de la traducción. Con respecto a los participantes, encontramos cuatro tipos diferentes: 1. el autor del texto original, que en este caso es desconocido;⁶⁵⁷ 2. el emisor de la traducción, que se trata de la empresa de aventura en Reino Unido que necesita la traducción; 3. el iniciador, que en este caso es el profesor de la asignatura por encontrarnos en un ámbito

⁶⁵⁶ A pesar de que las asignaturas en las que se ha llevado a cabo el experimento no incluyen en su programación la traducción turística, estos alumnos han trabajado este tipo de traducción en asignaturas anteriores, concretamente en las asignaturas de segundo y tercer curso de licenciatura de Traducción general B (inglés) y C (alemán e italiano), por lo que partimos de que conocen las características propias de este tipo de traducción especializada (uso de adjetivos positivos y atrayentes, lenguaje metafórico, oraciones comparativas y superlativas, etc.). Para revisar las características del lenguaje turístico, remitimos al capítulo tres (cf. 3.3.4.1.).

⁶⁵⁷ Recordemos que una de las características de los textos turísticos consiste en la frecuente omisión del nombre del autor del texto, ya sea este el autor original o el traductor.

académico, y 4. el destinatario final, que hace referencia al grupo de usuarios al que está destinada la traducción del texto, es decir, a turistas españoles que se interesen por la práctica de esta actividad de aventura en Reino Unido. Después de especificar los participantes, se indica el medio en que aparecerá la traducción, que en este caso es en la página web de la empresa de aventura, es decir, será un texto escrito en formato electrónico. Seguidamente, encontramos el código, que es la lengua española de España, y el contexto, que hace referencia a la situación real que denota la necesidad de la traducción, que en este caso se basa en el hecho de que cada vez más turistas españoles se interesan por la práctica de actividades de aventura en Reino Unido. Por último, hemos incluido aquí el elemento de temporalidad a fin de indicar el periodo de publicación del texto origen (entre 2008 y 2010) y la necesidad, por tanto, de mantener las posibles referencias temporales que aparezcan en el texto.

Por último, encontramos las indicaciones referentes al tiempo y forma de la entrega. En primer lugar, se especifica el momento en el que se debe entregar la traducción, que en este caso es a la finalización del seminario práctico; y, en segundo lugar, la forma en la que se debe entregar el trabajo. Con respecto a este segundo aspecto, se indica el medio de entrega (a través de una tarea creada para tal efecto en la asignatura virtual correspondiente a la asignatura en la que se realiza el seminario práctico: *Traducción especializada inglés/español-español/inglés (b): textos científico-técnicos*), la codificación que tienen que utilizar los sujetos para nombrar el archivo con objeto de vincular de forma unívoca los cuestionarios y las traducciones realizadas y, finalmente, las normas de presentación de la traducción, que serán las mismas que se aplican a la elaboración de trabajos de la asignatura en la que se realiza el experimento.

Finalmente, para terminar con la fase de preparación del experimento, pasaremos a describir el entorno web en el que se ha llevado a cabo el experimento.

4.3.1.4.1.5. Preparación del entorno web

Como hemos indicado anteriormente, este experimento se ha llevado a cabo a través de la organización de un seminario teórico-práctico en los grupos A y B de la asignatura de *Traducción especializada inglés/español-español/inglés (b): textos científico-técnicos* del último curso de la licenciatura de Traducción e Interpretación de

la Universidad de Málaga en el curso académico 2010/2011. Para la planificación y ejecución de este seminario, se contó con un espacio virtual dentro de las asignaturas virtuales alojadas en el Campus Virtual (CV) de la Universidad de Málaga⁶⁵⁸ creadas para esta asignatura, correspondientes a los dos grupos A y B.

En este espacio virtual, se pusieron a disposición de los sujetos los diferentes materiales que hemos indicado anteriormente, a saber:

- Textos para la traducción de las dos partes del experimento.
- Especificaciones del encargo de traducción.
- Cuestionarios para las dos partes del experimento.

Con respecto a estos últimos, preferimos utilizar la opción de crear cuestionarios personalizables del CV en vez de copias en papel u otra opción, a fin de recopilar los datos de los sujetos de forma más fácil y mejorar el tratamiento de los datos con posterioridad. A continuación, se observa un fragmento del cuestionario 1 que se elaboró en el CV:

⁶⁵⁸ El Campus Virtual de la Universidad de Málaga es un entorno virtual desarrollado a partir de la plataforma de enseñanza *Moodle* (*Modular Object-Oriented Dynamic Learning Environment*; en español, Entorno Modular de Aprendizaje Dinámico Orientado a Objetos), un sistema de libre distribución para la gestión de entornos de enseñanza-aprendizaje virtual, que se engloba dentro de las *Learning Management System*. Se creó en 2002 en la Universidad Tecnológica de Curtin. En la Universidad de Málaga, este CV se encuentra dividido por facultades, por lo que las asignaturas del Grado y de la Licenciatura de Traducción e Interpretación están ubicada en la página de la Facultad de Filosofía y Letras. La URL de este centro es el siguiente: <<http://filosofia.cv.uma.es/>>.

Vista Todas las respuestas (34) Ajustes avanzados Preguntas Vista previa

Cuestionario 1

•1 Indique su código:

•2 Edad:

•3 Lengua materna:

•4 ¿Qué tipo de fuente terminológica ha utilizado durante la traducción del texto facilitado? Selecciónelas.

- Diccionario/glosario monolingüe especializado (L1)
- Diccionario/glosario monolingüe especializado (L2)
- Diccionario/glosario bilingüe especializado
- Diccionario/glosario multilingüe especializado
- Diccionario visual monolingüe
- Diccionario visual bilingüe
- Sección imágenes de un motor de búsqueda (por ej. Google)
- Búsquedas en motores de búsqueda (por ej. Google)
- Tesaurus
- Corpus paralelo (textos originales y sus traducciones)
- Corpus comparable (textos originales en ambas lenguas)
- Base de datos terminológica

Figura 109. Fragmento del cuestionario 1 elaborado en el CV.

Además de estos tres materiales, se incluyó también un breve resumen del seminario, donde se indicaban los principales objetivos del seminario, un breve resumen del currículum de la ponente del seminario, la presentación realizada en la parte teórica del seminario y la lista de alumnos con la codificación que debían seguir para nombrar sus traducciones y sus cuestionarios. Por último, relacionado con la herramienta *Trandix*, se facilitó un vídeo de funcionamiento de la herramienta junto con un archivo comprimido que incluía la herramienta y los dos archivos necesarios para su utilización (un archivo en formato .txt con la traducción que debían realizar y un archivo en formato .tbx con la base de datos terminológica extraída de *ProTermino*).

De esta manera, el espacio virtual destinado a este seminario presenta el siguiente aspecto:

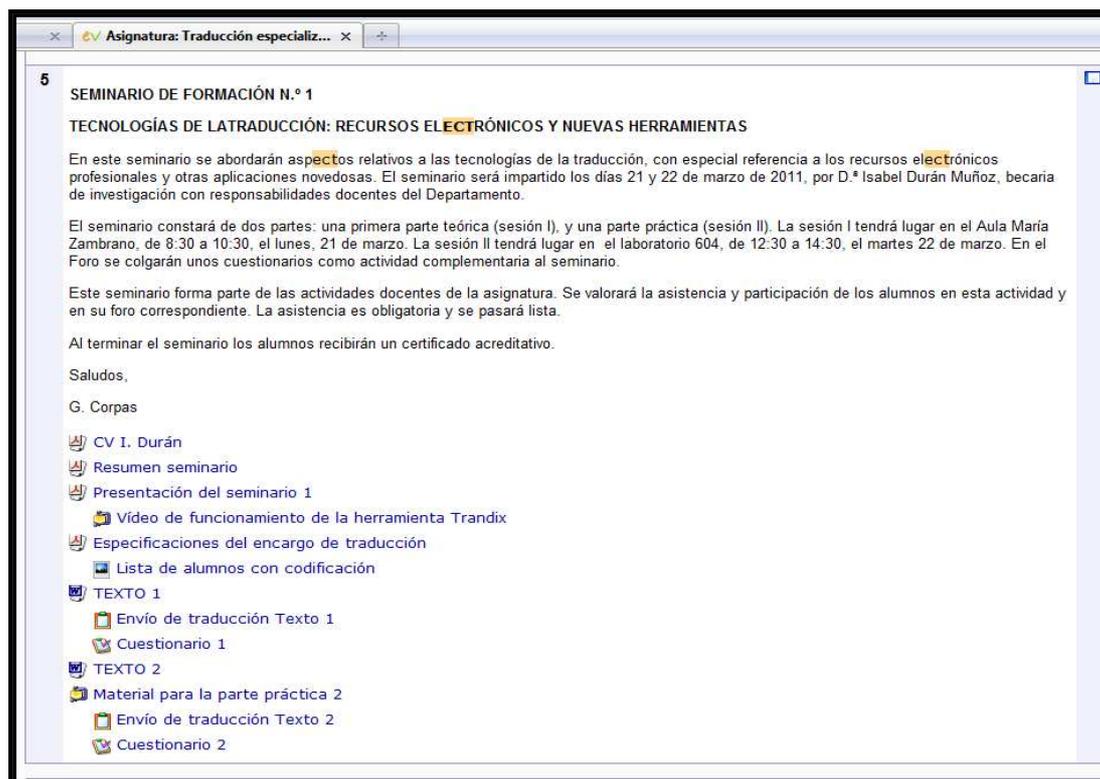


Figura 110. Espacio virtual destinado al seminario teórico-práctico.

El material contenido en este espacio no estuvo visible desde el primer momento, sino que se fueron abriendo los materiales correspondientes conforme fueron pasando las fases del experimento, tal y como veremos a continuación en la fase de ejecución del experimento.

4.3.1.4.2. Fase de ejecución

Antes de llevar a cabo el seminario, se realizaron pruebas con los materiales elaborados y necesarios para llevarlo a cabo mediante una prueba piloto, que sirvió para determinar el correcto funcionamiento de la herramienta *Trandix* en las aulas de informática donde se iba a desarrollar el experimento y la adecuación de los fragmentos para las traducciones. Asimismo, se procedió a planificar la distribución del tiempo y a concretar las distintas partes del experimento. Tras la prueba piloto, se determinaron los diferentes aspectos comprobados y se descartaron los posibles problemas de uso de *Trandix* en las aulas de informática. Así pues, se procedió a ponerlo en práctica.

La parte práctica del seminario se llevó a cabo en los respectivos horarios de clase de los dos grupos de la asignatura arriba señalada, a saber: *Traducción especializada inglés/español-español/inglés (b): textos jurídicos y socioeconómicos*; mientras que la primera parte se realizó de forma conjunta. De esta manera, los seminarios se desarrollaron con cada grupo en dos días distinto del segundo cuatrimestre del curso académico 2010-2011:

- Traducción especializada inglés/español-español/inglés (b): textos científico-técnicos (Grupo A y B): día 21 de marzo.
- Traducción especializada inglés/español-español/inglés (b): textos científico-técnicos (Grupo A): día 22 de marzo.
- Traducción especializada inglés/español-español/inglés (b): textos científico-técnicos (Grupo B): día 28 de marzo.

Estas sesiones tuvieron una duración de dos horas cada una y se llevaron a cabo en las aulas de informática de la Facultad de Filosofía y Letras, puesto que todas están equipadas con ordenadores y con conexión a Internet, a excepción de la primera sesión, que se organizó en la sala María Zambrano de la misma facultad. Durante esta primera sesión, asistieron juntos los dos grupos de la asignatura y se les proporcionó a los alumnos una introducción teórica acerca de las nuevas tecnologías de la traducción relacionadas con la competencia documental del traductor y se les facilitó información sobre los diferentes tipos de fuentes de información que se pueden encontrar en Internet y que son muy útiles para la realización de traducciones (tanto en un ámbito académico como profesional), haciendo especial hincapié en las fuentes de información terminológica y sus formas más innovadoras. Dentro de esta parte, se les presentó la herramienta *Trandix* y se les describieron sus principales funcionalidades y características y sus ventajas a la hora de emplearla en traducción. También en esta parte se les comunicó a todos los sujetos que el objetivo de los seminarios guardaba un fin investigador y se les informó acerca del experimento, sus diferentes fases y su finalidad. Asimismo, se les pidió que dieran su consentimiento a través de unos acuerdos de investigación para que sus traducciones pudiesen ser utilizadas con fines investigadores, aclarando que, por razones de confidencialidad, se mantendría el

anonimato. Finalmente, esta primera sesión finalizó con un tiempo dedicado a preguntas por parte del público asistente.

Para las otras dos sesiones se hizo una división de grupos debido al número limitado de puestos de ordenadores en las aulas de informática y, por tanto, la sesión para el grupo A del día 22 de marzo estaba organizada de la misma manera que la del grupo B del día 28 de marzo aunque en aulas diferentes. A continuación, especificamos la distribución temporal que se realizó durante estas sesiones:

- Se comenzó la sesión indicando en qué iba a consistir y explicando dónde se encontraba el material que tenían que utilizar para la ejecución del experimento, así como la codificación que debían utilizar a la hora de entregar su traducción y los cuestionarios. Esta parte introductoria duró unos 15 minutos.
- Traducción 1: Los alumnos realizaron una traducción directa inglés-español sin la ayuda de la herramienta *Trandix*, aunque con la posibilidad de consultar cualquier otro material disponible en línea. Se les concedió un tiempo máximo de 20 minutos. Tras finalizar la traducción, entregaron el documento a través de una tarea habilitada para este fin en el espacio disponible para el seminario en la asignatura virtual de la asignatura *Traducción especializada inglés/español-español/inglés (b): textos científico-técnicos del CV*.
- Después de terminar la primera traducción, se les pidió a los alumnos que rellenaran el cuestionario disponible a tal efecto. La duración de esta parte fue de unos 15 minutos.
- A continuación, se procedió a indicar a los alumnos los pasos que tenían que hacer para empezar a trabajar con la herramienta *Trandix* y se les recordó el modo de uso de la herramienta y la posibilidad de visualizar el vídeo sobre su funcionamiento. Esta parte tuvo una duración de unos 15 minutos.
- Traducción 2: Los alumnos realizaron una traducción directa inglés-español con la ayuda de la herramienta *Trandix*, así como con acceso libre a Internet. Para esta parte, se les concedió de nuevo un tiempo máximo de 20 minutos. El proceso de entrega de la traducción consistió en la entrega del documento a través de otra tarea habilitada para este fin en la asignatura virtual del Campus Virtual.

- Finalmente, se les pidió a los alumnos que rellenaran el cuestionario preparado en el experimento para esta segunda, lo que duró alrededor de 15 minutos también.

Conforme se fue avanzando en el experimento, se fueron abriendo los enlaces a la información correspondiente. De esta manera, en primer lugar solo estuvo disponible la parte relacionada con el seminario en general, es decir, el resumen y el currículum de la ponente; en segundo lugar, se permitió la visualización de la presentación del seminario y del vídeo de la herramienta *Trandix*; en tercer lugar, se abrieron los enlaces a los materiales de la primera parte práctica del seminario (especificaciones del encargo, lista de alumnos con codificación, texto 1, envío de la tarea y cuestionario 1); y, por último, se abrieron el resto de enlaces relacionados con la segunda traducción, es decir, con la traducción con la herramienta. Antes de esta parte, se recordó a los alumnos cómo funcionaba *Trandix* y los pasos que debían dar para utilizarla durante la traducción.

Durante esta fase de ejecución, no se dio ninguna incidencia, a excepción de la falta de espacio que se encontró en el primer aula de informática donde tuvo lugar la parte práctica del experimento, lo cual obligó a que algunos de los sujetos tuvieran que compartir el mismo puesto de ordenador y, por tanto, a realizar la traducción en parejas.

4.3.1.4.3. Fase de evaluación

En este último apartado describiremos los resultados obtenidos en nuestro experimento a través de la evaluación de las traducciones realizadas por los traductores semiprofesionales seleccionados para este estudio. Pero, antes de pasar a presentar el método de evaluación seguido y los resultados obtenidos, determinaremos algunos aspectos de la evaluación que pueden resultar complejos y dar lugar a confusiones, a saber: qué se entiende por evaluación y qué diferencias presenta con respecto a conceptos afines y cuáles son los métodos de evaluación que se pueden emplear en este contexto.

4.3.1.4.3.1. Evaluación de traducciones y conceptos relacionados

En este apartado, trataremos de definir qué se entiende por «evaluación de traducciones», a fin de poder concretar el trabajo que realizaremos con los resultados obtenidos en nuestro experimento. No obstante, para poder definir este concepto, nos resulta imprescindible delimitarlo con respecto a otros conceptos afines con los que puede llegar a confundirse. En este caso, procederemos a definir cuatro conceptos, a saber: la crítica, la revisión, la corrección de pruebas y la evaluación de traducciones.

De acuerdo con Horguelin (en Parra Galiano, 2005: 15), autor de la primera monografía sobre la revisión de traducciones,⁶⁵⁹ podemos intervenir en un texto de formas diferentes en función de la fase en la que este se encuentre:

- Si se trata de un producto semiacabado o borrador de texto, la intervención puede consistir en una *revisión* o en una *corrección de pruebas*.
- Si se trata de un producto acabado, la intervención puede consistir en una *evaluación* o en una *crítica*.

Partiendo de esta clasificación, realizaremos una presentación breve de estos conceptos, que, como hemos observado, se diferencian principalmente por su fase de actuación. De esta manera, vemos que los conceptos *revisión* y *evaluación*, que a menudo dan lugar a confusiones, quedan claramente delimitados por esta taxonomía.

En primer lugar, trataremos el concepto de *revisión*, que se podría definir de la siguiente manera:

Una lectura atenta de un texto de llegada (TL), considerado como producto semiacabado o borrador de traducción, que realiza la misma persona que ha traducido el texto (traductor) o un tercero (revisor), generalmente, comparándolo con el texto de origen (TO) y utilizando para ello una serie de criterios establecidos *a priori*, con la finalidad de comprobar si se han cumplido las especificaciones del encargo de traducción y realizar las oportunas correcciones y mejoras en el TL, antes de proceder a su entrega al cliente (Parra Galiano, 2005: 18).

Por lo general, la figura del revisor, ya sea mediante una autorrevisión o un revisor diferente, recae en un traductor con la competencia adecuada en las lenguas de origen y

⁶⁵⁹ Esta obra es *Practique de la révision*, y fue publicada en 1978.

de destino que examina el texto traducido para garantizar que el texto cumple con el objetivo previsto en cuanto a la terminología, corrección gramatical y de convenciones, el registro, el estilo, la función y la presentación requerida. En definitiva, el revisor trata de corregir, mejorar o enmendar el texto de llegada con objeto de alcanzar un resultado exento de errores y acorde con las directrices establecidas en el encargo antes de entregárselo al cliente.

Dentro de la revisión, se pueden distinguir tres tipos diferentes según la finalidad que se persigue: la revisión pragmática, la revisión formativa y la revisión pedagógica (Horguelin, 1985: 10; Brunette, 2000: 174; Parra Galiano, 2005: 20-21). Como decíamos, cada una de ellas hace referencia a un objetivo diferente, a saber:

- Revisión pragmática: comparación exhaustiva del texto meta con el texto origen, con la finalidad de mejorar la traducción, sin necesidad de que el revisor tenga que consultar al traductor ni justificar ante él modificaciones o correcciones realizadas.
- Revisión formativa: revisión que se practica en el ámbito profesional con la finalidad de corregir y de mejorar la traducción, además de perfeccionar la técnica del traductor que la ha realizado.
- Revisión didáctica: ejercicios de revisión que se realizan en el ámbito académico como parte del aprendizaje del estudiante de traducción.

El segundo concepto que trata con el producto semiacabado como la revisión es la corrección de pruebas, la cual constituye el último control de calidad al que se somete un texto antes de pasar a su publicación definitiva. Esta tarea se asemeja a la anterior por lo que respecta al objetivo común: mejorar la traducción en cuanto a terminología, estilo, registro, etc. y detectar cualquier posible error, pero se distinguen en un aspecto básico: el corrector no tiene competencia de efectuar modificaciones en el texto, sino de señalarlas según una serie de signos de corrección acordados y devolverlas al revisor, al autor original o al traductor para que este proceda a las oportunas correcciones y mejoras en el texto (Parra Galiano, 2005: 43).

Según Martínez de Sousa (1999: 188-191), la corrección de pruebas puede ser principalmente de tres tipos, a saber: corrección de estilo, corrección de concepto y corrección tipográfica.

- Corrección de estilo: la revisión literaria de la traducción con objeto de garantizar la calidad requerida, tanto desde un punto de vista lingüístico, gramatical y ortográfico, como desde la semántica y el léxico.
- Corrección de concepto: la revisión monolingüe de la traducción mediante la lectura atenta del texto por un experto en la materia en cuestión encargado de atender a la terminología, a la adecuación del lenguaje, al destinatario y al fondo del asunto y de recomendar las correcciones pertinentes, en su caso.
- Corrección tipográfica: la revisión de las pruebas de imprenta para señalar erratas ortográficas y tipográficas del texto, así como otros desarreglos que se producen al convertir los formatos de los textos.

Como decíamos, en la corrección la competencia de modificar el texto no recae en el corrector, sino en el autor original, el revisor o el traductor. El corrector únicamente señala los errores, indica recomendaciones de modificación, etc. en el texto y devuelve el texto al sujeto correspondiente en cada caso.

Una vez que hemos presentado las técnicas que se basan en el producto semiacabado, pasaremos a describir las técnicas sobre productos acabados, entre las que se encuentra la evaluación.

En primer lugar, trataremos la *crítica de traducciones*, también denominada crítica traductológica o crítica translatoria, una técnica que se considera también afín a la técnica de evaluación y que, de hecho, puede entenderse en su acepción más general como sinónimo de evaluación (Parra Galiano, 2005: 29). La mayoría de los estudios que abordan esta técnica tratan la crítica de traducciones desde dos puntos de vista diferentes: por un lado, como veíamos anteriormente, en un sentido amplio en el que la asemejan a la evaluación de traducciones; y, por otro, en un sentido más concreto en el que la consideran un concepto diferente que hace referencia a la valoración de la traducción de textos literarios o, en una aproximación más general, de textos publicados (Martínez Melis, 1997; Waddington, 2000; Martínez Melis y Hurtado Albir, 2001). En

este sentido, la *crítica de traducciones* se entendería como la valoración razonada que surge de la comparación de un texto de partida con su traducción o sus diferentes versiones traducidas y publicadas. Como señala Hurtado Albir (1990), el objetivo de este tipo de evaluación es el de juzgar una traducción, debatir sus pros y sus contras y, en algunas ocasiones, proponer soluciones. En nuestro caso, consideraremos la técnica de crítica de traducciones como se define en su segunda acepción, ya que nos parece oportuno realizar una distinción con la técnica de evaluación y tener en cuenta, por tanto, el hecho de que sus objetivos no son totalmente compartidos, aunque el procedimiento pueda ser similar.⁶⁶⁰

Finalmente, llegamos a la técnica de *evaluación*, que, como hemos venido viendo hasta el momento difiere, por una parte, de la *revisión* y de la *corrección de traducciones* por lo que respecta al producto de base (texto semiacabado o acabado) y, por otra, de la *crítica de traducciones*, por basarse esta en textos publicados, principalmente literarios. Asimismo, además de la distinción con base en el producto objeto de interés y del tipo de texto analizado, la evaluación conlleva un juicio de valor, al que se llega aplicando criterios cuantitativos o cualitativos. Para que una evaluación sea lo más objetiva posible, es necesario que esté basada en una plantilla de evaluación, la cual, además de aportar la objetividad necesaria, permita establecer *a priori* los rasgos de la traducción que se van a evaluar para todos los posibles evaluadores que intervienen en el proceso, permita seguir un proceso homogéneo y justo para todos los evaluados y, además, ahorre tiempo a los evaluadores. Estas plantillas pueden ser de dos tipos básicamente: por un lado, plantillas analíticas, cuyo objetivo consiste en analizar la traducción atendiendo a distintos parámetros (léxicos, sintácticos, morfológicos, semánticos, de sentido, etc.) que han sido categorizados previamente y ofrecer una valoración final basada en el cómputo de los errores y aciertos encontrados en el texto; y, por otro, plantillas holísticas (o rúbrica), que consisten en la evaluación global del texto, sin prestar atención a errores concretos, para clasificarlo según una tabla de niveles de calidad (de excelente a pésimo, por ejemplo) diseñada previamente. En las siguientes páginas, trataremos con mayor detenimiento estos dos tipos de plantillas y presentaremos la utilizada para la evaluación de nuestro experimento.

⁶⁶⁰ Para profundizar en las técnicas de crítica de traducciones, remitimos a Parra Galiano (2005: 37 y ss.).

4.3.1.4.3.2. Métodos de evaluación

Evaluar cualquier tipo de texto, en nuestro caso una traducción, implica necesariamente determinar la calidad del texto que se ha de evaluar. El problema estriba en determinar qué se entiende por una buena traducción y, por tanto, qué se considera una traducción “de buena calidad”, teniendo en cuenta que este concepto depende del género textual y de su contexto histórico.⁶⁶¹ En este contexto actual, donde impera la visión funcionalista de la traducción, es decir, la función que presenta el texto original y su traducción en una situación comunicativa determinada (con una finalidad, unos destinatarios, etc.) cabe recordar la siguiente cita de Sager (1985: 121):

There are no absolute standards of translation quality but only more or less appropriate translations for the purpose for which they are intended.

De esta manera, queda patente que una traducción de calidad se considera una traducción que cumple la función que se ha establecido en su encargo y que se adecúa, por tanto, a la situación en la que se encuentra, a las expectativas de los usuarios, a las instrucciones del cliente y al código lingüístico utilizado en cada caso.

En este contexto, y después de haber consultado los criterios de evaluación propuestos por diferentes autores de renombre (Nobs, 2005; Collados Aís, 1998; House, 1997; Chesterman, 1993; Sager, 1989; Reiss, 1986), podemos establecer que los criterios más eficaces para medir la calidad de forma general son los siguientes:

⁶⁶¹ Desde la Antigüedad hasta la actualidad ha existido una dicotomía acerca de la manera de traducir: traducción literal o traducción libre, lo que ha provocado una sucesión de periodos en los que imperaba una forma u otra y, por tanto, en los que se consideraba como una traducción “de calidad” un tipo u otro de traducción. En el Imperio romano, con la figura de Cicerón, se defendía la traducción libre («no hay que traducir *verbum pro verbo*» [Hurtado 2001:105]), al igual que ocurría durante la Edad Media en Occidente, donde aparecieron incluso contribuciones como las cinco reglas de Dolet (en Hurtado 2001:108-109), que pueden considerarse válidas para este periodo actual: El traductor debe 1. comprender el sentido y el tema; 2. conocer la lengua del autor y la lengua a la que traduce; 3. evitar traducir palabra por palabra; 4. no emplear palabras demasiado próximas al latín; y 5. procurar la armonía del discurso. También durante el siglo XVII se extiende por Europa una manera de traducir surgida en Francia, basada en la modificación del texto en pro de la adaptación, de la distancia lingüística y cultural, que se conoce como las *belles infidèles*. Sin embargo, en la segunda mitad del siglo XVII surge una corriente que exige mayor exactitud y fidelidad al traductor y una traducción palabra por palabra. Esta corriente se repetirá durante el siglo XIX, mientras que en el siglo xx se vuelve a asistir a un giro de perspectiva. En este periodo comienza la corriente funcionalista, que defiende que la calidad de la traducción depende de su adaptación al contexto y a su función como texto (aunque siempre respetando el original), es decir, se pasa a considerar la calidad de la traducción según su adecuación funcional a la situación comunicativa y al contexto.

adecuación pragmática y funcional, legibilidad, corrección gramatical y comportamiento profesional (Durán Muñoz, 2008a).⁶⁶²

En el caso de la adecuación pragmática, se hace referencia principalmente al contexto y a los usuarios. En este caso, la traducción debe ajustarse a las expectativas de los usuarios en cuanto a contenido y forma, es decir, la información que se ofrece debe ser útil y accesible para un receptor del texto meta, debe ajustarse a las normas y convenciones de la cultura meta, debe ofrecer un estilo y un registro normal en la cultura de llegada para ese tipo textual y debe contar con la terminología y fraseología propias del campo de especialidad en cuestión.

En segundo lugar, la adecuación funcional está relacionada con el criterio anterior, aunque se basa en las funciones que desempeña el texto meta. El traductor debe adecuar también el contenido y la forma a los usuarios pero, en esta ocasión, con respecto a la función que debe desempeñar el texto traducido en la cultura meta. En otras palabras, el traductor debe tener en cuenta las funciones características de un determinado género textual en la cultura meta a fin de conseguir la adecuación funcional en el texto de llegada, de la misma manera que se ha cumplido en el texto origen. En algunos casos, las funciones del texto origen y del texto meta pueden variar, pero esto deberá estar recogido en las especificaciones del encargo de traducción.

El criterio de legibilidad hace referencia a la claridad y la comprensibilidad del texto meta. Para cumplir este criterio, el traductor deberá conseguir que su traducción sea clara y comprensible para sus receptores y, por tanto, deberá perseguir la coherencia y cohesión necesarias para que esto se produzca. Ante la carencia de estos aspectos, la traducción de cualquier tipo textual provocará una incomprensión parcial o, incluso, total por parte de los usuarios finales.

Con respecto a la corrección gramatical, es evidente que toda traducción debe ajustarse a las normas gramaticales, ortográficas y ortotipográficas de la lengua de llegada, por lo que una carencia de ese seguimiento provocará una falta de calidad absoluta en cualquier texto traducido y, por tanto, una evaluación negativa.

⁶⁶² Esta autora también incluye el tratamiento no verbal como criterio de evaluación al trabajar en el contexto del discurso turístico, aunque para establecer unos criterios generales no nos parece pertinente incluirlo.

Por último, el criterio de comportamiento profesional hace referencia principalmente al cumplimiento por parte del traductor de los aspectos acordados en el encargo, a saber: el tiempo, el presupuesto de la traducción y las herramientas utilizadas (normas, glosarios, memorias de traducción, etc.).

Con objeto de evaluar la adecuación de las traducciones con respecto a estos criterios, encontramos sistemas de evaluación de diferente naturaleza, que principalmente se pueden dividir en dos tipos: por un lado, sistemas analíticos y, por otro, sistemas holísticos. Según Chamosa (1997: 36), el primero tiene en cuenta «el grado de corrección con que el traductor ha sido capaz de resolver los problemas técnicos de trasvase de códigos lingüísticos» mientras que el segundo pretende «emitir un juicio de valor sobre la calidad del producto terminado, en tanto que forma parte del sistema cultural de la lengua de llegada». Asimismo, podemos encontrar un tercer modelo que combina ambos sistemas y que se conoce como sistema mixto, el cual permite reducir la posible subjetividad y, por tanto, garantizar la fiabilidad del sistema de evaluación utilizado (Waddington, 2000; Beeby, 2000).⁶⁶³ En estos modelos, especialmente en el primero, el error constituye la clave de la evaluación, ya que esta irá directamente vinculada a la presencia o ausencia de errores.⁶⁶⁴

Como veíamos, el primer sistema indicado, el sistema analítico, se centra en la corrección lingüística. En la mayoría de estos sistemas se distinguen dos fases principalmente: una inicial, en la que se categorizan los fenómenos encontrados en el texto, y una segunda en la que dichos fenómenos reciben una valoración, normalmente numérica (o convertible a números), que sustenta el cómputo que da lugar al juicio global de la traducción. Con objeto de que ese resultado final sea lo más objetivo, fiable y claro posible se suele emplear un conjunto de instrumentos de medición de errores de diferente naturaleza, entre los que destacan las plantillas de evaluación.⁶⁶⁵ De este modo, se pretende reducir la subjetividad de la evaluación, intrínseca en cualquier

⁶⁶³ Para profundizar en los modelos mixtos que se han aplicado en investigaciones recientes, remitimos a Conde Ruano (2009: 197 y ss.).

⁶⁶⁴ Como veíamos en el capítulo segundo, el «error» se define como «una equivalencia inadecuada para la tarea traductora encomendada» (Hurtado Albir, 2001: 289), es decir, como una ruptura de las reglas de coherencia de un texto meta, sean estas de tipo lingüístico (gramatical, terminológico, de congruencia y precisión semántica, etc.) o extralingüísticas (cultural, funcional y de registro).

⁶⁶⁵ Además de las plantillas de evaluación, también se utilizan otros instrumentos como tests de respuesta múltiple, cuestionarios, entrevistas personales, diarios, etc. (cf. Delisle, 1993; Hurtado Albir, 1999).

actividad humana. Actualmente existen numerosas plantillas de evaluación que contienen diferentes clasificaciones de errores según la aproximación que realice el autor y según la exhaustividad de la plantilla, aunque la mayoría coinciden en recoger los errores más frecuentes encontrados en traducción, a saber: errores de comprensión y transmisión del sentido original, errores de expresión en la lengua de llegada (errores gramaticales, ortotipográficos, terminológicos, de coherencia, etc.) y errores por inadecuación funcional y adaptación cultural. Asimismo, cada vez es más común encontrar en estas plantillas apartados que puntúan positiva y negativamente las decisiones de los traductores, es decir, no solo los errores de traducción reciben una puntuación (en este caso, negativa), sino también los aciertos ante alguna dificultad encontrada (en este caso, positiva).⁶⁶⁶

A continuación, presentamos como ejemplo de plantilla de evaluación analítica la plantilla propuesta por Toledo Báez (2009), una plantilla muy completa en la que se observan los diferentes criterios de evaluación más utilizados en este tipo de sistemas.⁶⁶⁷

⁶⁶⁶ Para consultar diferentes plantillas de evaluación, remitimos a Waddington (2000), Conde Ruano (2009) y Toledo Báez (2009).

⁶⁶⁷ Esta plantilla tiene como base principal la plantilla de evaluación propuesta por Hurtado Albir (1995), aunque también toma elementos del baremo de corrección del proyecto Multilingual eLearning in LANGuage Engineering (MeLLANGE), un proyecto Leonardo da Vinci desarrollado desde 2004 a 2007 cuyo objetivo principal consistió en adaptar la formación de los traductores y otros profesionales lingüísticos a las nuevas necesidades surgidas a raíz del proceso de globalización actual. URL: <<http://mellange.eila.jussieu.fr/>>.

| PLANTILLA DE EVALUACIÓN ANALÍTICA |
|--|
| <p>1. Inadecuaciones que afectan a la comprensión del texto original</p> <p>1.1. Contrasentido</p> <p>1.2. Falso sentido</p> <p>1.3. Sin sentido</p> <p>1.4. Adición innecesaria de información</p> <p>1.5. Supresión innecesaria de información</p> <p>1.6. No mismo sentido</p> <p>1.7 Inadecuación de variación lingüística</p> <ul style="list-style-type: none"> - Inadecuación de registro lingüístico (formal/informal) - Inadecuación de estilo |
| <p>2. Inadecuaciones que afectan a la expresión en la lengua de llegada</p> <p>2.1. Ortografía</p> <p>2.2. Puntuación</p> <p>2.2. Gramática (sintaxis, género, número, verbo)</p> <p>2.3. Léxico</p> <ul style="list-style-type: none"> - Barbarismos, calcos - Usos inadecuados (registro, regionalismos) - Empleo de terminología no adecuada <p>2.4. Textual</p> <ul style="list-style-type: none"> - Incoherencia, falta de lógica - Mal encadenamiento discursivo, uso indebido de conectores <p>2.5. Estilística</p> <ul style="list-style-type: none"> - Formulación no idiomática - Formulación defectuosa - Muy literal - Demasiado libre |
| <p>3. Inadecuaciones funcionales</p> <ul style="list-style-type: none"> - Inadecuación a la función textual prioritaria del original - Inadecuación a la función de la traducción |
| <p>4. Aciertos</p> <ul style="list-style-type: none"> - Buena equivalencia - Muy buena equivalencia |

Tabla 53. Plantilla analítica en Toledo Báez (2009).

Por último, también es posible que las plantillas de evaluación contengan una clasificación interna de errores basada en errores graves y errores leves mediante la cual otorguen una mayor o menor puntuación negativa según el tipo de error detectado.

| TIPO DE ERROR | CATEGORÍA | PENALIZACIÓN |
|----------------------|-----------|--------------|
| Sintaxis | Leve | -0,25 |
| | Grave | -0,5 |
| Precisión/coherencia | Leve | -0,25 |
| | Grave | -0,5 |
| Terminología | Leve | -0,25 |
| | Grave | -0,5 |
| Registro | Leve | -0,25 |
| | Grave | -0,5 |
| Ortografía/formato | Leve | -0,25 |
| | Grave | -0,5 |
| Ortografía | - | -1,00 |
| Omisión | - | -1,00 |
| Sentido | - | -1,00 |

Tabla 54. Plantilla analítica del proyecto de innovación educativa Diseño de una plantilla electrónica para la evaluación y revisión de proyectos de Traducción especializada según las normas del sector.⁶⁶⁸

El segundo tipo de sistema son los sistemas holísticos (o de evaluación global). Estos, al contrario que los anteriores, no se basan en aspectos concretos de la traducción (errores lingüísticos, de sentido, etc.) sino en la adecuación de la traducción de forma general. Estos modelos defienden el hecho de que, en muchas ocasiones, el resultado obtenido mediante una evaluación analítica no se corresponde con lo que el evaluador siente (McAlester, 2000: 235), es decir, con la imagen de calidad que el propio evaluador se ha formado de la traducción. De esta manera, los autores han propuesto modelos basados en niveles más o menos exhaustivos que indican la calidad de la traducción en función de una serie de criterios. A continuación, presentamos como ejemplo el modelo holístico de De Rooze (2003: 54), que tiene una estructura bastante sencilla:

⁶⁶⁸ El proyecto de innovación educativa Diseño de una plantilla electrónica para la evaluación y revisión de proyectos de Traducción especializada según las normas del sector (N.º ref. PIE07/049. Inv. Principal: G. Corpas Pastor) fue concedido por la Dirección de Enseñanza Virtual y Laboratorios Tecnológicos de la Universidad de Málaga para los años 2007-2008. Tenía como objetivo la creación y desarrollo de una plantilla para la revisión y evaluación de traducciones especializadas que reflejara la situación actual del mercado laboral de la traducción profesional.

| NIVEL | EXPLICACIÓN |
|-------------------------|--|
| 1. Muy deficiente: | Tarea muy incompleta. Transmisión muy defectuosa; no valdría la pena revisar. Manifiesta incapacidad para expresarse bien en español. |
| 2. Mala/Deficiente: | Tarea a la mitad. Transmisión desvirtuada; revisión profunda para ser profesional. Casi todo suena a traducción. |
| 3. Aceptable/Buena: | Casi completa. Transmisión casi completa, revisión media para ser profesional. Gran parte del texto se lee como un original. |
| 4. Muy buena/Excelente: | Tarea completa. Transmisión completa de la información relevante; revisión mínima para ser profesional. Se lee como un original. Puede tener variantes estilísticas no preferidas pero aceptables. |

Tabla 55. Plantilla holística de De Rooze (2003).

Otros modelos, en cambio, presentan mayor complejidad y un sistema de puntuación con valores positivos y negativos, como el que presenta Toledo Báez (2009: 590).

| NIVEL | CALIDAD DE TRANSMISIÓN | CALIDAD DE EXPRESIÓN | PUNTUACIÓN |
|---------|--|---|------------|
| Nivel 5 | Transmisión completa de información del texto origen; revisión mínima (matices). El traductor funciona igual que un traductor profesional con experiencia. | Todo o casi todo se lee como un original en la lengua meta; errores mínimos de terminología, de gramática u ortografía. | +5 |
| Nivel 4 | Transmisión casi completa, alguna omisión no significativa; revisión mediana. Puede trabajar sin supervisión. | Gran parte se lee como un original en la lengua meta; algún error aislado de terminología, gramática y ortografía. | +3 |
| Nivel 3 | Transmisión de la(s) idea(s) general(es) pero con lagunas e imprecisiones; revisión a conciencia. Puede trabajar independientemente. | Algunos fragmentos se leen como un original en la lengua meta, pero otros suenan a traducción; bastantes errores de terminología, gramática y ortografía. | +1 |
| Nivel 2 | Transmisión desvirtuada por omisiones o interpretaciones erróneas; revisión muy profunda. El traductor necesita supervisión. | Casi todo suena a traducción; múltiples errores de terminología, gramática y ortografía. | 0 |
| Nivel 1 | Transmisión muy defectuosa; no valdría la pena revisar. La traducción es incoherente en su mayor parte. | Manifiesta incapacidad para expresarse adecuadamente en la lengua meta. | -1 |

Tabla 56. Plantilla holística de Toledo Báez (2009: 590).

Una vez que hemos presentado los principales sistemas de evaluación y hemos ofrecido algunos ejemplos de plantillas de dichos sistemas, continuaremos con la presentación de nuestra plantilla mixta de evaluación, que hemos utilizado para realizar la evaluación de la calidad de las traducciones realizadas en nuestro seminario práctico.

4.3.1.4.3.3. Propuesta de plantilla mixta de evaluación

Desde nuestro punto de vista, un buen método de evaluación debe prestar atención tanto a los aspectos concretos de la traducción mediante una evaluación analítica como a la calidad global que presenta la traducción que se está evaluando. Sólo de esta forma se alcanzará una evaluación válida, objetiva y fiable, reduciendo así la subjetividad. Siguiendo esta línea, nuestro método de evaluación consiste en una combinación de los dos sistemas de evaluación, es decir, se trata de un sistema mixto que contiene elementos propios de un sistema analítico y un sistema holístico. Asimismo, presenta otras características específicas:

- Establece puntos positivos y negativos.
- Diferencia entre faltas graves y leves y, por tanto, distinguen dos puntuaciones negativas diferentes: 0,25 para las leves y 0,5 para las graves.⁶⁶⁹

Para llevar a cabo la propuesta de este sistema mixto, nos hemos basado en las plantillas de evaluación citadas anteriormente y la hemos adaptado a nuestras necesidades, eliminando y añadiendo criterios que nos parecen más adecuados para la evaluación de traducciones. En general, hemos intentado simplificar las plantillas de Toledo Báez, puesto que consideramos que una evaluación tan exhaustiva resulta a menudo tediosa, larga e innecesaria, a la par que hemos completado las otras plantillas presentadas (plantilla del proyecto de Innovación educativa PIE07/049 y de De Rooze, 2003), con objeto de conseguir una evaluación completa y precisa.

De esta manera, el resultado final de la plantilla de evaluación analítica es el siguiente:

⁶⁶⁹ En caso de que se encontrara un error que se considerara muy grave se puntuaría con -1,00 punto.

| TIPO DE ERROR | CATEGORÍA | PENALIZACIÓN |
|--|----------------|--------------|
| Gramática (sintaxis, género, número, etc.) | Leve (GRAM) | -0,25 |
| | Grave (GRAM+) | -0,5 |
| Terminología | Leve (TERM) | -0,25 |
| | Grave (TERM+) | -0,5 |
| Registro | Leve (REG) | -0,25 |
| | Grave (REG+) | -0,5 |
| Ortotipografía | Leve (ORTO) | -0,25 |
| | Grave (ORTO+) | -0,5 |
| Puntuación | Leve (PUNT) | -0,25 |
| | Grave (PUNT+) | -0,5 |
| Estilo | Leve (EST) | -0,25 |
| | Grave (EST+) | -0,5 |
| Coherencia y cohesión textual (COH) | - | -0,5 |
| Ortografía (ORT) | - | -0,5 |
| Omisión (OMS) | - | -1,00 |
| Sentido (SENT) | - | -1,00 |
| Función textual (FUNC) | - | -1,00 |
| Aciertos | Bien (BIEN) | +0,5 |
| | Muy bien (EXC) | +1,00 |

Tabla 57. Plantilla analítica propia.⁶⁷⁰

Por su parte, la evaluación holística se realizaría también otorgando puntuaciones positivas y negativas en función de la calidad global de la traducción. A continuación, presentamos nuestra propuesta, basada en la plantilla de Toledo Báez (2009) y De Rooze (2003):

⁶⁷⁰ Junto a cada criterio de evaluación se observa una abreviatura, la cual será utilizada para indicar los errores y aciertos en la traducción.

| NIVEL | EXPLICACIÓN | PUNTUACIÓN |
|---------|--|------------|
| Nivel 4 | Transmisión completa de información del texto origen; revisión mínima y sobre todo con recomendaciones. Todo o casi todo se lee como un original en la lengua meta. Exenta prácticamente de errores. | +3 |
| Nivel 3 | Transmisión casi completa; revisión media. Gran parte se lee como un original; algún error aislado de terminología, gramática y ortografía. | +1 |
| Nivel 2 | Transmisión desvirtuada por omisiones o interpretaciones erróneas; revisión muy profunda. Casi todo suena a traducción y hay múltiples errores de terminología, gramática y ortografía. | 0 |
| Nivel 1 | Transmisión muy incompleta y defectuosa; no valdría la pena revisar. Manifiesta incapacidad para expresarse bien en español. | -1 |

Tabla 58. Plantilla holística propia.

Una vez presentadas las dos plantillas de evaluación, pasaremos a evaluar la calidad de las traducciones del experimento que se ha llevado a cabo en el seminario organizado para tal efecto, con objeto de determinar la validez de la herramienta *Trandix* con respecto al tiempo de trabajo como a la calidad alcanzada de las traducciones.

4.3.1.4.3.4. Resultados del experimento

En este apartado, mostraremos los resultados obtenidos en nuestro experimento, así como los criterios utilizados para la recogida de datos cuantitativos que nos permitirán confirmar o descartar la hipótesis de que la herramienta *Trandix* constituye un recurso terminográfico de gran ayuda a la traducción.

En primer lugar, debemos indicar que el número total de traducciones obtenidas tras los seminarios asciende a un total de 61 traducciones, de las cuales 12 tuvieron que realizarse en parejas debido a las limitaciones de espacio de la primera aula asignada para realizar el experimento, como hemos visto anteriormente. Todas estas traducciones son anónimas por motivos de confidencialidad, aunque se les ha asignado un código identificativo para vincular los cuestionarios realizados con sus respectivas traducciones.

Una vez presentado el número de traducciones, pasaremos a presentar los criterios de carácter cuantitativo utilizados para medir la aplicación de *Trandix* en la traducción:

- Criterio de finalización de la tarea, que se tendrá en cuenta si se ha finalizado o no la traducción encargada en el tiempo dado para ello, en este caso 20 minutos.
- Criterio de tiempo de realización, que se observará si la finalización de la tarea y la subida del texto se ha producido antes de la finalización del tiempo dado para la traducción (20 minutos).
- Criterio de calidad, que se determinará mediante la aplicación de la plantilla mixta de evaluación analítica y holística.

El primer criterio indicado, el relacionado con la finalización de la tarea, nos servirá de filtro para la aplicación del resto de criterios, es decir, si la tarea está finalizada pasará a evaluarse por el criterio de calidad pero si no lo está, se desechará y, por tanto, no se tendrá en cuenta para su posterior evaluación. Partimos de la base de que si no se ha terminado la tarea encomendada en el tiempo dado, dicha tarea presenta una calidad baja y, por tanto, exenta de valor en la evaluación.

El criterio relativo al tiempo de realización de la traducción se considera un criterio útil que tiene en cuenta si algún alumno es capaz de terminar su traducción antes del tiempo dado para la realización de las tareas (en este caso, 20 min.). Si se recogieran casos de este tipo, se recopilaría información acerca de la rapidez de traducción de los alumnos con y sin la herramienta *Trandix*.

Por último, el tercer criterio está relacionado con la calidad de las traducciones y, por tanto, se emplearán la plantilla mixta diseñada en el apartado anterior para evaluar las diferentes traducciones de un modo analítico y holístico. En este criterio se realizará también un estudio acerca de la calidad de las traducciones con relación al tiempo utilizado para realizarlas con objeto de determinar si el hecho de haber terminado antes del tiempo proporcionado para realizar la traducción actúa en beneficio o, por el contrario, en detrimento de la calidad ofrecida. En otras palabras, comprobaremos si la calidad de la traducción aumenta, disminuye o se mantiene si se realiza de forma más rápida. Por un lado, puede ocurrir que se cuente con más tiempo

para revisar la traducción y, por tanto, para corregir cualquier posible error detectado; pero por otro, podría ser que la velocidad diera lugar a más errores.

Además de estos tres criterios cuantitativos, contamos con un cuarto criterio basado en la satisfacción de los sujetos durante la utilización de la herramienta *Trandix* y su comparación con respecto al protocolo de trabajo sin dicha herramienta. Este criterio de satisfacción se aplica a través de los cuestionarios elaborados para el experimento y tiene naturaleza cualitativa, puesto que cuenta con preguntas de opinión, aunque también tiene una naturaleza cuantitativa, puesto que hay preguntas de valoración del 1 al 5, entre otras.

En los apartados siguientes desarrollaremos cada uno de los criterios arriba señalados y aportaremos los datos recogidos de cada criterio.

4.3.1.4.3.4.1. Criterio de finalización de la tarea

Como decíamos anteriormente, este criterio es el primero que se aplica en esta evaluación, ya que sirve de filtro a la hora de considerar si una traducción debe ser evaluada por su calidad o no, es decir, si una traducción no se ha finalizado en el tiempo dado para tal efecto se obviará. Por el contrario, si una traducción se ha realizado en el tiempo establecido, pasará a ser evaluada por su calidad, el tercer criterio de aplicación. Este criterio de finalización de la tarea nos indica un aspecto muy importante con respecto al uso o no de la herramienta *Trandix*, que consiste en confirmar la gran diferencia existente entre los resultados obtenidos con las traducciones realizadas sin la herramienta y con la herramienta, como exponemos a continuación.

En la primera tarea encomendada, los sujetos realizaron la traducción del texto sin la herramienta *Trandix* y, como resultado, obtuvimos que solo el 14,75% de los sujetos consiguieron entregarla completa al 100%. Por el contrario, la segunda tarea, esta vez con la herramienta, ocurrió totalmente lo contrario: solo el 13,1% la entregaron sin finalizar al 100% mientras que el resto, el 86,9%, la entregó completamente terminada. Teniendo en cuenta la gran diferencia entre el número de traducciones finalizadas para la tarea 1 y la tarea 2, decidimos que era conveniente contabilizar también aquellas traducciones completadas en un 80% del total, con el fin de no

descartar la mayoría de las traducciones realizadas durante la primera tarea y, poder, así, establecer comparaciones con respecto a la calidad obtenida en las traducciones de ambas tareas.

Como resultado de esta ampliación del análisis, observamos que el 34,4% de las traducciones de la primera tarea habían sido completadas, de las cuales el 14,75% correspondían a traducciones finalizadas y el resto a traducciones completadas en un 80%.

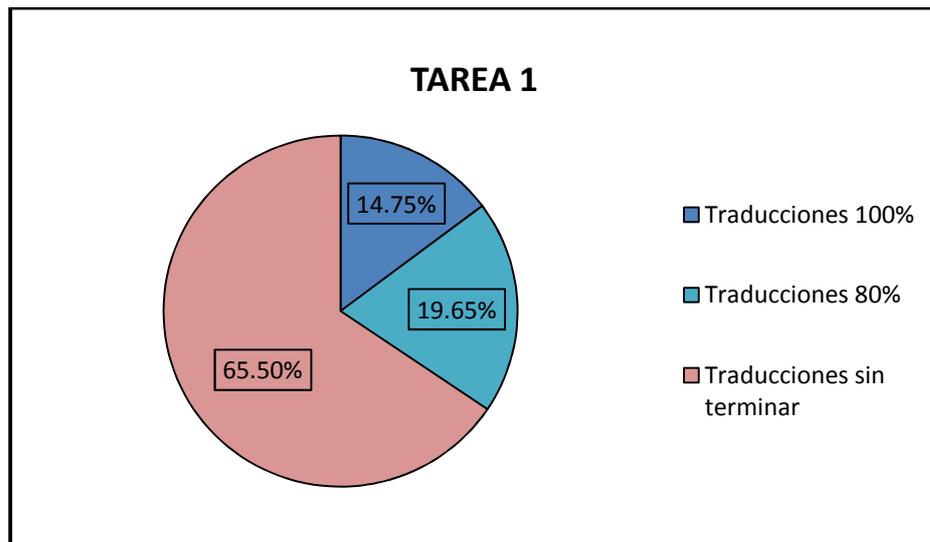


Gráfico 11. Traducciones terminadas y no terminadas para la tarea 1.

Por su parte, en la segunda tarea se contabilizaron el 96,6% de las traducciones realizadas con la herramienta *Trandix*, de las cuales el 88,3% eran traducciones totalmente finalizadas y 8,3 completas al 80%. En otras palabras, ocurrió lo contrario de lo que había sucedido en la primera tarea, ya que la gran mayoría finalizó la tarea a tiempo y solo el 3,4% de los sujetos no tuvo suficiente tiempo para completarla.

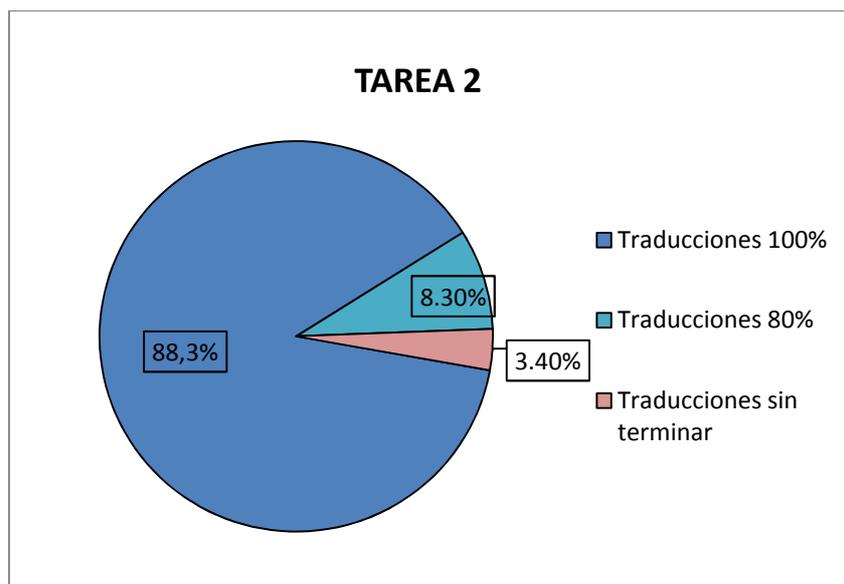


Gráfico 12. Traducciones terminadas y no terminadas para la tarea 2.

Este criterio nos permite observar las diferencias con respecto al tiempo que existen entre la traducción asistida con la herramienta *Trandix* y sin ella. Gracias a este experimento, podemos extraer como conclusión el hecho de que la ayuda terminológica que ofrece nuestra herramienta, junto con sus opciones de búsqueda y flexibilidad, permiten reducir el tiempo de traducción de un texto gracias al acceso rápido y preciso a la terminología y otros recursos externos, lo que supone un aumento del número de palabras traducidas en el mismo tiempo. Teniendo en cuenta que los traductores profesionales trabajan en función de número de palabras traducidas, consideramos que esta conclusión resulta esencial para demostrar la validez de *Trandix* en este sentido.

4.3.1.4.3.4.2. Criterio de tiempo de realización

Como indicábamos anteriormente, este criterio consiste en medir el tiempo que tardan los alumnos en caso de que estos terminen antes del tiempo dado para tal efecto. Cabe destacar que ningún alumno finalizó su traducción antes del tiempo proporcionado cuando se encontraban trabajando en la primera tarea (sin la herramienta). De hecho, como anécdota, podemos añadir que todos los sujetos quedaron sorprendidos cuando se les informó de la finalización del tiempo dado. Por el contrario, hubo un elevado número de sujetos que terminaron la segunda tarea (con la herramienta) antes de tiempo,

concretamente el 68,86% de los sujetos. En el gráfico siguiente se puede observar claramente la proporción:

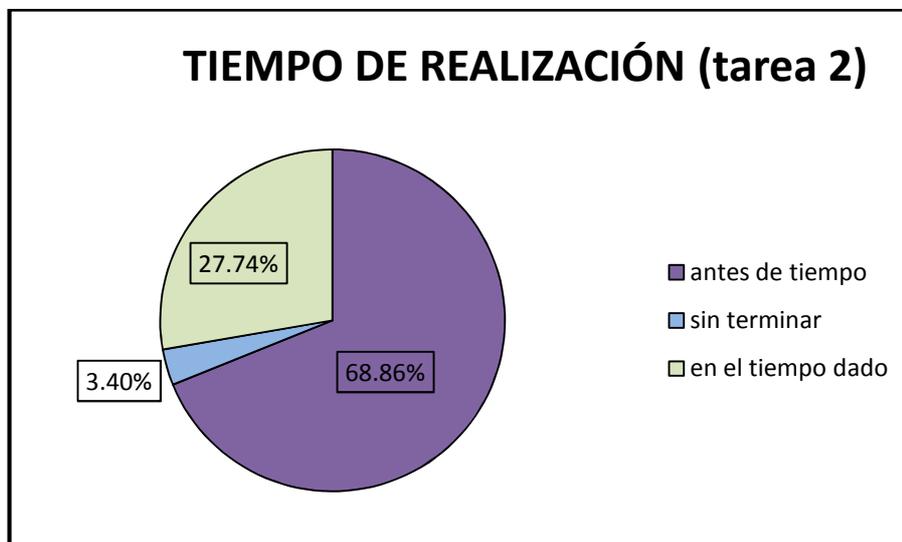


Gráfico 13. Tiempo de realización de traducciones de la tarea 2.

Con este resultado, confirmamos de nuevo el hecho de que la herramienta *Trandix* sirve de gran ayuda a la hora de realizar búsquedas terminológicas, tanto dentro como fuera de la herramienta, lo que agiliza el trabajo documental y reduce el tiempo de traducción. Todo lo cual se traduce directamente en un aumento del número de palabras traducidas o, al menos, en una reducción del tiempo requerido para finalizar un encargo de traducción.

4.3.1.4.3.4.3. Criterio de calidad

Los dos criterios utilizados hasta el momento solo han tenido en cuenta el tiempo de realización de las tareas, que ha resultado ser muy positivo para la herramienta *Trandix*, y han dejado de lado la calidad de las traducciones. La calidad que presentan estas traducciones es, a la vez que el tiempo, un criterio muy relevante en esta investigación, puesto que nos indica si realmente la herramienta permite mejorar la calidad de las traducciones a la par que el tiempo dedicado a su realización.

Para aplicar este criterio utilizaremos la plantilla mixta que hemos presentado anteriormente, la cual consta de una parte analítica, basada en aspectos concretos de la traducción, y de otra holística, basada en la evaluación global. De esta manera, evaluaremos cada una de las traducciones realizadas por los traductores semiprofesionales a fin de comprobar si los textos traducidos con la herramienta *Trandix* como recurso poseen más o menos errores y, por tanto, mayor o menor calidad que los traducidos sin la herramienta.

Para llevar a cabo esta evaluación, hemos utilizado el programa *Markin*,⁶⁷¹ desarrollado por M. Holmes en 1996 y cuya función principal consiste en facilitar al docente la corrección y evaluación de trabajos entregados en formato electrónico. El programa ofrece unos juegos de botones de corrección por defecto, que pueden ser modificados o creados por el usuario. En nuestro caso, nos hemos visto obligados a crear un juego de botones para efectuar la evaluación de acuerdo con nuestras plantillas, puesto que los juegos incluidos en el programa no se ajustaban a nuestras necesidades de evaluación analítica y holística.



Figura 111. Juego de botones creado para la evaluación mixta.

En este juego de botones, se encuentran recogidas las abreviaturas creadas *ad hoc* para realizar nuestra evaluación, que se clasifican según su puntuación positiva o

⁶⁷¹ Este programa permite la corrección y evaluación de traducciones mediante cuatro posibilidades diferentes: 1. Anotaciones, donde se especifican errores o aciertos concretos; 2. Comentarios, con los que se ofrecen explicaciones sobre algún aspecto o anotación; 3. Retroalimentación, donde el docente ofrece su comentario final al trabajo corregido, y 4. Nota, que permite calificar de diferentes formas (letras, porcentajes, calificaciones del 1 al 10, etc.). URL: <<http://www.cict.co.uk/markin/index.php>>. Para ver otro ejemplo reciente de aplicación de *Markin*, remitimos a Toledo Báez (2009), que lo emplea para evaluar traducciones realizadas con y sin la ayuda de resúmenes automáticos.

negativa.⁶⁷² De la parte analítica, encontramos como criterios negativos los siguientes: GRAM (error de gramática leve) y GRAM+ (error gramática grave), COH (error de coherencia o cohesión), TERM (error de terminología leve), TERM+ (error de terminología grave), REG (inadecuación del registro leve), REG+ (inadecuación del registro grave), ORTO (error ortotipográfico), ORTO+ (error ortotipográfico grave), PUNT (error de puntuación), EST (inadecuación del estilo), OMS (omisión), SENT (error de sentido) y FUNC (inadecuación de la función textual); y de la parte positiva, BIEN (Acierto) y EXC (muy acertado). Por su parte, la evaluación holística cuenta con los siguientes botones: para los negativos, N. 1 (nivel 1) y N. 2 (nivel 2); y para los positivos, los botones N. 3 (nivel 3) y N. 4 (nivel 4).

A continuación, presentaremos dos ejemplos de traducciones completadas de las tareas del experimento (con y sin herramienta) que hemos evaluado mediante la aplicación de la plantilla mixta de evaluación de la calidad, con objeto de ilustrar el procedimiento seguido con todas las traducciones. Según el resultado de la aplicación de la plantilla mixta, otorgaremos una nota de Apto y No apto, para distinguir entre los que ofrecen mayor o menor calidad. Posteriormente, mostraremos un recuento de las notas de Apto y No apto de todas las traducciones obtenidas en el experimento.

En primer lugar, presentaremos una traducción evaluada realizada para la tarea 1. En ellas, se pueden observar los errores detectados mediante la indicación de las abreviaturas de cada error según hemos indicado en el juego de botones anteriormente.

- Traducción 1. Traducción realizada para la tarea 1 y, por tanto, sin la utilización de la herramienta *Trandix*.

⁶⁷² En la parte superior se encuentran los valores negativos, mientras que en la parte inferior se observan los positivos.

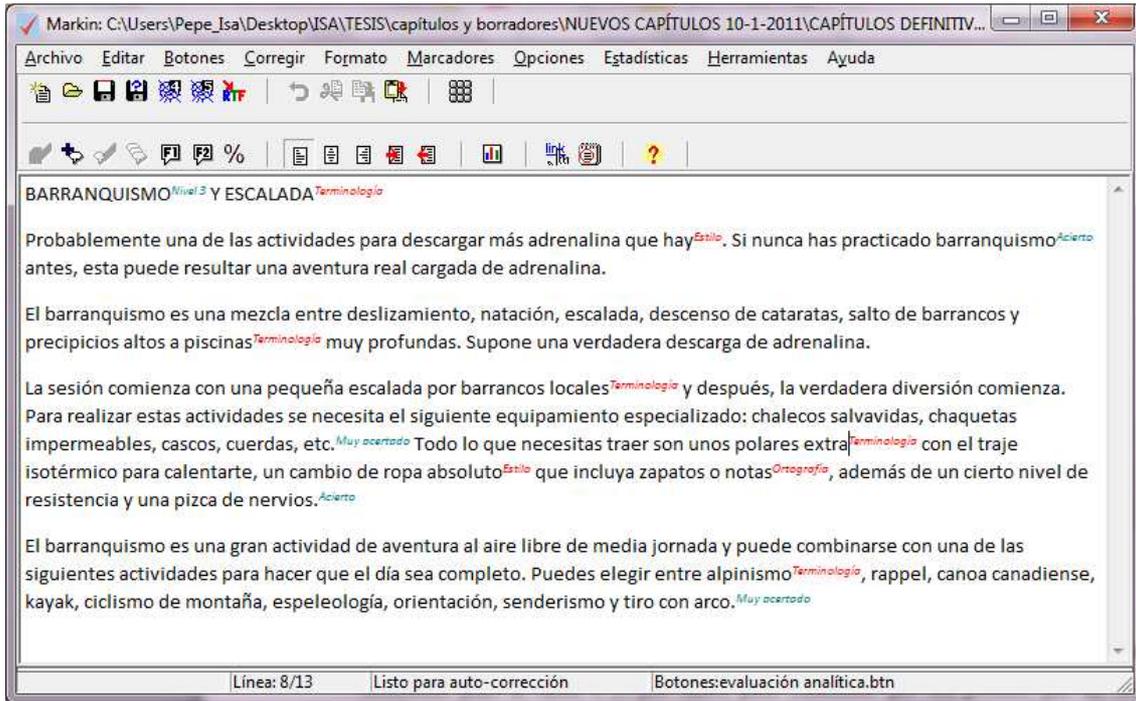


Figura 112. Evaluación y corrección con *Markin* de una traducción sin la herramienta *Trandix*.

Como se puede observar, se encuentran escasos errores en esta traducción, a la que se le ha evaluado con el Nivel 3 de la parte holística, es decir, una traducción muy completa y casi correcta. La mayoría de los errores encontrados se refieren a la terminología, el aspecto más complicado de la traducción, aunque también se encuentran dos errores de estilo y un error de ortografía, debido a una errata del texto. No obstante, se han señalado algunos aciertos y traducciones muy acertadas al traducir partes del texto, lo que compensa estos errores marcados y permite evaluar la traducción de forma positiva con la nota de Apto.

A continuación, mostramos la tabla completa de los errores cometidos por el traductor semiprofesional tomado como ejemplo.

| Número de casos | Anotación Larga | Categoría | Valor | Total |
|-----------------|--|-----------|-------|-------|
| 5 | Terminología | | 1 | 5 |
| 2 | Estilo | | 1 | 2 |
| 1 | Ortografía | | 1 | 1 |
| 2 | Acierto | | 1 | 2 |
| 2 | Muy acertado | | 1 | 2 |
| 1 | Transmisión casi completa. Gran parte se lee | | 1 | 1 |

Totales
 Negativo: 8
 Positivo: 5

Volver a contar
 Cerrar

Tabla 59. Conjunto de errores encontrados en la evaluación mixta para la tarea 1.

Esta traducción ha sido elegida por ser una de las mejores traducciones obtenidas para la tarea 1 y, por tanto, para confirmar que era posible, por un lado, completar la tarea encomendada en el tiempo dado para tal efecto, y, por otro, obtener una buena calidad de la traducción. No obstante, pocos alumnos la realizaron de forma satisfactoria como hemos podido observar con anterioridad.

- Traducción 2. Traducción realizada para la tarea 2 y, por tanto, con la utilización de la herramienta *Trandix*.

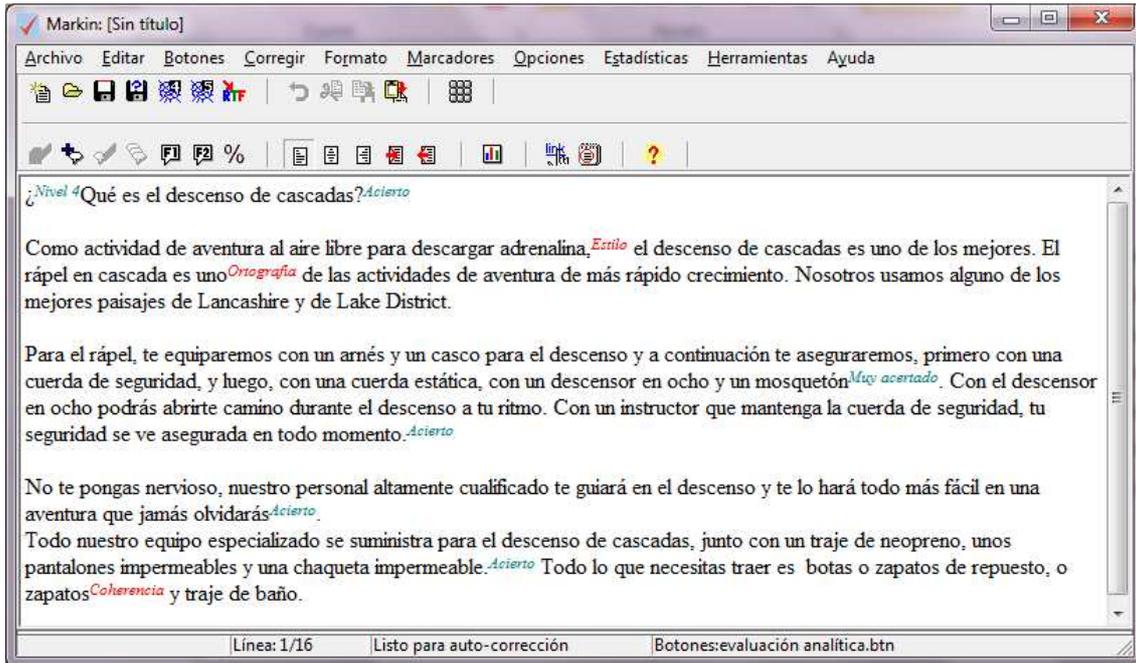


Figura 113. Evaluación y corrección con *Markin* de una traducción con la herramienta *Trandix*.

Esta traducción, al contrario que la anterior, no se ha seleccionado por ser una de las mejores traducciones conseguidas en esta tarea, sino por tratarse de la traducción realizada por el mismo sujeto anterior. Al haber seleccionado las traducciones del mismo sujeto, podemos comparar claramente las mejoras o no aportadas por la herramienta *Trandix*.

En esta traducción, se encuentran escasos errores, por lo que se le ha evaluado con el Nivel 4 de la parte holística, es decir, una traducción muy completa, correcta y prácticamente exenta de errores. A la inversa de lo que ocurría en el caso anterior, donde la mayoría de los errores encontrados se referían a la terminología, aquí no encontramos ningún error terminológico sino tres errores de diferente naturaleza: un error de estilo, un error de ortografía (debido posiblemente a una errata) y un error de coherencia, al repetir la misma información dos veces. Asimismo, hemos señalado los aciertos y traducciones muy acertadas del texto, lo que ha aumentado su calidad. Por tanto, esta traducción también ha obtenido la puntuación de Apto.

| Número de casos | Anotación Larga | Categoría | Valor | Total |
|-----------------|---|-----------|-------|-------|
| 1 | Coherencia | | 1 | 1 |
| 1 | Estilo | | 1 | 1 |
| 1 | Ortografía | | 1 | 1 |
| 4 | Acierto | | 1 | 4 |
| 1 | Muy acertado | | 1 | 1 |
| 1 | Nivel 4. Transmisión completa. Todo o casi to | | 3 | 3 |

Totales
 Negativo: 3
 Positivo: 8

Volver a contar
 Cerrar

Tabla 60. Conjunto de errores encontrados en la evaluación mixta para la tarea 2.

Después de haber comparado las traducciones realizadas por el mismo sujeto con y sin la ayuda de *Trandix*, podemos concluir que la herramienta permite una mejora sustancial en la calidad de la traducción, ya que reduce los errores terminológicos y de fraseología. No obstante, no puede evitar que los traductores cometan errores de estilo, coherencia, etc., fruto principalmente de la falta de revisión de las traducciones. De esta manera, la mejora en la calidad de las traducciones, sumado a la reducción del tiempo de realización, nos indica los beneficios que aporta *Trandix* a la labor traductora. No obstante, antes de avanzar conclusiones, presentaremos a continuación los resultados obtenidos después de evaluar todas las traducciones recopiladas con el experimento.

Como indicamos anteriormente, el total de traducciones obtenidas en el seminario práctico realizado para llevar a cabo nuestro experimento ascendió a 61 traducciones, aunque solamente las traducciones completadas pasaron el filtro del primer criterio y, por tanto, han sido evaluadas en esta fase. En total, hemos contado con 8 traducciones evaluadas para la primera tarea y 59 para la segunda tarea, de las que se han obtenido las siguientes calificaciones: para la primera tarea, 5 Aptos y 3 No aptos, y para la segunda tarea, 57 Aptos y 2 No aptos, de acuerdo con los errores y aciertos marcados en el programa *Markin* y la evaluación holística realizada.

Las traducciones consideradas como no aptas en la primera parte del experimento (sin la herramienta) han presentado principalmente errores terminológicos, además de algunos errores de sentido, gramaticales, ortográficos y de puntuación encontrados. Por su parte, las traducciones realizadas en la segunda parte del experimento (con la herramienta) han presentado algunos errores gramaticales, ortográficos y de puntuación, pero han eliminado casi por completo los errores de sentido y terminológicos, lo que ha provocado un aumento de calidad de la traducción. Además, si tenemos en cuenta que mucho de los errores gramaticales, ortográficos y de puntuación podrían haberse eliminado con una revisión de la traducción, podemos prever un aumento mayor en la calidad. Por todo esto, cabe destacar el aumento de calidad que se consigue con *Trandix*, además del tiempo que se reduce durante la traducción, como vimos anteriormente.

A modo de conclusión, y uniendo los criterios de calidad y de tiempo presentados hasta el momento, podemos ilustrar el resultado obtenido después de esta evaluación cuantitativa con el siguiente gráfico, donde se observa el total de Aptos y No aptos obtenidos: 5 Aptos y 56 para la primera tarea (sin herramienta) y 57 Aptos y 4 No aptos para la segunda tarea (con herramienta).

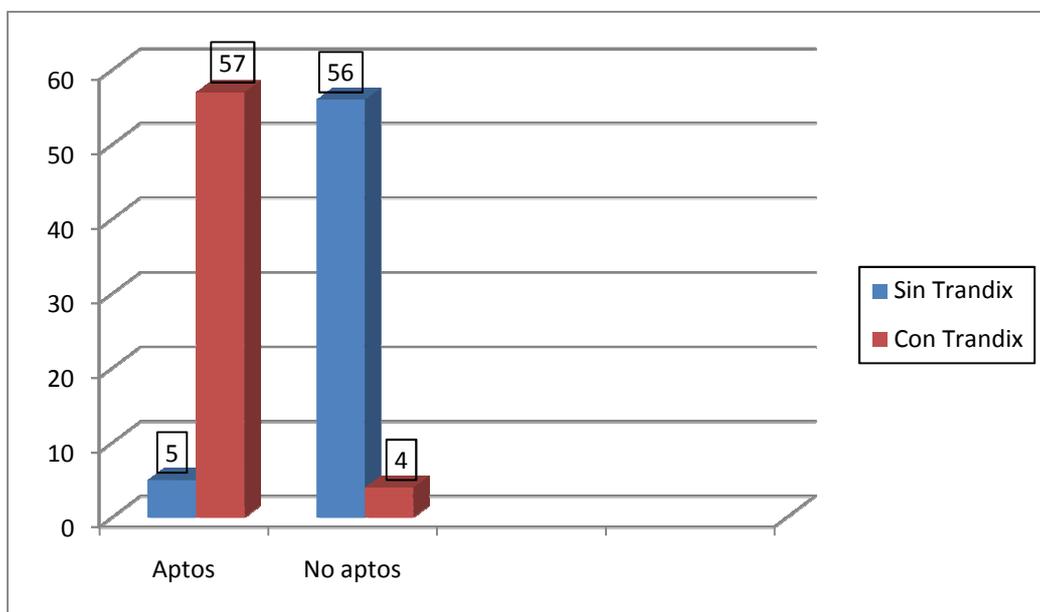


Gráfico 14. Total de traducciones Aptas y No aptas para las tareas 1 y 2.

4.3.1.4.3.4.4. Criterio de satisfacción

Finalmente, llegamos al criterio de satisfacción de la herramienta *Trandix*, un criterio que combina una evaluación cuantitativa y cualitativa y que se mide a través de las respuestas obtenidas en los cuestionarios realizados durante el experimento. Como indicamos anteriormente, se llevaron a cabo dos cuestionarios (cf. Anexo VI): uno después de la realización de la primera fase del experimento, que estaba dirigido a conocer el protocolo de traducción que seguían los sujetos y su opinión sobre la traducción realizada, y otro, después de la segunda fase, cuyos principales objetivos eran conocer su opinión sobre la herramienta *Trandix* y su nivel de satisfacción con la traducción realizada y con la ayuda prestada por la herramienta. Llegados a este punto, cabe destacar que los cuestionarios fueron totalmente anónimos, aunque fueron codificados con el mismo código indicado a los alumnos para el envío de las tareas, ya que nos resultó interesante tener la oportunidad de consultar la traducción realizada por un sujeto y, a la vez, su opinión ofrecida.

El resultado de los cuestionarios se muestra directamente en el mismo Campus Virtual en forma de porcentajes y gráficos, como se puede observar en la siguiente figura:

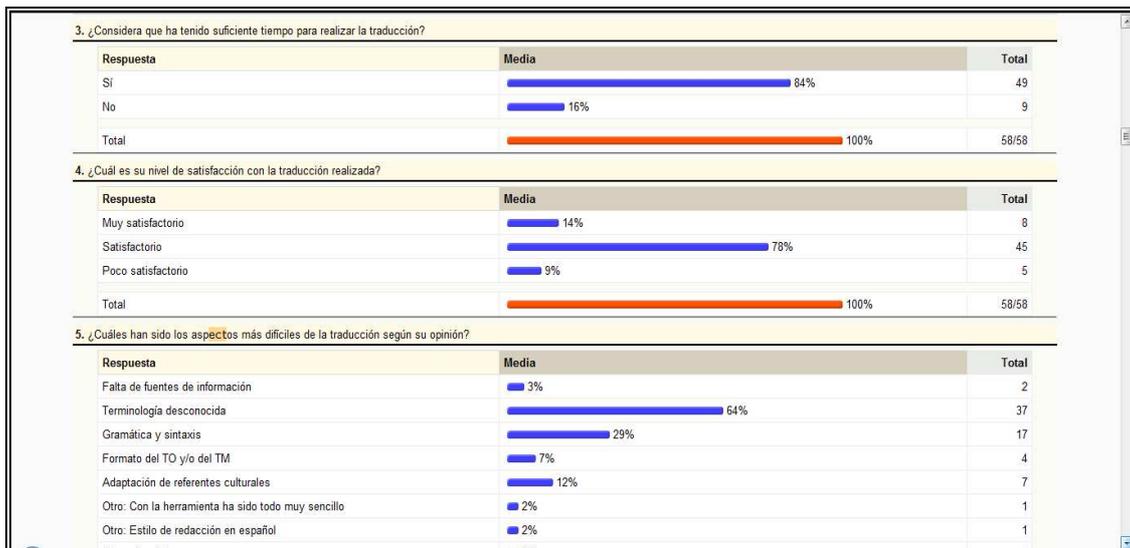


Figura 114. Ejemplo de resultados del cuestionario 2 en forma de gráfico.

En primer lugar, comentaremos las preguntas que se encuentran repetidas en los dos cuestionarios diseñados para el experimento, las cuales hacen referencia especialmente a la opinión de los sujetos sobre el texto traducido, el tiempo disponible, los aspectos más difíciles, etc. Estas preguntas nos permitirán comparar la metodología seguida para realizar la primera tarea y la metodología con *Trandix* para la segunda tarea y, así, poder sacar conclusiones de las mejoras o inconvenientes que muestra nuestra herramienta.

A continuación, presentamos una tabla que contienen estas preguntas y los resultados obtenidos en cada una de ellas tanto en el cuestionario de la primera tarea (sin la herramienta) como en el de la segunda tarea (con la herramienta).

| Preguntas repetidas en ambos cuestionarios | CUESTIONARIO 1 (sin <i>Trandix</i>) | CUESTIONARIO 2 (con <i>Trandix</i>) |
|--|--|--|
| 6. ¿Le ha resultado fácil la traducción que ha realizado? Indique del 1 al 5 el nivel de dificultad que ha encontrado (siendo 1 muy fácil y 5 muy difícil). | 2 → 21,9% 3 → 60,9% 4 → 9,4% 5 → 7,8% | 1 → 32,76% 2 → 44,83% 3 → 17,24% 4 → 5,17% |
| 7. ¿Cuáles han sido los aspectos más difíciles de la traducción según su opinión? <input type="checkbox"/> Falta de fuentes de información <input type="checkbox"/> Terminología desconocida <input type="checkbox"/> Gramática y sintaxis <input type="checkbox"/> Formato del TO o del TM <input type="checkbox"/> Adaptación de referentes culturales <input type="checkbox"/> Otro | Falta de fuentes de información: 33% Terminología desconocida: 98% Gramática y sintaxis: 11% Formato del TO o del TM: 5% Adaptación de referentes culturales: 10 Otro: 5% | Falta de fuentes de información: 5% (3 personas) Terminología desconocida: 68% Gramática y sintaxis: 29% Formato del TO o del TM: 11% Adaptación de referentes culturales: 12% Otro: 6% |
| 9. ¿Considera que ha tenido suficiente tiempo para realizar la traducción? <input type="checkbox"/> Sí <input type="checkbox"/> No | NO: 100% | SÍ: 84% NO: 16% |
| 10. ¿Ha encontrado fácilmente la información que buscaba? Indique del 1 al 5 el nivel de dificultad que ha encontrado (siendo 1 muy fácil y 5 muy difícil). | 1 → 6,2% 2 → 18,5% 3 → 46,1% 4 → 16,9% 5 → 12,3% | 1 → 53,1% 2 → 31,0% 3 → 3,3% 4 → 8,3% 5 → 4,3% |
| 11. ¿Cuál es su nivel de satisfacción con la traducción realizada? <input type="checkbox"/> Muy satisfactorio <input type="checkbox"/> Satisfactorio <input type="checkbox"/> Poco satisfactorio <input type="checkbox"/> Nada satisfactorio | Satisfactorio: 13% Poco satisfactorio: 63% Nada satisfactorio: 19% | Muy satisfactorio: 14% Satisfactorio: 78% Poco satisfactorio: 9% |

Tabla 61. Resultados obtenidos de las preguntas repetidas incluidas en los dos cuestionarios.

Una vez mostrados estos resultados, presentaremos las conclusiones extraídas de estos datos, centrándonos sobre todo en los aspectos más relevantes:

- El nivel de dificultad en ambas traducciones (pregunta 6) no se ha encontrado excesivamente alto, puesto que los porcentajes obtenidos para los niveles 4 y 5 (los más altos) no son elevados. Sin embargo, cabe señalar que los datos apuntan a que la primera tarea presenta un nivel de dificultad mayor que la segunda, puesto que, por un lado, la suma de los niveles del 3 al 5 (el 5 es el más difícil) presentan el mismo total que la suma de los niveles 1 y 2 de la segunda tarea, y, por otro, el nivel 1 no recibe ninguna puntuación en la primera tarea. En este contexto, consideramos oportuno recordar aquí que los textos seleccionados para llevar a cabo estas tareas presentan las mismas características con respecto al género y la temática, número de palabras, número de términos, etc., lo que indica la valoración subjetiva de los sujetos provocada posiblemente por la mayor dificultad que han encontrado a la hora de traducir el primer texto (sin la herramienta).
- Según los sujetos, el aspecto más complicado de las dos traducciones es, sin lugar a dudas, la terminología de los textos (pregunta 7). En esta pregunta de multirrespuesta queda patente dicha dificultad con los resultados obtenidos: un 98% para el primer texto y un 68% para el segundo texto, seguidos de lejos por los demás aspectos indicados. De nuevo aquí, debemos recordar la similitud de los dos textos con respecto a la terminología, ya que, a pesar de contar con el mismo número de términos y dificultades muy parecidas, el aspecto terminológico en la primera traducción presenta un porcentaje mayor que en la segunda. También cabe destacar el aumento de los porcentajes de los demás aspectos en el segundo cuestionario: el referente a la gramática y sintaxis y el relativo a la adaptación cultural, quizá debido a la pérdida de peso de los problemas terminológicos de la primera parte.
- El tiempo de realización de la traducción (pregunta 9) nos indica claramente la ventaja que proporciona *Trandix* con respecto a la metodología utilizada en la primera parte. Aquí, el 100% de los sujetos indica que no ha tenido suficiente tiempo para realizar la primera traducción, mientras que, para la segunda traducción, solo el 16% indica que no le ha dado tiempo frente al 84% que confirma que sí. De nuevo aquí, podemos confirmar la ventaja de

Trandix con respecto a la reducción del tiempo de traducción, su principal fortaleza.

- Con respecto a la búsqueda de información (pregunta 10), se invierten de nuevo las cifras obtenidas: para la primera traducción, el porcentaje mayor se obtiene en el nivel intermedio (3), mientras que para la segunda traducción el porcentaje mayor se recoge en el nivel que indica mayor facilidad (1). Confirmamos aquí también las ventajas que ofrece con respecto al acceso de información y, por tanto, de documentación.
- Por último, cabe destacar la valoración que realizan los propios sujetos de su traducción (pregunta 11). De nuevo, vemos aquí que los datos más positivos los encontramos en la segunda traducción (con *Trandix*), ya que aquí el 78% de los sujetos están satisfechos con su traducción frente al 13% de la primera traducción; y, más importante aún, el 14% se encuentra muy satisfecho con su resultado, mientras que en la primera traducción ningún sujeto había seleccionado esta opción. Por último, también es relevante destacar que en la primera traducción el 19% de los sujetos estaba poco satisfecho, mientras que en la segunda ninguno optó por esta respuesta. En otras palabras, la calidad de las traducciones de la segunda parte del experimento (con *Trandix*) fue superior a la calidad de las traducciones de la primera parte, al menos desde el punto de vista de los propios sujetos.

De forma general, podemos resumir lo anterior destacando el aspecto positivo que se observa en todos los resultados obtenidos con el segundo cuestionario frente a los datos del primero, tanto en lo que respecta al tiempo de realización de la traducción como a la hora de buscar información y de valorar la propia traducción. En definitiva, observamos la satisfacción de los sujetos con la herramienta *Trandix*, tanto en lo que respecta al tiempo como a la calidad.

Una vez presentados los datos de las preguntas repetidas en ambos cuestionarios, pasaremos a comentar las preguntas del segundo cuestionario relativas a la herramienta *Trandix* propiamente.

En primer lugar, comentaremos las preguntas de carácter cuantitativo, con objeto de ilustrar porcentajes exactos y datos totalmente objetivos. En este sentido,

presentamos nuevamente una tabla con las preguntas y respuestas obtenidas en este segundo cuestionario.

| Preguntas cuantitativas | CUESTIONARIO 2 (con <i>Trandix</i>) |
|--|---|
| 7. ¿Le ha resultado fácil el uso de la herramienta <i>Trandix</i> ? Indique del 1 al 5 el nivel de dificultad que ha encontrado (siendo 1 muy fácil y 5 muy difícil). | 1 → 67,3% 2 → 31% 3 → 1,7% |
| 8. ¿Le ha ayudado la herramienta <i>Trandix</i> a realizar la traducción? <input type="checkbox"/> Totalmente <input type="checkbox"/> Mucho <input type="checkbox"/> Bastante <input type="checkbox"/> Poco <input type="checkbox"/> Nada | Totalmente → 50% Mucho → 38% Bastante → 12% |
| 14. Si tuviera que elegir entre la metodología que utiliza normalmente para realizar sus traducciones y la que le ofrece la herramienta <i>Trandix</i> , ¿cuál preferiría? <input type="checkbox"/> <i>Trandix</i> <input type="checkbox"/> La metodología anterior | <i>Trandix</i> : 97% La metodología anterior: 3% |
| 16. Si la herramienta <i>Trandix</i> estuviera disponible, ¿la utilizaría para sus traducciones? <input type="checkbox"/> Sí. <input type="checkbox"/> No | Sí: 100% |

Tabla 62. Resultados obtenidos de las preguntas cuantitativas del segundo cuestionario.

Además de estas cuatro preguntas incluidas en la tabla, encontramos otra de carácter cuantitativo (pregunta de puntuación del 1 al 5, donde 1 es muy malo) referente a diferentes aspectos de *Trandix*, a saber: facilidad de manejo, facilidad de acceso a la información, flexibilidad, contenido e interfaz. Como se puede ver en la siguiente figura, la media de todos estos aspectos se encuentra en torno al nivel 4, es decir, un nivel bueno, siendo la facilidad de manejo el aspecto que mayor puntuación recibe.

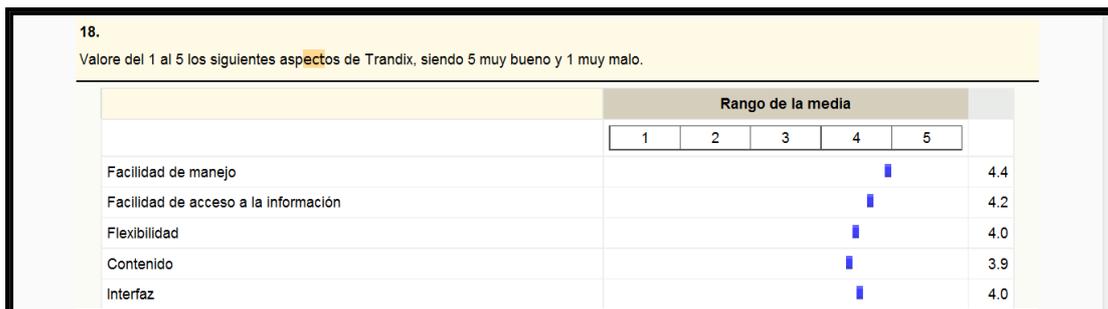


Figura 115. Resultados de la pregunta 18 del cuestionario 2.

Con estos resultados, se observa el alto nivel de satisfacción por parte de los sujetos encuestados, puesto que el 100% de ellos indica que utilizaría la herramienta si estuviera disponible y el 97% la prefiere frente a la metodología utilizada para llevar a cabo la primera tarea (sin la herramienta). Asimismo, cabe destacar que todos los sujetos consideraron fácil su utilización a la par que útil la ayuda prestada por *Trandix* e incluso el 50% considero que le había ayudado *Totalmente durante la traducción*. Por último, también resulta relevante la buena puntuación recibida en la pregunta 18 al valorar los diferentes aspectos de *Trandix* propuestos en el cuestionario.

En general, con estos datos podemos confirmar de nuevo las ventajas que ofrece *Trandix* como fuente terminológica y de ayuda a la documentación en comparación con la metodología utilizada sin la herramienta, en esta ocasión reforzando los beneficios que ofrece desde un punto de vista del uso.

Por último, pasaremos a tratar las preguntas del segundo cuestionario de naturaleza cualitativa, es decir, aquellas en las que los sujetos tienen la oportunidad de aportar su opinión y, por tanto, sus críticas o recomendaciones.

En primer lugar, encontramos las preguntas relativas a los aspectos positivos y negativos que consideran los sujetos. En general, las aportaciones de los sujetos se pueden resumir en los siguientes puntos:

- Ahorro de tiempo al realizar búsquedas terminológicas.
- La terminología ofrecida es fiable y, por tanto, se ahorra tiempo en la comprobación.
- Es posible editar el texto en el mismo programa, lo que ahorra tiempo.

- El interfaz es sencillo y el programa es fácil de manejar.
- Permite traducir de forma más rápida.
- Permite realizar búsquedas externas directamente desde el programa, lo que ahorra tiempo al traductor.
- Permite seleccionar los campos de la entrada terminológica.
- Al ahorrar tiempo, permite que el traductor pueda concentrarse en otras cuestiones tan importantes o incluso más que la terminología.
- Elimina la infoxicación.⁶⁷³

Por su parte, encontramos también una serie de comentarios acerca de lo negativo de la herramienta, que se presentan a continuación:

- Sólo se encuentra disponible para el ámbito del turismo de aventura y para tres lenguas (español, inglés y alemán).
- Debería estar vinculado a un mayor número de buscadores que pudieran abrirse en la misma interfaz del programa.
- En ocasiones, reconoce algunos términos con categoría gramatical errónea. Por ejemplo, «kit» en el texto es utilizado como verbo pero el programa lo reconoce como sustantivo.
- La falta de opciones propias de un editor de texto, a saber: corrector ortográfico, cambio de colores, etc.
- No tiene opción de *Deshacer*.
- No permite exportar un archivo directamente, sino que es necesario copiar y pegar el texto en el procesador habitual.
- Se deberían mejorar las búsquedas externas en los diccionarios y añadir más fuentes.
- Estaría bien que incluyera imágenes en las entradas, además de toda la información terminológica.

⁶⁷³ Recordemos que la *infoxicación* es un neologismo que hace referencia al exceso de información que encontramos actualmente en la Red y que nos obliga a conocer las técnicas para evaluar y seleccionar la información útil y fiable.

Relacionada con esta pregunta, encontramos la pregunta 15, que hace referencia a los cambios que proponen los sujetos y que se encuentran muy relacionados con los aspectos negativos presentados anteriormente. Los comentarios de los sujetos se pueden resumir en tres asuntos básicamente, a saber:

- Ampliación de campos de especialidad e idiomas.
- Ampliación de opciones para incluir aquellas propias de los procesadores de textos como Word (color, opción *Deshacer*, corrector de errores, etc.)
- Opción para añadir más enlaces a otros recursos.

Para finalizar el cuestionario, se les ofreció la oportunidad de dar su opinión o aportar cualquier comentario que desearan. La gran mayoría de los usuarios utilizó esta pregunta para mostrar su satisfacción por haber probado esta herramienta y por los resultados obtenidos, además de aprovechar para ofrecer sus ánimos a los creadores y sus deseos para que la herramienta esté pronto disponible. A modo de ejemplo, mostramos alguno de estos comentarios:

| 19. ¿Desea añadir algún comentario sobre la herramienta <i>Trandix</i> ? | |
|--|---|
| nº | Respuesta |
| 1 | ¡Enhorabuena y suerte! |
| 2 | . |
| 1 | A demás de los recursos utilizados, he accedido a un diccionario de sinónimos. |
| 1 | Apuesto por su desarrollo para un uso futuro. |
| 1 | Creo que ha sido una gran idea el desarrollo de una herramienta como esta, ya que nos facilitará bastante la tarea de traducción. |
| 1 | Debería estar preparada para cualquier tipo de temática |
| 1 | Es una herramienta bastante útil, con el único inconveniente de que no se encuentre disponible y de que no disponga de terminología para todos los campos. |
| 1 | Es una herramienta eficaz. Creo que podría ofrecer otras fuentes dónde realizar las búsquedas pero, por lo general, es una herramienta bastante útil. |
| 1 | Es una herramienta excelente. |
| 1 | Es una herramienta muy útil |
| 1 | Esperemos que esté disponible para su utilización en un futuro. |
| 1 | Espero que no sea de pago y se subvencione con anuncios o algo así ;) |
| 1 | Espero que pronto esté disponible. |
| 1 | Estamos satisfechas con el resultado. |
| 1 | Ha sido muy agradable utilizar la herramienta. Espero poder contar con ella en un futuro. |
| 1 | Hemos disfrutado mucho con su uso. |
| 1 | Me gustaría que cuando esté disponible, se comunique de ello para que lo sepamos y espero que no sea a un precio muy elevado. |
| 1 | Me ha encantado. |
| 1 | Me ha gustado mucho esta herramienta y me ha resultado muy útil a la hora de realizar la traducción. Considero que es un recurso fiable, cómodo y rápido de utilizar. |
| 1 | Me ha parecido una herramienta muy útil para futuras traducciones. |

Figura 116. Ejemplos de comentarios obtenidos en la pregunta 19 del cuestionario 2.

Una vez que hemos expuesto los resultados obtenidos en los dos cuestionarios referentes al criterio de satisfacción de los sujetos del experimento, así como los demás criterios comentados (criterio de calidad, criterio de tiempo de realización y finalización de las tareas), pasaremos a indicar las conclusiones extraídas.

4.3.1.4.3.4. Conclusiones del experimento y trabajo futuro con *Trandix*

La herramienta *Trandix* ofrece una serie de beneficios al proceso de traducción que han quedado patentes durante todo el apartado anterior. Principalmente, su aporte más relevante consiste en la reducción del tiempo en el que se realiza la traducción gracias a disminuir el tiempo de documentación y de búsqueda terminológica. Todo esto, a su vez, repercute positivamente en la calidad de la traducción como se ha observado, puesto que si el traductor necesita menos tiempo para llevar a cabo la traducción contará con más tiempo para realizar la revisión del texto traducido y, por tanto, para alcanzar una mejor calidad. Asimismo, podemos destacar aspectos de *Trandix* más concretos, que se han recogido también como puntos positivos en las respuestas de los sujetos del experimento, como son las búsquedas en recursos externos a la herramienta, la posibilidad de seleccionar la información que aparece en la entrada terminológica de los términos, así como la opción de editar el texto directamente en el programa.

Todo esto nos permite concluir que la herramienta *Trandix* ofrece una forma de búsqueda terminológica y de apoyo documental muy satisfactorio para los traductores, puesto que reduce el tiempo de trabajo, permite dedicar más tiempo a la revisión, agiliza el trabajo de búsqueda y tiene flexibilidad para seleccionar la información deseada.

Con vistas a una posible ampliación en el futuro de las funcionalidades de *Trandix* y, con objeto de algunos subsanar los aspectos negativos indicados por los sujetos del experimento, nos hemos planteado una serie de líneas de actuación que comentamos a continuación.

En primer lugar, y en consonancia con muchos de los comentarios recibidos por los sujetos, consideramos que nuestra aplicación informática no debería limitarse a un dominio de especialidad (en este caso, el turismo de aventura), por lo que nos parece

indispensable la constante actualización e inclusión de fuentes de información, concretamente recursos terminológicos, de diferentes campos de especialidad. De esta manera, la lista de recursos terminológicos iría creciendo de forma constante para abrir los campos de actuación de la herramienta y, así, se agilizaría el proceso de consulta durante la traducción especializada de un número mayor de dominios.

En segundo lugar, y relacionado con lo anterior, nos parece imprescindible también ampliar el número de lenguas de trabajo de la herramienta. En principio, hemos optado por incluir únicamente español, inglés y alemán, por ser las lenguas de trabajo de esta investigación, pero nuestro objetivo es ampliar este número a otras lenguas, como pueden ser el francés, el italiano o el portugués. De la misma manera que la anterior, las fuentes de información se ampliarían también a dichas lenguas.

En tercer lugar, nos parece oportuno, después de haber encontrado tantas dificultades derivadas del lematizador seleccionado (a pesar de ser uno de los más utilizados en el campo del PLN), seguir buscando otros lematizadores que nos permitan solucionar los problemas encontrados principalmente con la lengua española y, en caso de no encontrar ninguno que satisfaga nuestras necesidades, desarrollar uno propio que se ajuste a las necesidades de nuestra herramienta.

Y, por último, consideramos que sería relevante estudiar la posibilidad de implementar este sistema en las memorias de traducción más utilizadas, como *SLD Trados*, *Déjà Vu X*, *Across*, entre otras. Hoy en día, la gran mayoría de estos sistemas de memorias de traducción han integrado módulos que permiten la gestión terminológica por parte del traductor, algunas de las cuales incluso están disponibles como herramientas independientes, como es el caso de *SDL MultiTerm 7*, *TermStar XV* o *Uniterm*. El funcionamiento de estos programas integrados en un sistema de memoria de traducción consiste en la interacción directa con el editor del sistema para que todos los términos que aparecen en el segmento del texto de origen en el que el traductor está trabajando se busquen automáticamente en la base de datos. De esta forma automática, todos los términos que se encuentren en la base de datos terminológica, introducidos previamente por el traductor, se podrán visualizar en una lista que aparecerá en una ventana separada junto con su equivalente en la lengua de llegada, y se podrán seleccionar para traducir el término en el texto origen. El sistema es muy similar al que

presentábamos antes como los antecedentes de *Trandix* y de nuestra misma aplicación, con la diferencia de que estas herramientas son alimentadas normalmente por los propios traductores que las utilizan y a partir de la terminología con la que trabajan.

En este contexto, debemos recordar que los traductores trabajan siempre contrarreloj y que, por tanto, el trabajo terminológico que realizan suele realizarse durante el proceso de traducción y pocos dedican más tiempo después de terminar la traducción para completar las fichas terminológicas, para crear nuevas o modificar las antiguas. De esta manera, nos parece que sería interesante estudiar la posibilidad de integrar un sistema de búsqueda textual como el que proponemos en este trabajo de investigación, que permita agilizar la búsqueda de información terminológica, elaborada de forma sistemática y según una metodología firme y coherente, se facilitaría el trabajo de los traductores y se evitaría que estos dedicaran excesivo tiempo, por un lado, a la elaboración de fichas terminológicas dentro de sus herramientas de traducción asistida por ordenador (TAO), como con *Multiterm* o *TermStar*, y, por otro, a la consulta de multitud de recursos terminológicos externos y descentralizados. Es decir, con la incorporación de una aplicación informática como *Trandix* a las herramientas TAO creemos que se agilizaría el trabajo de los traductores a la hora de buscar información terminológica y se ahorraría el tiempo de estos traductores a la hora de elaborar sus propias fichas. Además, se fomentaría el intercambio terminológico a través de archivos TBX, por lo que las empresas que contratan a los traductores podrían facilitarles los archivos en dicho formato con las bases de datos terminológicas adecuadas para el proyecto en cuestión, así como los mismos traductores podrían intercambiarse dichos archivos. No obstante, recordamos que este punto es simplemente una idea inicial y que, por tanto, no se ha trabajado aún en profundidad.

A modo de cierre, podemos decir que la herramienta *Trandix* permite mejorar el trabajo del traductor, agilizando la búsqueda documental y terminológica y reduciendo el tiempo de trabajo, por todo lo cual vemos cumplido uno de los objetivos principales de esta tesis doctoral.

CONCLUSIONS

Throughout this research work we have been able to confirm our main initial hypothesis, which states: Professional translators require terminological resources created specifically to them as user group, since they present concrete needs and requirements that are different from other user groups, together with the other methodological sub-hypotheses thanks to the achievement of our main and secondary goals. These main achievements and contributions of our work are specified below so as to point out the features that we consider the most relevant and innovative. This is followed by the general conclusions drawn from our research and the future lines of work that we pretend to initiate as a continuation of this research.

1. In chapter 1 we provided a thorough review of major work in terminology field by presenting the differences and similarities between traditional and modern theories in terminology but also the linguistic disciplines which have significantly influenced the current terminological lines of research.
2. The point stated above leads us to the practical branch of terminology, Terminography, and two of its current lines of research: corpus-based Terminography and ontology-based Terminography. Also in chapter 1 we present

a revision of the main features of corpora and ontologies in a terminological context regarding the following aspects:

- 2.1. Most common taxonomies of corpora and ontologies with reference to their use in Terminography.
- 2.2. Advantages and disadvantages that corpora and ontologies bring to Terminography.
- 2.3. Current computer applications to be used in terminological work within these two lines of research.
- 2.4. General and specific criteria which delimitate the compilation of corpora or design of ontologies in a terminological context.
- 2.5. Revision of projects working on the basis of corpora and ontologies in a terminological context.

3. After having reviewed the literature and the advantages/disadvantages about corpora in Terminography, we can confirm one of our sub-hypotheses established at the beginning of the research, which states: *Multilingual corpora (comparable and parallel), virtually compiled by means of electronic resources, are valuable terminological and phraseological sources which enable us to make up for deficiencies encountered in traditional terminological resources (dictionaries, databases, etc.).*

4. We must highlight a relevant contribution included in this first chapter: the proposal of the term *Ontoterminography* so as to distinguish our methodology from other terms and methodologies used in this context such as *Ontoterminology* (Roche et al. 2009), which defends an onomasiological approach in Terminography close to the traditional assumptions, and *Termontography* (Kerremans et al. 2003), which is based on the Sociocognitive Terminology Theory. Our approach upholds the use of corpora as the starting point of any terminological resource, since it is a valuable tool for acquiring linguistic and expert knowledge about a domain, and also combines Temmerman's assumptions with the Communicative Terminology Theory's assumptions (Cabré Castellví 2003).

5. In chapter 2 we have carried out a study about translators' competences regarding the European norm UNE EN-15038 (2006). As one of the most relevant competences, we have put special emphasis on the research competence and therefore, we have carried out a revision of the different information sources which translators may use while they are translating and searching for information, especially on different types of terminological resources. After having reviewed the most relevant proposals to date, we have established a taxonomy to classify these terminological resources based on ten criteria: 1) presentation (glossary, vocabulary, dictionary, etc.), 2) number of languages (monolingual, bilingual), 3) medium (CD-ROM, online, paper), 4) content organisation (alphabetically, onomasiologically. other), 5) standardisation (prescriptive, descriptive), 6) target audience, 7) type of terminological units (paradigmatic, syntagmatic), 8) delimitation in specialised domain, 9) information source, and 10) user purposes (decode, re-encode).

6. As the number of terminological resources is continuously growing but, at the same time, their quality is often questioned, we considered it was necessary to design a protocol to evaluate this type of resources and establish a scale of reliability to quantitatively define their quality. Therefore, in chapter 2 we reviewed different evaluation protocols and templates and designed our own template for assessment since none of the examined fulfilled our expectations. Based on this template we were able to assess terminological resources (in paper and electronic) and decide whether we use it or not with reference to the grade obtained from the application of the scale of reliability.

7. Also in chapter 2, and as a continuation of the previous contribution, we carried out a survey among professional translators in order to know their needs and expectations regarding terminological resources. This survey was translated into different versions and distributed worldwide among associations, enterprises, distributional lists, etc. so as to reach as many translators as possible (cf. Appendix III). Before carrying out the survey the literature related to this matter was checked but very few researches working on the same topic as ours were found. The works reviewed addressed their studies either to dictionary users in general or second-language students or translation students, but few included

professional translators as their population sample. In this sense, we consider that this contribution of our PhD becomes very relevant, since it shows objective and quantitative data about professional translators' needs and expectations about terminological resources. From this study, we can draw some conclusions:

7.1. The results obtained in this research clarify the needs and expectations that professional translators present as real users. We now know more about their opinions regarding the current terminological resources and also about their "ideal" resource.

7.2. Translators are not real experts in the numerous and different domains they work in and thus, their translation process is mainly based on all the terminological resources they consult. Hence, they need appropriate and specific resources in order to satisfy their needs and to be able to provide high quality results in their translations.

7.3. The results obtained from this research differ from the conclusions drawn in previous studies (based on trainee translators and second language students), which defended the needs to include linguistic, pragmatic and semantic information in resources for translators. With our research we observed that professional translators consider semantic information (semantic relations, semantic frames or domains, etc.) as desirable data but not as essential data, *i.e.* they do not see this information as essential but only as complementary. Consequently, we observe how trainee translators need different information to professional translators and therefore, it is necessary to distinguish between each user group' needs in order to elaborate adequate resources for them.

7.4. Also, it is relevant to take professional translators' needs into consideration in order to provide quick and easy access to information in online resources and include good and concrete definitions together with pragmatic information (context, tips of use, information about false friends, etc.), which help them understand the source term and correctly translate it.

8. Chapter 2 ends with the achievement of the central goal of this research work, which is concerned with the design of a systematic and descriptive methodology with a linguistic-cognitive approach to elaborate adequate terminological resources for translators taking into account the criteria of multidimensionality and dynamism in specialised languages. In order to achieve this main goal, we firstly undertake a revision of the literature regarding the most relevant methodologies to carry out ontoterminographical works and, secondly, we present the practical application of our methodology proposal and the phases and sub-phases in which it is divided. As said above, this methodology is based on several aspects: 1. The user group are translators, 2. It is based on the new line of terminography called Ontoterminography, 3. It is systematic and descriptive, 4. It employs semantics frames to represent domains and their concepts, and 5. It takes multidimensionality and dynamism of specialised languages into account. Taking these aspects into account, the methodology is divided in six main phases: 1. Project specification and analysis of the domain, 2. Compilation of the specialised corpus, 3. Ontoterminographical management, 4. Elaboration of the ontoterminographical database, 5. Verification, and 6. Exportation and edition, which are tackled from a general point of view, that is, taking into account the different possibilities that terminologies could choose depending on tools, goals, etc. of their specific projects. Within this methodological proposal we also find a set of contributions, which are exposed below:

8.1. In the first phase we indicate the procedure that should be followed in order to precisely specify the project and analyse the study domain with regard to different levels: linguistic, economic, translation, among others.

8.2. Secondly, we define the steps to be followed in order to compile a corpus which satisfy necessities in a terminological context.

8.3. The next phase provides several contributions related to the different sub-phases in which it is divided: 1. It completes a revision of the different possibilities to extract candidate terms and possible tools, 2. It presents the most recent researches to extract equivalents from corpora (comparable and parallel), 3. It indicates the possibilities to extract semantic relations

and the methodologies to do it, and 4. It states the steps to elaborate domain ontologies based on semantic frames.

8.4. In the fourth phase we detail how to elaborate a terminological database based on an ontoterminographical work and provide relevant contributions to this research regarding the different sub-phases: 1. We undertake a revision of the database fields proposed by the ISO 12620 (1999) and the organisation of these fields in a database entry, and indicate the fields that should be included in a resource for translators (with regard to the results obtained in our survey), 2. We propose the methodology to write definitions based on the application of semantic frames at a microstructural level, and 3. We indicate the steps to select proper contextual examples to be included in the database.

8.5. Finally, the last phase offers another relevant contribution regarding the needs of innovation that is found in the elaboration of terminological resources nowadays. In this phase we carry out a review about several theories which defend the needs of flexibility and adaptability to users, and consequently we propose several options to improve the edition of these resources and new ways of searching and consulting terminological information.

9. Chapter 3 puts into practice the methodological proposal with its first phases and presents its application to a concrete specialised domain: adventure tourism. By doing this, we achieve a set of results:

9.1. We specify the pragmatic-linguistics variables (user group, domain, languages, etc.) and human and technological resources of our terminological project. In this sense, we make two types of contributions: on the one hand, we exemplify the steps to be followed when a terminological project is starting and, on the other hand, we define the variables and resources of our own terminological project.

9.2. In order to select the tools to be used in our terminological project we undertake a revision of the possible applications and make a distinction between terminology management systems and ontoterminology

management systems. Also we indicate the advantages and disadvantages of each system and justify the election of the system used in our project.

9.3. We perform an exhaustive analysis of the tourist sector from different viewpoints: economic, considering tourism as an economic sector; linguistic, considering it as the language of tourism; translation, as a specialised translation of tourist texts, and legislative, as an object to be regulated by law. In this realm, we obtain enough information about the tourist sector so as to recognise necessities, common features (especially linguistic features), limitations, etc. and therefore, to acquire specialised knowledge for our terminological project in this domain. Firstly, we analyse the concept <tourism> and its features as a whole, and afterwards we concrete our study on the adventure tourism. In this phase, we present the features of this domain and also the current controversy regarding its delimitation from other near tourist segments such as active tourism, nature tourism, ecotourism, etc., with which it is usually confused.

9.4. In this sense, our main contribution is to propose and justify a delimitation of this segment (together with the activities included in it) as well as a conceptual organisation of the tourism sector with reference to the adventure tourism and its near tourist segments (cf. 9.3.).

9.5. Related to the previous point, we also make searches for terminological resources about the adventure tourism and assess them by using the template for assessment designed in Chapter 2. The aim of this analysis consists of determining whether there are adequate resources for translation in this domain but also to specify the level of reliability that they reach. After having assessed all the resources found for this domain, we can confirm one of the secondary hypotheses established at the beginning of our study, which states: *There is a lack of high-quality, complete and adequate terminological resources for the tourist sector in general and for the adventure tourism in particular to carry out translation of specialised texts in this domain.*

9.6. The analysis of the adventure tourism domain finishes with the conceptualisation of the domain, that is, with the proposal of a representation of the domain categories. This representation is carried out once we have acquired specialised knowledge about the domain and always with the support of domain experts. The representation that we propose is based on semantic frames and therefore, it shows a frame-based form. Thanks to this we can perfectly observe the relations established between concepts that appear in the same context and clearly understand how specialised domains are organised.

9.7. Finally, we find the last contribution in Chapter 3: the compilation of the corpus AVENTURIS, a multilingual corpus (Spanish, English, and German) of promotional texts from webpages in the subdomain of adventure tourism. The multilingual corpus comprises one comparable subcorpus (AVENTURIS 1), containing texts originally written in the three working languages, and one parallel subcorpus (AVENTURIS 2), containing one component in original language (Spanish) and texts translated into the other languages (English and German). The compilation of this corpus is essential in terminography, and of course in ontoterminography, since it is the basis of the work that will be carried out afterwards. In this sense, this corpus must be compiled according to the general and specific criteria for a terminological context proposed in Chapter 1 and present representation for its specialised domain.

10. In Chapter 4, we contribute with the elaboration of the multilingual terminological database (Spanish, English, and German) for the adventure tourism sector and its exportation by carrying out the remaining phases of our methodology. Throughout the elaboration of the database, we also achieved a set of goals and provide several contributions to this research:

10.1. Thanks to the automatic extraction of candidate terms provided by the system we have selected to work with, *ProTermino*, we have obtained the most representative terminology of the adventure tourism domain and analysed its features regarding the different types of terms found

(neologisms, variants, foreign words, etc.). In other words, we have been able to study the terminology of this domain and its main features and access the most representative terminology, which is been used for the elaboration of the terminological database. This result occupies a valuable role since there are few terminological resources for this domain and most of them are of a very low quality according to the assessment performed. In this sense, our contribution is not limited to the fact that is been acquired in a semi-automatic way by means of a computer application but also to the fact that there is no other resource, to the best of our knowledge, that is concerned with the most representative terminology of the adventure tourism segment.

10.2. Another important contribution is related to the extraction of translation equivalents from corpora: The use of cognates to find equivalents is an innovative procedure that has been applied in this research and it has offered very good results. The system used to carry out our work, *ProTermino*, includes this functionality and thus, we have been able to employ it in our work. As said, the achieved results have been successful and also the advantages that it presents (ease, speed, etc.) prove it as a recommended method to extract equivalents.

10.3. The proposal of the main semantic relations between concepts found in the adventure tourism domain and the extraction of semantic patterns that represented these relations are also considered as an achievement in this phase. The semantic relations were established when the domain conceptualisation was carried out in Chapter 3, but now after a deep analysis of the adventure tourism domain we are able to concrete these relations and include new ones. Also, thanks to the search of term concordances we can confirm the relations established and also detect the semantic patterns which represent these relations.

10.4. Once we have determined the relations and detected the main patterns, we created our domain ontology at a conceptual level. This ontology, as said above, shows a frame-based form and thus, indicates the

context in which concepts are related to other concepts. Following the conceptual level, we continued to work at a terminological level, that is, we completed the ontology with the terms validated for the adventure tourism domain. This domain ontology is also another relevant contribution to this research since it is used both to organise the domain knowledge at a macrostructural level and to write definitions at a microstructural level.

10.5. Related to the previous point, we also provide a methodology to write definitions in a systematic and coherent way by using the domain ontology as the basis. That is, we create definitional templates for each conceptual category that has been recognised in the adventure tourism domain and prove the validity of these templates by writing sample definitions for each category. In this sense, we test the sub-hypothesis which states: *The use of definitional templates based on semantic frames for a specialised domain provides the possibility to write definitions in a systematic homogeneous and complete way.*

11. According to the above achieved results included in Chapter 4, we can also confirm another sub-hypothesis of our research: Ontoterminography combined with the Semantic Frame (Fillmore, 1977, 1982, 1985) is an adequate methodology to elaborate terminological databases for translators, since it enables us to contextualise concepts and terms, write homogeneous definitions, show clearly relations between concepts, facilitate access to terms from other terms and their definitions, among others.

12. Also in Chapter 4, we encounter a very relevant contribution regarding the exportation and edition of the terminological database, which confirms our last hypothesis: *A terminological resource which provides a flexible terminological management by translators and which includes different types of searches (including contextual searches) will enhance translator's tasks and will help achieve better results regarding time and quality.* Here we present the design and implementation of an innovative computer application to make searches and consultation of terminological information called *Trandix*. This tool pretends to fulfill the gap that is found in terminological resources regarding searches and

consultation. Most of the resources available offer traditional types of searches and consultation and, although their content can be excellent, they do not fulfill the requirements stated by translators (or other users). In this sense, we propose a new flexible tool that offers an innovative way of searching and consulting information which pretends to satisfy translators' needs and expectations. In order to design and implement this tool, we have also provided some other results:

12.1. We analysed translators' needs, the function theory, by Tarp (2008a), and proposals for electronic multifunctional dictionaries (Spohr, 2009) so as to design a coherent and useful tool with adequate functionalities. By doing this we can confirm one of our sub-hypotheses related to this: *The different current types of searches and consultation of terminological resources do not fulfill translators' needs as they require faster, easier and complemented ways of accessing information.*

12.2. We carried out a revision of the previous systems that can be considered as its ancestors. Also we compare the functionalities included in these systems and the ones that are found in *Trandix* in order to examine the advantages and disadvantages of our system and detect weaknesses.

12.3. To prove the efficiency and adequacy of the tool *Trandix*, and confirm our hypotheses, we organised a seminar with *semi-professional* translators (undergraduate students in their 4th year in Translation and Interpreting at the University of Málaga) who were able to test the tool and give quantitative and qualitative data by means of a survey. This seminar was divided in three parts: the first one consisted of an introduction to new technologies and tools supporting the research phase of translation (information sources in general, comprehensive assistants including *Trandix*, etc.); the second one was devoted to a translation of an original promotional text in English about an activity of the adventure tourism sector without *Trandix*, but with the possibility to use any resource on the Internet, and the last one was organised as another practical session in which the subjects had to carry out a translation with the same features as in the previous session but this time with the tool *Trandix*.

12.4. In order to assess the results obtained in the practical sessions, we established four criteria, which measure different aspects of the tool and the procedure of translating with it. These criteria were the following: execution of the tasks, time of execution, quality and users' satisfaction criteria. Within this context, we also make a review of the different assessment methods and proposed a template to assess translations based on different previous assessment templates.

12.5. Based on the results of the experiment and their assessment, we can confirm the benefits that our tool offers to translation regarding time of execution, efficiency, ease of use, translation quality, among other aspects indicated in Chapter 4. Therefore, comparing it with a methodology without *Trandix*, the improvements achieved are evident.

Once we have presented the contributions of this research, we continue to summarise the general conclusions drawn from our work:

1. Translators require terminological resources designed for them as a specific user group, since they present concrete needs and requirements, which are different from other groups. In this sense, resources designed to them must include both linguistic and pragmatic information so that translators can successfully rewrite the original message of any text.
2. Related to the one above, it is necessary to continue innovating in terminological resources for translators (like *Trandix*) so as to fulfill their concrete needs and expectations during the translation process.
3. Frame semantic-based ontoterminography can be considered as an excellent methodology to systematically elaborate resources for translators, since it facilitates concept representation in contexts, both at macrostructural level (regarding domain) and microstructural (referred to definitions) and enables access information in an efficient way.
4. The quantity of information available nowadays is so vast that assessment is essential to select reliable sources and therefore, translators must be trained in documentary search, *i.e.*, in learning how to locate, validate and correctly use the

information sources offered by the library and the new technologies during the translation activity.

5. Tourism is a very important economic sector in Spain, even during the current economic crisis, and hence it requires a special attention in all their aspects, including translation of tourist texts.
6. Tourist translation in general, and especially in Spain, is not properly estimated both at academic and professional levels. It is still considered as a non-specialised language, what provokes a lack of high-quality translations, a lack of training translators in tourist texts, a lack of resources, presence of numerous mistakes, etc.
7. Adventure tourism is a growing segment in the tourism sector and, as such, its multidimensionality, dynamism and neologisms are frequent. In this sense, we underline the need to unify opinions and create a coherent and organised domain in different spheres: tourist experts, legislators, administrations, agencies, etc. Otherwise, this segment will continue growing in a disorderly manner as it has been doing it to date.

Finally, we present the possible future lines of research emerged during this research:

1. The empirical study carried out with semi-professional translators to test the adequacy of the tool *Trandix* should be repeated with professional translators. The result would serve as an accurate indicator of the advantages of the application *Trandix* as a terminological resource in real translation markets.
2. Another line of future research related to our application is to extend the domain and working languages. To its implementation, we have focused on a very specific domain (adventure tourism) and languages (Spanish, English, and German), because they were the relevant data for our terminological project. However, as any TBX file could be uploaded and used on the application, it would be of particular interest to update *Trandix* and include other information resources for external searches so that it could be used for other specialised texts.

3. Related to the previous point, it could be also useful to repeat the same empirical study, but with other domains and languages so as to prove whether the tool offers good results.
4. Also, it could be possible to test other methods for term and equivalent candidate extractions, like the methodology proposed by Le Serrec et al. 2010, Gamallo (2007, 2008) and Gamallo and Pichel (2005) or Ha et al. (2008), and compare results with the cognate methodology used in this research.
5. Regarding the difficulties found in *Trandix* due to the lemmatizer employed (*TreeTagger*), especially with the Spanish language, we consider it is necessary to test other options or even implement a new one.
6. Another future line of research, also related to *Trandix*, is to study the possibilities of implementing this system into translation memories, like *SLD Trados*, *Déjà Vu X*, *Across*, among others. Nowadays, most of these computer-assisted translation tools include terminological management systems, some of which are also available as independent tools, like *SDL MultiTerm 7*, *TermStar XV* or *Uniterm*. These systems are helpful while translating but present several limitations: they cannot be interchanged between different systems, must be fed by translators, etc. In this sense, we consider it could be useful to integrate a textual search system like *Trandix*, which enables users to improve terminological searches and consultation, provides them with a database developed in a systematic and coherent way by TBX files and saves their time by avoiding searches in numerous resources and by eliminating their needs to create terminological records in their terminological management systems.
7. And finally, we also pretend to do some changes in the tool *ProTermino*, especially regarding the functionalities and design of the ontology editor, the access to the concordancer module and the possibility to export images in TBX, so as to eliminate current limitations in its use and improve its possibilities.

Most of these future directions will be developed in the current research project *Ecosistema: espacio único de sistemas de información ontológica y tesauros sobre el medio ambiente* (FFI2008-06080-C03-03/FILO; 2008-2011), directed by Dr. Corpas

Pastor and Dr. Faber, where the most innovative lines in terminology, ontology, and automatic summarization will emerge.

BIBLIOGRAFÍA

- Aarts, J. y Meijs, W. (ed.) 1984. *Corpus Linguistics I: Recent Developments in the Use of Computer Corpora*. Ámsterdam: Rodopi.
- Abaitua Odriozola, J. 2002. «Tratamiento de córpora bilingües», en M. A. Martí Antonín y J. Llisterra Boix (eds). *Tratamiento del lenguaje natural*. Barcelona: Universidad Autónoma de Barcelona. 61-90.
- Abercrombie, D. 1965. *Studies in phonetics and linguistics*, Londres: Oxford University Press.
- Abián, M.A. 2005. *Ontologías: qué son y para qué sirven* [en línea]. <<http://www.wshoy.sidar.org/index.php?2005/12/09/30-ontologias-que-son-y-para-que-sirven>> [Consulta: 13/04/2011].
- ACT. 2005. *Estudio de situación de mercado español de servicios profesionales de traducción*. Madrid: SpainBase.
- Agirre, Ansa, O.; Hovy, E.H. y Martínez, D. 2000. «Enriching very large ontologies using the WWW», en S. Staab, A. Maedche, C. Nedllec y P. Wiemer-Hastings (eds.). *Proceedings of the Workshop on Ontology Learning of the European Conference on Artificial Intelligence (ECA100)*. CEUR Workshop Proceedings. Vol. 31. Berlín. 25-30.
- ; E.; Arregi, X.; Artola, X.; Díaz De Ilarraza, A.; Sarasola, K. y Soroa, A. 2000. «A Methodology for Building Translator-oriented Dictionary Systems». *Machine translation*, 15 (4). 295-310.

- Ahmad, K. 1995. «Pragmatics of Specialist Terms: The Acquisition and Representation of Terminology», en P. Steffens (ed.) *Machine Translation and the Lexicon. Proceedings of the 3rd International EAMT Workshop*. Berlín y Nueva York: Springer Verlag. 51-76.
- ; Holmes-Higgin, P. y Sibte Abidi, R. 1994. «A description of texts in a corpus: “virtual” and “real” corpora», en W. Martin, W. Meijs, M. Moerland, E. Ten Pas, P. Van Sterkenburg y P. Vossen (eds.). *EURALEX 1994 Proceedings*. Ámsterdam: Vrije Universiteit. 390-402.
- ; Martin, W.; Hölter, M. y Rogers, M. 1995. *Specialist Terms in General Language Dictionaries* [en línea], University of Surrey Technical Report CS-95-14. <ftp://ftp.ee.surrey.ac.uk/pub/research/AI/pointer/diction.ps.gz> [Consulta: 13/04/2011]]
- y Rogers, M. 2001. «Corpus Linguistics and Terminology Extraction», en S. E. Wright y G. Budin (eds.). Vol. 2. 725-760.
- Alcaraz Varó, E. 2000. *El inglés profesional y académico*. Madrid: Alianza.
- , Hughes, B. y Campos, M. Á. 2006. *Diccionario de términos de turismo y ocio: Inglés-Español, Spanish-English*. 2ª ed. act. Barcelona: Ariel.
- y Martínez Linares, M^a. A. 1997. *Diccionario de lingüística moderna*. Barcelona: Ariel.
- Alcina Caudet, M. A. 2009. «Metodología y tecnologías para la elaboración de diccionarios terminológicos onomasiológicos», en A. Alcina, E. Valero y E. Rambla (eds.). 33-58.
- ; Soler Puertes, V. y Estellés Palanca, A. 2005. «Internet como instrumento para la documentación en terminología y traducción», en D. Sales Salvador (ed.), *La Biblioteca de Babel. Documentarse para traducir*. Granada: Comares. 221-242.
- ; Valero, E. y Rambla, E. (eds.) 2009. *Terminología y sociedad del conocimiento*. Berna: Peter Lang.
- Almahano Güeto, I. 2002. *El contrato de viaje combinado en alemán y español: las condiciones generales. Un estudio basado en corpus*. Tesis doctoral. Málaga: Universidad de Málaga.
- Alonso Ímaz, M. C. 2005. «El lenguaje de la publicidad en la prensa alemana actual. Anglicismos y otros recursos», en E. Hernández y L. Sierra (ed.), *Lenguas para fines específicos*. Vol. VIII. 197-207. Madrid: Universidad de Alcalá.
- Alvar Ezquerro, M. 1980. «¿Qué es un diccionario? Al hilo de unas definiciones académicas», en *LEA*, II/1, 103-118.
- 1996. «Los diccionarios del español en nuestros días», en *Cuadernos Cervantes*, 11 nov.-dic. 9-14.
- y Corpas Pastor, G. 1994. «Criterios de diseño para la creación de corpora», en M. Alvar Ezquerro y J. A. Villena Ponsoda (coords.). *Estudios para un corpus del*

- español*. Anejo n.º 7 de *Analecta Malacitana*. Málaga: Servicio de Publicaciones de la Universidad de Málaga.
- Amat Noguera, N. 1978. *Técnicas documentales y fuentes de información*. Barcelona: Bibliograf.
- Ananiadou, S. 1994. «A Methodology for Automatic Term Recognition». *Proceedings of the 15th International Conference on Computational Linguistics (COLING'94)*. Kyoto (Japón). 1034-1038.
- Antia, B.E. 2000. *Terminology and language planning: an alternative framework of practice and discourse*. Amsterdam y Filadelfia: John Benjamins.
- Arano, S. 2003. «La ontología: una zona de interacción entre la Lingüística y la Documentación» [en línea]. *Hipertext.net*, n.º 2. <<http://www.hipertext.net>> [Consulta: 13/04/2011].
- Ard, J. 1982. «The use of bilingual dictionaries by ESL students while writing». *ITL, Review of Applied Linguistics*, 58, 1-27.
- Arnaudov, T. y Mitkov, R. 2008. «Smarty - Extendable Framework for Bilingual and Multilingual Comprehension Assistants». *Proceedings of the 6th International Conference on Language Resources and Evaluation, LREC'08*. Marrakech. 3287-3292.
- Arntz, R. y Picht, H. 1995. *Introducción a la terminología*. Madrid: Fundación Germán Sánchez Ruipérez. Madrid: Pirámide, D.L.
- Arribas Baño, M. A. 2004. *La representación del significado en el diccionario bilingüe de negocios: una propuesta de tratamiento*. Tesis doctoral. Universidad de Valladolid.
- Aspas Aspas, J. M. 2000. *Los deportes de aventura. Consideraciones jurídicas sobre el turismo activo*. Sta. Cruz de Tenerife: Prames.
- . 2004. «Los deportes de aventura, ¿deporte o turismo?», en F. J. Melgosa (ed.). *Derecho y Turismo*. Salamanca: Ediciones Universidad Salamanca. 289-325.
- Atkins, S. (ed.) 1998. *Using Dictionaries. Studies of Dictionary Use by Language Learners and Translators. Lexicographica Series Mayor 88*. Tubinga.
- , Clear, J. y Ostler, N. 1992. «Corpus Design Criteria». *Literary and Linguistic Computing*, 7(1). 1-16.
- y Rundell, M. 2008. *The Oxford Guide to Practical Lexicography*. Oxford: Oxford University Press.
- y Varantola, K. 1998. «Language Learners Using Dictionaries: the Final Report on the EURALEX/AILA. Research Project on Dictionary Use», en Atkins, S. (ed.) 1998. 21-82.
- Auger, P. 1988. «La terminologie au Québec et à travers le monde», en *OLF-STQ. Actes du sixième colloque OLF-STQ de terminologie. L'ère nouvelle de la terminologie*, 27-59. Quebec: Gouvernement du Québec.

- y Rousseau, L. 1987. *Metodologia de la recerca terminologica*. Sant Adrià de Besòs: Departament de cultura de la Generalitat de Catalunya.
- Auger, A. 1997. *Repérage des énoncés d'intérêt définitoire dans le bases de données textuelles* [en línea]. Tesis doctoral. Universidad de Neuchâtel. <<http://doc.rero.ch/record/473?ln=fr>> [Consulta: 13/04/2011].
- y Barrière, C. 2008. «Pattern-based approaches to semantic relation extraction». *Terminology*, 14 (1). 1-19.
- Austermühl, F. 2001. *Electronic Tools for Translators*. Manchester: St. Jerome.
- Baker, M. 1993. «Corpus Linguistics and Translation Studies: Implications and Applications», en M. Baker, G. Francis y E. Tognini-Bonelli (ed.). *Text and Technology: In Honour of John Sinclair*. Ámsterdam y Filadelfia: John Benjamins.
- . 1995. «Corpus Linguistics and Translation Studies: An Overview and Some Suggestions for Future Research». *Target*, 7 (2), 223-243.
- Bajo Pérez, E. 2000. *Los diccionarios. Introducción a la lexicografía del español*. Gijón: TREA.
- Barlow, M. 1996. «Corpora for Theory and Practice». *International Journal of Corpus Linguistics*, vol. 1 (1). 1-38.
- Barrière, C. 2001. «Investigating the Causal Relation in Informative Texts». *Terminology*, 7 (2). 135-154.
- . 2002. «SeRT - A tool for Knowledge Extraction from Text». *Comunicación en Computational Linguistics in the North East (CLiNE)*. Montreal (Canadá).
- . 2004. «Building a concept hierarchy from corpus analysis». *Terminology*, 10 (2). 241-264.
- . 2006. «Semi-Automatic Corpus Construction from Informative Texts», en L. Bowker (ed.). 81-92.
- y Agbago, A. 2006. «TerminoWeb: A software environment for term study in Rich contexts». *Proceedings of the International Conference on Terminology, Standardisation and Technology Transfer (TSTT'2006)*. 103-113. Pekín (China).
- Baugh, S.; Harley, A. y Jellis, S. 1996. «The role of corpora in compiling the Cambridge International Dictionary of English». *International Journal of Corpus Linguistics*, 1. 39-59.
- Beckwith, R. 2005. *Web Page Evaluation Form* [en línea]. Bowling Green (Ohio): Bowling Green State University. <<http://www.bgsu.edu/offices/acen/writerslab/handouts/webeval.pdf>> [Consulta: 13/04/2011].
- Beeby, A. 2000. «Evaluating the Development of Translation Competence», en C. Schäffner y B. Adab (eds.). *Developing Translation Competence*. Ámsterdam: John Benjamins. 185-198.
- ; Berenguer, L.; Ensinger, D.; Fox, O.; Hurtado Albir, A.; Martínez Melis, N.; Orozco, M. y Presas, M. 2001. «La competencia traductora y su adquisición» [en

- línea]. *Quaderns. Revista de Traducció* 6, 39-45.
<<http://www.fti.uab.es/pacte/publicacions/Quaderns2001.pdf>> [Consulta: 13/04/2011]
- Bejoint, H. 1981. «The foreign student's use of monolingual English dictionaries: A study of language needs and reference skills». *Applied Linguistics*, 2/3. 207-222.
- Bell, R. T. 1991. *Translation and Translating*. Londres: Longman.
- Benbow, T.; Carrington, P.; Johannesen, G.; Tompa, F.W. y Weiner, E. 1990. «Report on the New Oxford English Dictionary User Survey». *International Journal of Lexicography*, 3 (3). 155-204.
- Benvenuti, N. 2002a. «La catalogazione delle risorse di internet» [en línea], en P. Capitán (ed.). *Semantica e terminologia nei portali. Tavola rotonda. Firenze* 11-12 Dicembre 2000. Florencia: Firenze University Press. 93-102.
<http://eprints.unifi.it/archive/00000155/00/Semantica_e_terminologia_nei_portali.pdf> [Consulta: 13/04/2011].
- . 2002b. «Die moderne lexikographische Funktionslehre. Diskussionsbeitrag zu neuen und alten Paradigmen, die Wörterbücher als Gebrauchsgegenstände verstehen». *Lexicographica. International Annual for Lexicography*, 18. 253-263.
- Berners-Lee, T. Hendler, J. y Lassila, O. 2001. «The Semantic Web». *Scientific American Magazine*.
- Bergenholtz, H. y Tarp, S. 1995. *Manual of specialised lexicography: the preparation of specialised dictionaries*. Ámsterdam y Filadelfia: John Benjamins.
- Biber, D. 1993. «Representativeness in Corpus Design». *Literary and Linguistic Computing*. 8 (4) 243-257..
- . 1994. «Representativeness in Corpus Design», en A. Zampolli, N. Calzolari y M. Palmer (eds.). *Current Issues in Computational Linguistics: In Honour of Don Walker*. Dordrech y Pisa: Kluwer y Giardini. 377-408.
- ; Conrad, S. y Reppen, R. 1998. *Corpus Linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.
- Blommfield, L. (1933). *Language*. Nueva York: Holt, Rinehart and Winston.
- Bogaards, P. 2005. «A propos de l'usage du dictionnaire de langue étrangère». *Cahiers de Lexicologie*, 52 (1). 131-152.
- Borrueco, R. M. 2007. «El texto publicitario como recurso pedagógico de aprendizaje significativo», en M. R. Borrueco (ed.). *El lenguaje publicitario en el turismo*. Junta de Andalucía. Consejería de Turismo, Comercio y Deporte.
- Boulangier, J. C. 1995. «Présentation: images et parcours de la socioterminologie». *Meta*, 40. 195-205.
- Bourigault D. 1992. «Surface Grammatical Analysis for the Extraction of the Terminological Noun Phrase». *Proceedings of the 14th International Conference on Computational Linguistics (COLING'92)*. Nantes (Francia). 977-981.

-
- . 1995. «LEXTER, a Terminology Extraction Software for Knowledge Acquisition from Texts». *Proceedings of the 9th Knowledge Acquisition for Knowledge Based System Workshop (KAW'95)*. Banff (Canadá).
- ; Jacquemin, C. y L'Homme, M.C. (eds.). 2001. *Recent Advances in Computational Terminology*. Amsterdam y Filadelfia: John Benjamins.
- Bowker, L. 1996. «Towards a Corpus-based Approach to Terminography». *Terminology*, 3 (1). 27-52.
- . 1997. «Multidimensional Classification of Concepts and Terms», en S. E. Wright y G. Budin (eds.). Vol. 1. 133-143.
- (ed.). 2006. *Lexicography, Terminology and Translation. Text-based studies in honour of Ingrid Meyer*. Ottawa: Ottawa University Press.
- y Pearson, J. 2002. *Working with Specialized Language: A practical guide to using corpora*. Londres: Routledge.
- Brekke, M. 2001. «LSP Lexicography and Terminography: A Complementary View», en F. Meyer (ed.). *LSP: Perspectives for New Millenium*. Tubinga: Gunter Narr Verlag.
- Brethauer, P. 2000. «Das Anforderungsprofil für Fachübersetzer/-dolmetscher und Voraussetzungen für Ihren bidirektionalen Einsatz», en M. Grosman, M. Kadric, I. Kovačić y M. Snell-Hornby (eds.). 147-152.
- Brosson, F. 2008. *Contenu du module « Le Français du Tourisme »* [en línea]. Agence universitaire de la Francophonie. <<http://fromulus.chez-alice.fr/documents/tourisme.pdf>> [Consulta: 13/04/2011].
- Brunette, L. 2000. «Towards a Terminology for Translation Quality Assessment. A comparison of TQA Practices». *The Translator*, 6 (2). 169-182.
- Bugnot, M. A. 2005. *Texto turístico y traducción especializada. Estudio crítico de un corpus español-francés sobre la Costa del Sol (1960-2004)*. Tesis doctoral. Málaga: Servicio de Publicaciones Universidad de Málaga.
- Cabré Castellví, M. T. 1993. *La terminología. Teoría, metodología, aplicaciones*. Barcelona: Antártida/Empúries.
- . 1998a. «Las Fuentes Terminológicas para la Traducción», en P. Nistal y J.M^a. Bravo Gozalo (coords.). *La Traducción: Orientaciones Lingüísticas y Culturales*. Universidad de Valladolid: SAE. 27-59.
- . 1998b. «Elementos para una teoría de la terminología: hacia un paradigma alternativo». *El Lenguaraz. Revista académica del Colegio de Traductores Públicos de la ciudad de Buenos Aires*. 1 (1). 59-78.
- . 1998c. «Una nueva teoría de la terminología: de la denominación a la comunicación», en M. T. Cabré Castellví. 2000. 109-127.
- . 1998d. «¿Es necesaria una teoría autónoma de la terminología?», en M. T. Cabré Castellví. 2000. 93-108.

-
- , 1999a. «Fuentes de información terminológica para el traductor», en M. Pinto y J. A. Cordón (ed.), *Técnicas documentales aplicadas a la traducción*. Madrid: Síntesis. 19-40.
- , 1999b. «Hacia una teoría comunicativa de la terminología: aspectos metodológicos», en M. T. Cabré. 2000a. 129-150.
- , 2000a. *La Terminología: Representación y Comunicación. Elementos para una teoría de base comunicativa y otros artículos*. Barcelona: IULA. Universidad Pompeu Fabra.
- , 2000b. «El traductor y la terminología. Necesidades y compromiso». *Panace@*, 1 (2). 2-4.
- , 2002. «Textos especializados y unidades de conocimiento: metodología y tipologización», en J. García y M. T. Fuentes (eds.). *Texto, terminología y traducción*. Barcelona: Almar. 122–187.
- , 2003. «Teorías de la terminología: de la prescripción a la descripción», en G. Adamo y V. Della Valle (eds.). *Innovazione lessicale e terminologie specialistiche. Serie Lessico Intellettuale Europeo*, 92. Florencia: Leo S. Olschki Editore. 168-188.
- , 2004a. «Estado actual de la investigación en terminología», en P. A. Fuertes Olivera (coord.). *Lengua y Sociedad: Aportaciones recientes en Lingüística Cognitiva, Lenguas en Contacto, Lenguajes de especialidad y Lingüística del Corpus*. Valladolid: Universidad de Valladolid. 193-203.
- , 2004b. «¿Lenguajes especializados o lenguajes para propósitos específicos?», en A. Van Hooft (dir.). *Textos y discursos de especialidad: el español de los negocios. Revista Foro Hispánico*, 26. 19-34.
- , 2004c. «La terminología en la traducción especializada», en C. Gonzalo García y V. García Yebra (eds.). 89-125.
- , 2005. «La Terminología, una disciplina en evolución: pasado, presente y algunos elementos de futuro» [en línea]. *Revista Debate Terminológico*, 1 (3) <http://www.riterm.net/revista/n_1/cabre.pdf> [Consulta: 13/04/2011].
- , 2006. «From Terminological Data Banks to Knowledge Databases: The Text as the Starting Point», en L. Bowker (ed.). 93-106.
- , 2007. «Términos y palabras en los diccionarios», en J. Cuartero y M. Otal Emsel (eds.). *Vernetzungen: Bedeutung in Wort, Satz und Text. Festschrift für Gerd Wotjak zum 65. Geburtstag*. Fráncfort: Peter Lang, 71-84.
- , 2011. «Presente y futuro de los estudios de terminología». *Ponencia de inauguración de las IX Jornadas de AETER*. Soria, 26-27 de noviembre.
- ; Bach, C.; Estopà, R.; Feliu, J.; Martínez, G. y Vivaldi, J. 2004. «The GENOMA-KB project: towards the integration of concepts, terms, textual corpora and entities». *IV International Conference on Language Resources and Evaluation*. 87-90.
- ; Doménech, M.; Morel, J. y Rodríguez, C. 2001. «Las características del conocimiento especializado y la relación con el conocimiento general», en

- IULATERM. La terminología científico-técnica.* Barcelona: Instituto Universitario de Lingüística Aplicada. Universitat Pompeu Fabra. 173-186.
- ; Estopà Bagot, R. y Vivaldi Palatresi, J. 2001. «Automatic Term Detection: A review of current systems», en D. Bourigault, C. Jacquemin y M.C. L’Homme (eds.). 53-88.
- Cáceres González, A. E. 2008. «La métrica de Levenshtein». *Revista de Ciencias Básicas*, 7 (2). 35-43.
- Calvi, M. V. 2000. *El léxico del turismo.* Universidad de Bérgamo [en línea]. <<http://www.ub.es/filhis/culturele/turismo.html>> [Consulta: 13/04/2011].
- . 2001. ‘El español del turismo: problemas didácticos’, *IV Congrés Internacional sobre llengües per a finalitats específiques. The language of International Communication, Special Seminar: Español de los negocios.* Publicacions de la Universitat de Barcelona. 299-304.
- . 2006a. «El uso de términos culturales en el lenguaje del turismo: los hoteles y su descripción», en M. V. Calvi y L. Chiechichetti, *Nuevas tendencias en el discurso de especialidad.* Berna: Peter Lang. 271-292.
- . 2006b. *Lengua y comunicación en el español del turismo.* Madrid: Arco Libros.
- y Bonomi, M. 2008. «El lenguaje del turismo: de los textos especializados a la Comunidad del viajero», en C. Navarro, R.M. Rodríguez Abella, F. dalle Pezze y R. Miotti (eds.). *La comunicación especializada.* Berlín: Peter Lang. 181-202.
- Carroll, J. 2003. «Parsing», en R. Mitkov (ed.). 233-248.
- Casas Gómez, M. 2006. «Modelos representativos de documentación terminográfica y su aplicación a la Terminología lingüística». *Revista de Lingüística y Lenguas Aplicadas*, 1. 25-36.
- Castillo Rodríguez, C. M. 2010. *El trabajo terminográfico en traducción: explotación de un corpus multilingüe de turismo de salud y belleza (español-francés/inglés/italiano).* Tesis doctoral. Universidad de Málaga.
- Chamosa, J. L. 1997. *Memes of Translation. The Spread of Ideas in Translation Theory.* Ámsterdam/Filadelfia: John Benjamins.
- Chanod, J.P. 1998. «Natural Language Processing at the Xerox Research Centre Europe» [en línea], en T. Erjavec y J. Gros (eds.). *JEZIKOVNE TEHNOLOGIJE ZA SLOVENSKI JEZIK: zbornik konference Language Technologies for the Slovene Language: proceedings of the conference.* Jozef Stefan Institute, Ljubljana (Eslovenia). <<http://nl.ijs.si/isjt98/zbornik/sdjt98-Chanod.pdf>> [Consulta: 13/04/2011].
- Chesterman, A. (1993). «From ‘Is’ to ‘Ought’: Laws, Norms and Strategies in Translation Studies». *Target*, 5 (1). 1-20.
- Chomsky, N. 1957. *Syntactic Structures.* La Haya: Mouton.
- . 1965. *Aspects of a Theory of Syntax.* Cambridge: MIT Press.

- Ciapuscio, G. 2003. *Textos especializados y terminología*. Barcelona: Instituto Universitario de Lingüística Aplicada, Universidad Pompeu Fabra.
- Cimiano, P. 2006. *Ontology learning and population from text: algorithms, evaluation and applications*. Berlín: Springer Verlag.
- Codina, L. 2000a. «Parámetros e indicadores de calidad para la evaluación de recursos digitales». *Actas de las VII Jornadas Españolas de Documentación (7: 2000 Bilbao)*. *La Gestión del Conocimiento: retos y soluciones de los profesionales de la información*. Bilbao: Universidad del País Vasco. 135-144.
- . 2000b. «Evaluación de recursos digitales en línea: conceptos, indicadores y métodos». *Revista Española de Documentación Científica*. 23 (1). 9-44.
- Cohen, E. 1972. «Toward a sociology of international tourism». *Social Research*, 39. 164-182.
- y Lefebvre, C. 2005. *Handbook of Categorisation in Cognitive Linguistics*. Elsevier.
- y Cooper, R. 1986. «Language and tourism». *Annals of Tourism Research*, 13. 533-563.
- Collados Aís, A. 1998. *La evaluación de la calidad en Interpretación simultánea. La importancia de la comunicación no verbal*. Granada: Comares Interlingua.
- Comité Européen de Normalisation (CEN). *European Standard. Translation Services – Service Requirements*. <<http://www.statsaut-translator.no/Files/Standard-15038-final-draft-en.pdf>> [Consulta: 13/04/2011].
- Condamines, A. 1994. «Terminologie et représentation des connaissances». *La Banque des Mots, numéro spécial 1994*. 29-44.
- . 1995. «Terminology: New Needs, new perspectives». *Terminology*, 2 (2), 219-238.
- y Rebeyrolle, J. 2001. «Searching for and Identifying Conceptual Relations via a Corpus-based Approach to a Terminological Knowledge Base (CTKB): Methods and Results», en D. Bourigault, C. Jacquemin y M.C. L’Homme (eds.). 127-148.
- Conde Ruano, J. T. 2009. *Proceso y resultado de la evaluación de traducciones*. Tesis doctoral. Universidad de Granada.
- Conrad, S. y Biber, D. 2000. «Adverbial marking of stance in speech and writing», en S. Hunston y G. Thompson (eds.). *Evaluation in Text: Authorial stance and the construction of discourse*. Oxford: Oxford University Press, 57-63.
- Consejería de Turismo, Comercio y Deporte. 2005. *Plan de marketing del turismo en Andalucía 2005* [en línea]. Junta de Andalucía. <http://www.juntadeandalucia.es/turismocomercioydeporte/ctcd-docs/custom_doc/PLAN_de_MARKETING_2005.pdf> [Consulta: 13/04/2011].
- . 2006a. *Plan de Calidad Turística de Andalucía 2006-2008*. Junta de Andalucía. <http://www.juntadeandalucia.es/turismocomercioydeporte/ctcd-docs/custom_doc/261_976_res_ejecutivo_pcta_snpf.pdf> [Consulta: 13/04/2011].
- . 2006b. *Plan Director de Marketing Turístico de Andalucía 2006-2008* [en línea]. Junta de Andalucía.

- <http://www.juntadeandalucia.es/turismocomercioydeporte/opencms/planificacion/turismo/plan_0008.html> [Consulta: 13/04/2011].
- . 2008. *Plan General del Turismo Sostenible de Andalucía 2008-2011* [en línea]. <http://www.juntadeandalucia.es/turismocomercioydeporte/export/sites/ctcd/archivos/planificacion/turismo/367_pgts_definitivo_con_medidas_11_octubre_2007.pdf> [Consulta: 13/04/2011].
- Contreras, J. y Martínez Comeche, J. A. 2008. *Tutorial ontologías* [en línea]. <http://www.sedic.es/gt_normalizacion_tutorial_ontologias.pdf> [Consulta: 13/04/2011].
- Corcho, O., Fernández-López, M., Gómez-Pérez A. y López-Cima, A. 2005. «Building Legal Ontologies with METHONTOLOGY and WebODE», en R. Bejamins, R. Casanovas, P. Gangemi y P. Selic (eds.). *Law and the Semantic Web*. Springer Verlag. 142-157.
- Cordón García, J. A.; López Lucas, J. y Vaquero Pulido, J. R. 2001. *Manual de investigación bibliográfica y documental. Teoría y práctica*. Madrid: Pirámide.
- Cornellà, A. 1999. «A mayor desarrollo informacional, menor infoxicación» [en línea]. *El profesional de la Información*, vol. 8 (9). <http://www.elprofesionaldelainformacion.com/contenidos/1999/septiembre/a_mayor_desarrollo_informacional_menor_infoxicacion.html> [Consulta: 13/04/2011].
- Corpas Pastor, G. 1997. *Manual de fraseología española*. Madrid: Gredos.
- . 2001. «Compilación de un corpus *ad hoc* para la enseñanza de la traducción inversa». *TRANS*, 5. 155-179.
- . 2002. «Traducir con corpus: de la teoría a la práctica», en M. García Palacios y T. Fuentes Morán (eds.), *Texto, terminología y traducción*. Salamanca: Almar. 189-226.
- . 2003a. «Turicor: Compilación de un corpus de contratos turísticos (alemán, español, inglés, italiano) para la generación textual multilingüe y la traducción jurídica», en J. M. Peinado y E. Ortega Arjonilla (coord.). *Actas del II Simposio Internacional Traducción, Texto e Interferencias. La investigación y la práctica profesional de la Traducción e Interpretación en los albores del siglo XXI*. Granada: Comares.
- . (ed.). 2003b. *Recursos documentales y técnicos para la traducción del discurso jurídico (español, alemán, inglés, italiano, árabe)*. Granada: Comares.
- . 2003c. *Diez años de investigación en fraseología: análisis sintáctico-semánticos, contrastivos y traductológicos*. Madrid: Lingüística Iberoamericana.
- . 2004a. «La traducción de textos médicos especializados a través de recursos electrónicos y corpus virtuales», en L. González y P. Hernández (eds.). *Las palabras del traductor. Actas del II Congreso Internacional "El español, lengua de traducción"*. Bruselas: Comisión Europea/ESLETRA. 137-164.
- . 2004b. «Localización de recursos y compilación de corpus vía Internet: aplicaciones para la didáctica de la traducción médica especializada», en C. Gonzalo García y V. García Yebra (eds.). 223-258.

- , 2008. *Investigar con corpus en traducción: los retos de un nuevo paradigma*. Fráncfort am Main: Peter Lang.
- , 2010. *La contratación turística electrónica multilingüe como mediación intercultural: aspectos legales, traductológicos y terminológicos* (HUM-892. Junta de Andalucía. 2006-2009). Memoria final del proyecto, remitida a la D.G. Investigación, Tecnología y Empresa de la Consejería de Economía, Innovación y Ciencia de la Junta de Andalucía.
- ; Leiva Rojo, J. y Varela Salinas, M. J. 2001. «El papel del diccionario en la formación de traductores e intérpretes: análisis de necesidades y encuestas de uso», en M. C. Ayala Castro (ed.). *Diccionarios y enseñanza*. Alcalá: Universidad de Alcalá.
- y Seghiri Domínguez, M. 2006. *El concepto de representatividad en la Lingüística del Corpus: aproximaciones teóricas y metodológicas*. Documento Técnico. Universidad de Málaga.
- y Seghiri Domínguez, M. 2007a. «Determinación del umbral de representatividad de un corpus mediante el algoritmo N-Cor». *Procesamiento del lenguaje natural*, 39. 165-172.
- y Seghiri Domínguez, M. 2007b. «Specialized Corpora for Translators: A Quantitative Method to Determine Representativeness» [en línea]. *Translation Journal*, 11 (3). <<http://www.translationjournal.net/journal/41corpus.htm>>. [Consulta: 13/04/2011]
- y Seghiri Domínguez, M. 2009. «Virtual Corpora as Documentation Resources: Translating Travel Insurance Documents (English-Spanish)», en A. Beeby, P. Rodríguez Inés y P. Sánchez Gijón (eds.). *Corpus Use and Translating*. Ámsterdam y Filadelfia: John Benjamins. 75-107.
- y Seghiri Domínguez, M. 2011/en prensa. *El concepto de representatividad en lingüística de corpus: aproximaciones teóricas y consecuencias para la traducción*. Málaga: Servicio de Publicaciones de la Universidad de Málaga.
- Costa, A.; Iniesta, A. y Torres, J.A. 1999. «Turismo activo y deportivo», en F. Bayón Mariné (coords.). *50 años de turismo español: un análisis histórico y estructural*. Madrid: Centro de Estudios Ramón Areces. 749-780.
- Croft, W. y Cruse, D. A. 2004. *Cognitive Linguistics*. Cambridge: Cambridge University Press.
- Crowther, J. (ed.). 1995. *Oxford Advanced Learner's Dictionary of Current English*. 5ª ed. Oxford: Oxford University Press.
- Cuenca, M. J. y Hilferty, J. 1999. *Introducción a la lingüística cognitiva*. Barcelona: Ariel.
- Cumming, G., Cropp, S. y Sussex, R. 1994. «On-line lexical resources for language learners: Assessment of some approaches to word definition». *System*, 22/3. 369-377.
- Dancette, J. 1995. *Parcours de traduction, étude expérimentale du processus de compréhension*. Lille: Presses Universitaires de Lille.

- y Castrillón, R. 2009. «Un diccionario enciclopédico y ontológico: el campo de la migración transnacional». *ONOMÁZEIN*, 20 (2). 65-86.
- y L'Homme, M. C. 2001. «Modélisation des relations sémantiques dans un dictionnaire spécialisé bilingue». *VI Journées scientifiques de l'AUPELF-UREF*, Beyrouth. 385-399.
- Dancette, J. y C. Réthoré. 1997. «Le dictionnaire bilingue (Anglais-Français) de la distribution: entre dictionnaire de langue et encyclopédie». *Meta* XLII (2). 229–243.
- y C. Réthoré. 2000. *Dictionnaire Analytique de la Distribution. Analytical Dictionary of Retailing*. Montreal: Les Presses de l'Université de Montréal.
- Dann, G. 1996. *The language of Tourism. A Sociolinguistic Perspective*. Oxon: CAB International.
- . 2007. «Revisiting the language of tourism: what tourists and tourees are saying», en C. de Stasio y O. Palusci (eds.). *The Language of Tourism. Turismo e Mediazione*. Milán: Unicopli. 15-32.
- De Baer, P., Kerremans, K. y Temmerman, R. 2006. «Developing Special Language Dictionaries with the Termontography Workbench», *12th EURALEX Conference* (demostración de software), Turín.
- de Bessé, B. 1997. «Terminological Definitions», en S.E. Wright y G. Budin (eds). Vol. 1. 63-74.
- de Carvalho, P. G. y Lourenço, R. 2009. «Turismo de prática desportiva: um segmento do mercado do turismo desportivo». *Revista Portuguesa de Ciências do Desporto*, 9 (2). 122-132.
- de Juan Alonso, J. M. 2004. «El turismo activo», en S. Antón Clave. *La actividad turística española en 2003*. Jaén: Asociación Española de Expertos Científicos de Turismo.
- de la Torre Padilla, F. 1989. *Introducción al Estudio del Turismo*. México: Continental, S.A.
- De Rooze, B. 2003. *La traducción, contra reloj*. Tesis doctoral. Granada: Universidad de Granada.
- Dejean, H. y Gaussier, E. 2002. «Une nouvelle approche a l'extraction de lexiques bilingues a partir de corpus comparables». *Lexicométrica*. 1–21.
- Delisle, J. 1980. *Translation: An Interpretive Approach*. Ottawa y Londres: University of Ottawa Press.
- . 1993. *La traduction raisonnée. Manuel d'initiation à la traduction professionnelle de l'anglais vers le français*. Ottawa: Les Presses de l'Université d'Ottawa.
- Denny, M. 2004. *Ontology Building: A Survey of Editing Tools* [en línea]. <http://ie.technion.ac.il/OntoBuilder/Data/10.OntoBuilder_Papers/Ontology_Editor_Survey_2004_Table_-_Michael_Denny.pdf> [Consulta: 13/04/2011].
- Deutsches Institut für Normung (DIN) 2331. 1993. *Begriffe und Benennungen; Allgemeine Grundsätze*. Berlín, Colonia: Beuth.

- Deville, G. 1989. *Modelization of Task-oriented Utterances in a Man-machine Dialogue System*. Tesis doctoral. Universidad de Amberes.
- y Herbigniaux, E. 1995. «Natural Language Modeling in a Machine Translation Prototype for Healthcare Applications: A Sublanguage Approach» [en línea]. *MTI-95. Proceedings of the Sixth International Conference on Theoretical and Methodological Issues in Machine Translation*. Lovaina (Bélgica): Katholieke Universiteit. 142-157. <<http://www.mt-archive.info/TMI-1995-Deville.pdf>> [Consulta: 13/04/2011].
- Drouin, P. 2004. «Detection of Domain Specific Terminology Using Corpora Comparison». *Proceedings of the fourth international Conference on Language Resources and Evaluation (LREC)*. Lisboa.
- Druetta, R. 2007. *Types de communication et types de textes* [en línea]. Universidad de Turín. <http://www.francaisunivers.unito.it/documents/types_druetta.pdf> [Consulta: 13/04/2011].
- Dubuc, R. 1980. *Manuel pratique de terminologie*. Montreal: Linguattech.
- y Lauriston, A. 1997. «Terms and Contexts», en S. E. Wright y G. Budin (eds.). Vol. 1. 80-88.
- . 2002. *Manuel pratique de terminologie*. Montreal: Linguattech.
- Durán Muñoz, I. 2008a. «La traducción turística y la evaluación de su calidad», en P. Sánchez Hernández, P. Pérez-Paredes, P. Aguado Jiménez y R. Criado Sánchez (eds.). *Researching and teaching specialized languages: New contexts, new challenges. Proceedings of the AELFE 2008 Conference*. Murcia: Editum, Universidad de Murcia. 374-383.
- . 2008b. *Catalogación y evaluación de las fuentes terminológicas del sector turístico (alemán-español)*. Málaga: Universidad de Málaga.
- . 2008c. «La necesidad de profesionalización en la traducción turística». *I Congreso Internacional Translation in the Era of Information*. Oviedo, 22-24 de octubre de 2008.
- . 2010a. «Specialized lexicographical resources: a survey of translators' needs», en S. Granger y M. Paquot (eds.). *eLexicography in the 21st century: New Challenges, new applications. Proceedings of ELEX2009. Cahiers du Cental*, vol. 7. Lovaina-La-Nueva. Presses Universitaires de Louvain. 55-66.
- . 2010b. «Herramientas para la elaboración de recursos ontoterminográficos: estado de la cuestión», en M. Ibáñez Rodríguez (ed.). *Lenguas de especialidad y terminología*. Granada: Comares. 111-123.
- . 2012/en prensa. «Traducción de textos turísticos promocionales: características y necesidades». *Hermeneus*.
- Duineveld, A. J., Stoter, R., Weiden, M.R., Kenepa, B. y Benjamins, V.R. 2000. «Wondertools? A comparative study of ontological engineering tools». *International Journal of Human-Computer Studies*, 52 (6). 1111-1133.
- Duvå, G. y Laursen, A. L. 1994. «Translation and LSP Lexicography: A User Survey», en H. Bergenholtz y S. Tarp (eds.). *Manual of Specialised Lexicography. The*

- preparation of specialised dictionaries*. Ámsterdam y Filadelfia: John Benjamins. 247-267.
- EAGLES. 1994. *Corpus Typology: A framework for classification*. Documento técnico de EAGLES 080294. 1-18.
- . 1995. *Evaluation of Natural Language Processing Systems* [en línea]. Documento técnico de EAGLES EAG-EWG-PR.2. <<http://www.issco.unige.ch/en/research/projects/ewg95//ewg95.html>> [Consulta: 13/04/2011].
- . 1996a. *Preliminary Recommendations on Corpus Typology* [en línea]. Documento técnico de EAGLES EAG-TCWG-CTYP/P. <<http://www.ilc.cnr.it/EAGLES96/corpusyp/corpusyp.html>> [Consulta: 13/04/2011].
- . 1996b. *Recommendations for the Morphosyntactic Annotation of Corpora* [en línea]. Documento técnico de EAGLES EAG-TCWG-MAC/R. <<http://www.ilc.cnr.it/EAGLES/annotate/annotate.html>> [Consulta: 13/04/2011].
- East, E. 2008. «Dictionary Use in Foreign Language Writing Exams. Impact and implications». *Language Learning & Language Teaching*, 22.
- Estopà, R. 1999. *Extracció de terminologia: elements per a la construcció d'un SEACUSE (Sistema d'Extracció Automàtica de Candidats a Unitats de Significació Especialitzada)*. Tesis Doctoral. Universitat Pompeu Fabra, Institut Universitari de Lingüística Aplicada.
- European Committee for Standardisation (CEN). *European Standard. Translation Services – Service Requirements* (versión en inglés) [en línea]. <<http://www.statsaut-translator.no/Files/Standard-15038-final-draft-en.pdf>>. [Consulta: 13/04/2011]
- Evans, V.; Bergen, B. K. y Zinken, J. 2007a. «The Cognitive Linguistics Enterprise: An Overview», en V. Evans, B. K. Bergen y J. Zinken (ed.). 2-36.
- , Bergen, B. K. y Zinken, J. (eds.) 2007b. *The Cognitive Linguistics Reader*. Londres: Equinox Publishing.
- Evans, V y Green, M. 2006. *Cognitive Linguistics: An Introduction*. Edimburgo: Edinburgh University Press.
- Faatz, A. y Steinmetz, R. 2002. «Ontology Enrichment with Texts from the WWW». *Proceedings of the 2nd Workshop on Semantic Text Mining*. Helsinki. 20-35.
- Faber Benítez, P. 2002. «Investigar en Terminología», en P. Faber Benítez y C. Jiménez Hurtado. 3-23.
- . 2004. «PUERTOTERM: la representación del conocimiento y la generación de recursos terminológicos dentro del dominio de la ingeniería de costas», en M. Á. García Peinado y E. Ortega Arjonilla (coord.). *Panorama actual de la investigación en traducción e Interpretación*. Granada: Atrio. 385-398.
- . 2009. «Terminología multilingüe y ontologías», en A. Alcina; E. Valero y E. Rambla (eds.). 289-308.

- y Jiménez Hurtado, C. (eds.). 2002. *Investigar en terminología*. Granada: Comares.
- ; León Araúz, P.; Prieto Velasco, J. A y Reimerink, A. 2007. «Linking images and words: the description of specialized concepts». *International Journal of Lexicography*, 20. 39-65
- ; León Araúz, P. y Prieto Velasco, J. A. 2009. «Semantic Relations, Dynamicity, and Terminological Knowledge Bases» [en línea]. *Current Issues in Language Studies*, 1. 1-23. <<http://www.academicpress.us/Journals/ebooks/J3/3.pdf>> [Consulta: 13/04/2011].
- ; López Rodríguez, C. I. y Tercedor Sánchez, M. I. 2001. «La utilización de técnicas de corpus en la representación del conocimiento médico». *Terminology*, 7 (2). 167-197.
- ; Márquez Linares, C. y Vega Expósito, M. 2005. «Framing Terminology: A Process-Oriented Approach» [en línea]. *META: Translators' Journal*, 50 (4). <<http://www.erudit.org/revue/meta/2005/v50/n4/019916ar.pdf>> [Consulta: 13/04/2011].
- ; Montero Martínez, S.; Castro Prieto, M. R.; Senso Ruiz, J.; Prieto Velasco, J. A.; León Arauz, P.; Márquez Linares, C. y Vega Expósito, M. 2006. «Process-oriented terminology management in the domain of Coastal Engineering». *Terminology* 12 (2). 189-213.
- Fauconnier, G & Turner, M. 2002. *The Way We Think: Conceptual blending and the Mind's Hidden Complexities*. Nueva York: Basic Books.
- Felber, H. 1984. *Terminology Manual*. Viena: Infoterm.
- Feldweg, H. 1997. «Accessing Foreign Languages with COMPASS». *Machine Translation*, 12. 153–174.
- . y Breidt, E. 1996. «COMPASS An Intelligent Dictionary System for Reading Text in a Foreign Language», en F. Kiefer y G. Kiss (eds.). *Papers in Computational Lexicography*. COMPLEX '96. Budapest (Hungría). 53-62.
- Feliu Cortés, J. 2004. *Relacions conceptuals i terminologia: Anàlisi i proposta de detecció semiautomàtica*. Tesis doctoral. Universidad Pompeu Fabra, Barcelona.
- ; Vivaldi Palatresi, J. y Cabré Castellví, M.T. 2002. *Ontologies: a review*. [en línea]. Barcelona: IULA. Universitat Pompeu Fabra. <<http://www.iula.upf.edu/repositori/02inf034.pdf>> [Consulta: 13/04/2011].
- Fennell, D.A. 1999. *Ecotourism: An Introduction*. Londres: Routledge.
- Fernández, M.; Gómez Pérez, A. y Juristo, N. 1997. «METHONTOLOGY: From Ontological Art Towards Ontological Engineering Workshop on Ontological Engineering». *Spring Symposium Series. AAAI97*, Stanford, USA.
- Fernández Arroyo, R. 1996. *El país frágil. Las montañas deben sobrevivir*. Desnivel, Madrid.

- Fernández Sola, A. F. 2011. *Aplicación para la consulta terminológica en el proceso de traducción*. Proyecto Fin de Carrera (lectura prevista para septiembre de 2011). Departamento de Lenguas y Ciencia de la Computación. Universidad de Málaga.
- Fletcher, W. H. 2004. «Facilitating the Compilation and Dissemination of Ad-Hoc Web Corpora» [en línea], en G. Aston, S. Bernardini y D. Stewart (eds.). *The Fifth International Conference on Teaching and Language Corpora*. Ámsterdam: Benjamins. 1-18.
<http://www.kwicfinder.com/Facilitating_Compilation_and_Dissemination_of_Ad-Hoc_Web_Corpora.pdf> [Consulta: 13/04/2011].
- Fillmore, C. 1968. «The case for case», en Bach, E. y Harms, R. (ed.). *Universals in Linguistic Theory*. Nueva York: Holt, Rinehart and Winston. 1-88.
- . 1977. «Scenes and Frames Semantics», en Zampolli, A. (ed.) *Linguistic Structures Processing (Fundamental Studies in Computer Science, 59)*. Ámsterdam: North Holland Publishing. 55-88.
- . 1982. «Frame Semantics», en The Linguistic Society of Korea (ed.), *Linguistics in the Morning Calm*. Seúl: Hanshin. 109-118.
- . 1985. «Frames and the semantics of understanding». *Quaderni di Semantica*, 6. 222-254.
- . 2003. «Double-decker definitions: The Role of Frames in Meaning Explanations». *Sign Language Studies*, 3 (3). 263-295.
- . 2007. «Frame Semantics», en V. Evans, B. K. Bergen y J. Zinken. 238-261.
- y Atkins, S. 1992. «Towards a Frame-Based Lexicon: the Semantics of Risk and its Neighbors. Frames, Fields and Contrasts», en A. Lehrer y E. Kittay (ed.) *New Essays in Semantic and Lexical Organization*. Hilldale.
- Firth, J. R. 1935. «The Technique of Semantics». *Transactions of the Philological Society*. 36-72.
- . 1957a. «A synopsis of Linguistic Theory, 1930-1955». *Studies in Linguistic Analysis*. Número especial de *Philological Society*. 1-32.
- . 1957b. *Papers in Linguistics*. Londres: Oxford University Press.
- Fischer, M. B. 2000. «Sprachgefühl und Welterfahrung - La traducción inversa de textos turísticos como ejercicio para fomentar la competencia lingüística» [en línea]. *Cultura e intercultura en la enseñanza del español como lengua extranjera*, 1 (1). <<http://www.ucm.es/info/especulo/ele/alcala.html>> [Consulta: 13/04/2011].
- Fischer, D. H., Möhr, W. y Rostek, L. 1996. «A modular, object-oriented and generic approach for building terminology maintenance systems», en C. Galinski y K. D. Schmitz (eds.) *Proceedings of the 4th International Conference on Terminology and Knowledge Engineering (TKE-96)*. Viena: INDEKS Verlag. 245-258.
- Fleischer, W. y Barz, I. 1995. *Wortbildung der deutschen Gegenwartssprache*. Tübinga: Max Niemeyer Verlag.

- Flowerdew, L. 2000. «Investigating referencial and pragmatic errors in a learner corpus», en L. Burnard y T. McEnery (eds.). *Rethinking Language Pedagogy from a Corpus Perspective*. Fráncfort: Peter Lang, 145-154.
- Fuentes Luque, A. 2005. «La traducción de promoción turística institucional: la proyección de la imagen de España», en A. Fuentes Luque (ed.), *La traducción en el sector turístico*. Granada: Atrio. 59-92
- Fuertes Olivera, P.A. y Velasco Sacristán, M. 2001. «A Critical Comparison of the Macrostructure and Microstructure of two Bilingual English-Spanish Dictionaries of Economics». *International Journal of Lexicography*, 14 (1): 31-55.
- Fundación EUROPARC. 2007. *Carta Europea de Turismo Sostenible en los espacios protegidos* [en línea]. <http://www.juntadeandalucia.es/medioambiente/web/Bloques_Tematicos/Patrimonio_Natural._Uso_Y_Gestion/Espacios_Protegidos/Parques_Naturales/Carta_europea_turismo_sostenible/Texto_CETS_oficial_2007.pdf> [Consulta: 13/04/2011].
- Fung, P. 1995. «Compiling bilingual lexicon entries from a non-parallel English-Chinese corpus». *Proceedings of the 14th Annual Meeting of Very Large Corpora*. Boston, Massachusetts. 173-183.
- Fuster, J. y Elizalde, B. 1995. «Riesgo y actividades físicas en el medio natural: un enfoque multidimensional». *Apunts: Educación Física y Deportes*, 41. Barcelona. 94-107.
- Gak, V.G. 1991. «Para una tipología de las tipologías de diccionarios», *Voz y Letra*, II/1. 105-115.
- Gamallo Otero, P. 2007. «Learning Bilingual Lexicons from Comparable English and Spanish Corpora». *Proceedings of Machine Translation Summit XI*, Copenhagen. 191-198.
- . 2008. «Evaluating two different methods for the task of extracting bilingual lexicons from comparable corpora». *Proceedings of LREC 2008 Workshop on Comparable Corpora*, Marrakech. 19-26.
- y Pichel Campos, J.R. 2005. «An Approach to Acquire Word Translations from Non-Parallel Text». *12th Portuguese Conference on Artificial Intelligence (EPIA '05)*, Évora (Portugal).
- y Pichel, J.R. 2008. «Learning Spanish-Galician Translation Equivalents Using a Comparable Corpus and a Bilingual Dictionary», *Lecture Notes in Computer Science*, vol. 4919, Springer Verlag. 423-433.
- Gambier, Y. 1991. «Travail et vocabulaires spécialisés: prolegomènes à une socioterminologie». *Meta*, 36 (1). 8-1.
- Gamero, S. 1998. *La traducción de textos técnicos (alemán-español)*. Géneros y subgéneros. Tesis doctoral. Universitat Autònoma de Barcelona.
- Gamper, J., Nejdil, W. y Wolpers, M. 1999. «Combining Ontologies and Terminologies in Information Systems» [en línea]. *TKE'99 Terminology and Knowledge*

- Engineering*. Insbruck. <<http://www.kbs.uni-hannover.de/paper/99/tke99/>> [Consulta: 13/04/2011].
- Garau, J.B.; Serra, A. y Robledo, M.A. 2000. «El turismo activo en destinos turísticos tradicionales». *AECIT: La actividad turística española en 1999*. Madrid: AECIT. 587-606.
- García de Quesada, M. 2001. *Estructura definicional terminográfica en el subdominio de la oncología clínica*. Vol. 14. Madrid: CSIC/Elies.
- García Palacios, J. 2003. «La terminología en sus diccionarios: terminología, diccionarios de especialidad y traducción», en E. Ortega Arjonilla (dir.), *Panorama actual de la investigación en Traducción e Interpretación*. Vol. I. Granada: Ábaco. 529-540.
- Gaudin, F. 1993. *Pour une socioterminologie. Des problèmes sémantiques aux pratiques institutionnelles*. Rouen: Université de Rouen.
- Geeraerts, D. (ed.). 2006. *Cognitive Linguistics: Basic Readings*. Berlín: Mouton de Gruyter.
- Gile, D. 1995. *Basic Concepts and Models for Interpreter and Translator Training*. Amsterdam: John Benjamins.
- Gillam, L. y Ahmad, K. 1996. «Knowledge-engineered terminology (data)bases», en C. Galinski y K. D. Schmitz (eds.). *Proceedings of the 4th International Conference on Terminology and Knowledge Engineering (TKE-96)*. Vienna. INDEKS Verlag. 205-214.
- , Tariq, M. y Ahmad, K. 2007. «Terminology and the construction of ontology», en F. Ibekwe-San Juan (ed.). *Application-Driven Terminology Engineering*. Amsterdam: John Benjamins. 49-73.
- Givón, T., 1995. *Functionalism and grammar*. Amsterdam/Filadelfia: John Benjamins.
- Godman, A. y Payne, E.M.F. 1981. «A taxonomic approach to the lexis of science», en L. Selinker et. al (eds.). *English for Academic and Technical Purposes: Studies in Honour of Louis Timble*. Rowley MA: Newbury House. 23-39.
- Goffman, E. 1974. *Frame Analysis*. Nueva York: Harper and Row.
- Gómez González-Jover, A. 2006. *Terminografía, lenguajes profesionales y mediación interlingüística. Aplicación metodológica al léxico especializado del sector industrial del calzado y de las industrias afines*. Tesis doctoral. Universidad e Alicante.
- y Vargas Sierra, C. 2004. «Aspectos metodológicos para la elaboración de diccionarios especializados bilingües destinados al traductor». *El español, lengua de traducción. II congreso internacional*. Bruselas: ESLETRA. 365-398.
- Gómez Guinovart, X. 2000. «Lingüística computacional», en Ramallo, F.; Rei-Doval, G. y Rodríguez, X. P. (ed.). *Manual de ciencias da linguaxe*. Vigo: Xerais. 221-268.

- y Sacau Fontenla, E. 2004. «Métodos de optimización de la extracción de léxico bilingüe a partir de corpus paralelos». *Procesamiento del lenguaje natural*, 33. 133-140.
- Gómez Pérez, A.(coord.). 2002. *Deliverable 1.3.: A survey on ontology tools* [en línea]. OntoWeb. <http://www.york-sure.de/publications/OntoWeb_Del_1-3.pdf> [Consulta: 13/04/2011].
- , Fernández López, M. y Corcho, Ó. 2003. *Ontological engineering: with examples from the areas of knowledge management, e-commerce and the Semantic Web*. Londres: Springer Verlag.
- , Fernández, M. y De Vicente, A. 1996. «Towards a Method to Conceptualize Domain Ontologies». *Workshop on Ontological Engineering. ECAI'96*. 41–51.
- Gonzalo García, C. 2004 [2003]. «DocuTradSo en línea: Fuentes de información para la actividad traductora» [en línea]. <<http://www.uva.es/DocuTradSo>> [Consulta: 13/04/2011].
- . 2004. «Fuentes de información en línea para la traducción», en C. Gonzalo García y V. García Yebra (eds.). 275-336.
- y Fraile Vicente, E. 2004. «Selección y evaluación de recursos lingüísticos en Internet para el traductor especializado», en C. Gonzalo García y V. García Yebra (eds.). 337-372.
- Gonzalo García, C. y García Yebra, V. (eds.). *Manual de documentación y terminología para la traducción especializada*. Madrid: Arco/Libros, S.L.
- Goytia Prat, A. 1995. «El papel de la animación turística en el turismo sostenible». *Turismo y tiempo libre: actividades, métodos y organización*. Universidad de Deusto, Bilbao. 202-217.
- . 2000. «Nuevas tendencias de ocio y turismo» [en línea]. *Travel Turisme*. Consellería de Turisme. Generalitat Valenciana. <<http://www.travelturisme.com/estudios/pdf/ocio-turismo.pdf>> [Consulta: 13/04/2011].
- Grosman, M., Kadric, M., Kovačič, I. y Snell-Hornby, M. (eds.). 2000. *Translation into Non-Mother Tongues. In Professional Practice and Training*. Tubinga: Stauffenburg Verlag.
- Gómez Guinovart, X. y Simões, A. 2009. «Parallel corpus-based bilingual terminology extraction». *Proceedings of the 8th International Conference on Terminology and Artificial Intelligence*, IRIT (Institut de recherche en Informatique de Toulouse). Université Paul Sabatier, Toulouse (Francia).
- Gruber, T. R. 1993. «A Translation Approach to Portable Ontologies». *Knowledge Acquisition*, 5 (2). 199-220.
- Ha, L. A., Fernández, G., Mitkov, R. y Corpas Pastor, G. 2008. «Mutual bilingual terminology extraction» [en línea]. *Proceedings of LREC'08*. Marrakech. 1-7. <http://clg.wlv.ac.uk/papers/463_paper.pdf> [Consulta: 13/04/2011].

- , Mitkov, R. y Corpas Pastor, G. 2008. «Mutual Terminology Extraction Using a Statistical Framework». *Procesamiento del lenguaje natural*, 41. 107-112.
- Haensch, G. 1997. *Los diccionarios del español en el umbral del siglo XXI*. Salamanca: Universidad de Salamanca.
- Halaoui, N. 1990. «Questions de méthode en terminologie des langues africaines». *Terminologies nouvelles*, 3. 5-24.
- Halliday, M. A. K. 1992. «Language as system and language as instance: the corpus as a theoretical construct», en J. Svartvik (ed.). *Directions in Corpus Linguistics. Proceedings of Nobel Symposium 82*, Estocolmo, 4-6 agosto. Berlín: Mouton de Gruyter. 61-78.
- Hanks, P. 1996. «Contextual Dependency and Lexical Sets». *International Journal of Corpus Linguistics*, vol. 1 (1). 75-98.
- . 2003. «Lexicography», en R. Mitkov (ed.). 48-69.
- Hardman, D. 2000. *The Reltef Specification* [en línea]. <<http://www.ttt.org/clsframe/reltef.html>> [Consulta: 13/04/2011].
- Hartmann, R. R. K. 1982. «Das zweisprachige Wörterbuch im Fremdsprachenerwerb. Studien zur neuhochdeutschen Lexicographie II», en H. E. Wiegand (ed.). *Germanistische Linguistik*, 3-6/80. Hildesheim: Georg Olms. 73-86.
- . 1983. «The bilingual learner's dictionary and its uses». *Multilingua*, 2 (4). 195-201.
- . 1995. «Pedagogical lexicography: some desiderata», en R. Dirven y J. Vanparrys (ed.). *Current approaches to the lexicon*. Fráncfort: Peter Lang. 405-411.
- y James, G. 1998. *Dictionary of Lexicography*. Londres: Routledge.
- . 2001. *Teaching and Researching Lexicography*. Harlow: Longman.
- Hatim, B. y Mason, I. 1997. *The translator as communicator*. Londres: Routledge.
- Hausmann, F.J. 1989-1991. *Wörterbücher. Dictionaries. Dictionnaires. Ein internationales Handbuch zur Lexikographie. An International Encyclopedia of Lexicography. Encyclopédie internationale de lexicographie*. Berlín-Nueva York: De Gruyter.
- Hawkins, D.E. y Lamoureux, K. 2001. «Global Growth and Magnitude of Ecotourism», en D. Weaver (ed.). 63-72.
- Hearst, M.A. 1992. «Automatic acquisition of hyponyms from large text corpora». *Proceedings of the 15th International Conference on Computational Linguistics (COLING-92)*. Nantes (Francia). 539-545.
- Heid, U. 1999. «A Linguistic Bootstrapping Approach to the Extraction of the Term Candidates from German Texts». *Terminology*, 5 (2). 161-181.
- Hernández Hernández, H. 2000. «El diccionario en la enseñanza de ELE. (Diccionarios de español para extranjeros)». *Actas del XI Congreso Internacional de ASELE*, Zaragoza. 93-103.

- Hernández Ramírez, H. y Saiz Noeda, M. 2007. «Ontologías mixtas para la representación conceptual de objetos de aprendizaje». *Procesamiento del lenguaje natural*, 38. 99-106.
- Herrero Jiménez, I. 2004. *Desarrollo de ontologías sobre bases de datos relacionales: ontología que soporta la búsqueda avanzada de información* [en línea]. Proyecto fin de carrera. Universidad Carlos III de Madrid. <http://e-archivo.uc3m.es/bitstream/10016/5493/1/PFC_Ignacio_Herrero_Jimenez_2004.pdf> [Consulta: 13/04/2011]
- Hewson, L. y Martin, J. 1991. *Redefining Translation. The variational approach*. Londres: Routledge.
- Higgleton, E. y Seaton, A. (eds.). 1995. *Times-Chambers Essential English Dictionary*. Edimburgo y Singapur: Chambers and Federal Publications.
- Hoffmann, L. 1976. *Kommunikationsmittel Fachsprache. Eine Einführung*. Berlín: Akademie-Verlag.
- . 1985. *Kommunikationsmittel Fachsprache*. Tübinga: Gunter Narr.
- . 1998. «Característiques dels llenguatges d'especialitat», en J. Brumme (ed.), *Llenguatges d'especialitat. Selecció de textos*. Barcelona: IULA. 21-69.
- Holmes, J. 1988. «The Name and Nature of Translation Studies». *Translated! Papers on Literary Translation and Translation Studies*. Rodopi: Ámsterdam. 67-80.
- Horguelin, P. A. 1985. *Pratique de la révision*. 2^a ed. Montreal: Linguattech.
- House, J. 1997. *Translation Quality Assessment: A model Revisited*. Tübinga: Narr.
- Hudson, S. 2003. *Sport and Adventure Tourism*. Nueva York: The Haworth Press.
- Hulstijn, J. H. y Atkins, S. 1998. «Empirical Research on Dictionary Use in Foreign-Language Learning: Survey and Discussion», en Atkins, S. (ed.). 1998. 121-141.
- Hunston, S. 2002. *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.
- Hurtado Albir, A. 1990. *La notion de fidélité en traduction*. París: Didier Érudition.
- . 1996. «La enseñanza de la traducción directa “general”. Objetivos de aprendizaje y metodología», en A. Hurtado Albir (ed.). *La enseñanza de la traducción*. Castellón de la Plana: Universitat Jaume I. 31-56.
- (dir.). 1999. *Enseñar a traducir. Metodología en la formación de traductores e intérpretes. Teoría y fichas prácticas*. Madrid: Edelsa.
- . 2001. *Traducción y Traductología*. Madrid: Cátedra.
- Instituto de Turismo de España (TURESPAÑA). 2008. *Estudio de productos turísticos. 9. Turismo de montaña* [en línea]. <<http://www.tourspain.es/NR/rdonlyres/BE4C3474-E960-4150-9316-60FD56181D82/18739/turismomontana200994.pdf>> [Consulta: 13/04/2011].
- Instituto Nacional de Estadística. 2006. *Boletín informativo del Instituto Nacional de Estadística* [en línea], vol. 3. <<http://www.ine.es>> [Consulta: 13/04/2011].

-
- International Organization for Standardization (ISO). ISO 3166. 1974. *Maintenance Agency (ISO 3166/MA) - ISO's focal point for country codes*. Ginebra: ISO.
- . ISO 1087. 1990. *Terminology-Vocabulary*. Ginebra: ISO.
- . ISO 12620. 1999. *Computer applications in terminology - Data categories*. Ginebra: ISO.
- . ISO 12200. 1994. *Computational aids in terminology - Terminology Interchange Format (TIF) - An SGML Application*. Ginebra: ISO.
- . ISO 1087-1. 2000. *Terminology Work – Vocabulary – Part 1: Theory and application*. Ginebra: ISO.
- . ISO 704. 2000. *Terminology Work. Principles and methods*. Ginebra: ISO.
- . ISO 639-1. 2002. *Codes for the representation of names of languages - Part 1: Alpha-2 code*. Ginebra: ISO.
- . ISO 16642. 2003. *Computer applications in terminology -- Terminological markup Framework*. Ginebra: ISO.
- . ISO 30042. 2008. *Systems to manage terminology, knowledge and content -- TermBase eXchange (TBX)*. Ginebra: ISO.
- IULA. 2008. *Metodología de trabajo en terminología*. Taller en línea. Barcelona: IULA. Universidad Pompeu Fabra.
- . 2010. *Resolución de problemas terminológicos*. Taller en línea. Barcelona: IULA. Universidad Pompeu Fabra.
- Jakobson, R. 1984. *Ensayos de lingüística general*. Barcelona: Ariel.
- Johansson, S. 1998. «On the Role of Corpora in Cross-Linguistic Research», en S. Johansson y S. Oksefjell (eds.). *Corpora and Cross-Linguistic Research*. Ámsterdam: Rodopi. 3-24.
- . 2007. «Contrastive Linguistics and Corpora», en Teubert, W. y Krishnamurthy, R. *Corpus Linguistics. Critical Concepts in Linguistics*. Vol. IV. Routledge: Londres/Nueva York. 213-226.
- Jung, L. 2000. *La escuela traductológica de Leipzig*. Granada: Comares.
- Kade, O. 1981. «Zufall und Gesetzmäßigkeiten in der Übersetzung». *Zeitschrift für Übersetzer, Dolmetscher und Sprachkundige*, 3. 70-90.
- Kay, M. 2003. «Introduction», en R. Mitkov (ed.). XVII-XX.
- Kelly, D. 1997. «The translation of texts from the tourist sector: textual conventions, cultural distance and other constraints». *TRANS*, 2. 33-42.
- . 2002. «Un modelo de competencia traductora: bases para el diseño curricular». *Puentes. Hacia nuevas investigaciones en la mediación intercultural*, 1. 9-20.
- . 2005. «‘Let’s Periko Ortega give you a sweet ride...’ o la urgente necesidad de profesionalizar la traducción en el sector turístico. Algunas propuestas para

- programasa de formación», en A. Fuentes Luque (ed.). *La traducción en el sector turístico*. Granada: Atrio. 155-170.
- Kernerman, L. 1996. «English learners' dictionaries: How much do we know about their use?», en M. Gellerstam, J. Järborg, S. G. Malmgren, K. Norén, L. Rogström y C. Røjder (ed.) *EURALEX 1996 Proceedings*. Papehnl: Göteborg.
- Kerremans, K., Temmerman, R. y Tummers, J. 2003. «Representing multilingual and culture-specific knowledge in a VAT regulatory ontology: support from the termontology approach», en *Lecture Notes in Computer Science*, 2889. 662-674.
- Kilgarriff, A. 2001. «Web as Corpus». *Proceedings of Corpus Linguistics 2001*. Lancaster. 342-344.
- Kilgarriff, A. y Grefenstette, G. 2003. «Introduction to the Special Issue on Web as Corpus». *Computational Linguistics*, 29 (3). 1-15.
- Kiraly, D. C. 1995. *Pathways to Translation. Pedagogy and Process*. Kent: The Kent State University Press.
- Kirk, E. 1996. «Evaluating Information Found on the Internet». *Baltimore: The Johns Hopkins University* [en línea]. <<http://www.library.jhu.edu/researchhelp/general/evaluating/>> [Consulta: 13/04/2011].
- Klavans, J. y Tzoukermann, E. 1990. «Combining Lexical Information from Bilingual Corpora and Machine-Readable Dictionaries». *Proceedings of the 13th International Conference on Computational Linguistics (COLING'92)*. Helsinki, Finlandia.
- Kleiber, G. 1995. *La Semántica de los prototipos. Categoría y sentido léxico*. Madrid: Visor.
- Kokourek, R. 1982. «Textes et Termes». *Meta*, 36 (1).
- Königs, H. G. 1989. «Die Übersetzerrelevante Textanalyse», en F. G. Königs (ed.), *Übersetzungswissenschaft und Fremdsprachenunterricht. Neue Beiträge zu einem alten Thema*. München: Goethe Institut. 121-145.
- Francis, W. N. y Kucěra, H. 1979 [1964]. *Brown corpus manual: Manual of information* [en línea]. 3.^a rev y amp. Providence (Rhode Island, EE.UU.): Brown University. <<http://gandalf.aksis.uib.no/icame/brown/bcm.html>> [Consulta: 13/04/2011].
- Labarta Postigo, M. «La importancia de la expresión oral y del aprendizaje de alemán para turismo» [en línea]. *III Encuentro de profesores de formación turística universitaria. ANESTUR*. <http://www.anestur.com/8.Maria_Labarta_Postigo.doc> [Consulta: 13/04/2011]
- Lakoff, G. 1982. *Categories and Cognitive Models*. Cognitive Science Program, Institute of Cognitive Studies. Universidad de California, Berkeley (California).
- . 1987. *Women, Fire and Dangerous Things. What Categories reveal about the Mind*. Chicago, Londres: The University of Chicago Press.
- y Johnson, M. 1980. *Metaphors We Live By*. Chicago: Universidad de Chicago.

- Lamarca Lapuente, M. J. 2006. *Hipertexto, el nuevo concepto de texto en la cultura de la imagen* [en línea]. Tesis doctoral. Madrid: Universidad Complutense de Madrid. <<http://www.hipertexto.info/>> [Consulta: 13/04/2011].
- Langacker, R. W. 1987. *Foundations of Cognitive Grammar*. Stanford: Stanford University Press.
- Larsen, L. 2001. *Information Literacy: The Web is not an Encyclopedia* [en línea]. <<http://www.oit.umd.edu/units/web/literacy/>> [Consulta: 13/04/2011].
- Laurén, C. y Picht, H. 1993. *Ausgewälte Texte zur Terminologie*. Viena: Termnet. IITF Infoterm.
- Lavid, J. 2005. *Lenguaje y nuevas tecnologías. Nuevas perspectivas, métodos y herramientas para el lingüista del siglo XXI*. Madrid: Cátedra.
- Latiesa, M., Rebollo, S. y Paniza, J.L. 2002. «Deporte, turismo y salidas profesionales», en Latiesa, M., Rebollo, S. y Paniza, J.L. (coord.). *Salidas profesionales en el campo del turismo deportivo*. Instituto Andaluz del Deporte. Junta de Andalucía, Málaga. 15-33.
- Le Serrec, A., L’Homme, M.C., Drouin, P. y Kraif, O. 2010. «Automating the compilation of specialized dictionaries». *Terminology*, 16 (1). 77–106.
- Lehrberger, J. 1986. «Sublanguage Analysis», en Grishman, R. y Kittredge, R. (eds.). *Analyzing Language in Restricted Domains*. Hillsdale, NJ: Lawrence Erlbaum Associates, 19-38.
- Lerat, P. 1988. «Terminologie et sémantique descriptive». *La banque des mots*, número especial. 11–30.
- . 1997. *Las lenguas especializadas*. Barcelona: Ariel.
- Leech, G. 1991. «The State of the Art in Corpus Linguistics», en K. Aijmer y B. Altenberg (eds.). *English Corpus Linguistics*. Londres: Longman, 8-29.
- . 1992. «Corpora and Theories of Linguistic Performance», en Svartvik, J. (ed.) (1992). *Directions in Corpus Linguistics. Proceedings of Nobel Symposium 82*, Estocolmo, 4-8 Agosto 1991. Berlín/Nueva Yor: Mouton de Gruyter. 105-134.
- . 2007. «Corpora», en Teubert, W. y Krishnamurthy, R. (ed.). Vol. 2.
- Lemay, C., L’Homme, M. C., Drouin, P. 2005. «Two methods for extracting “specific” single-word terms from specialized corpora: Experimentation and Evaluation». *International Journal of Corpus Linguistics*, 10 (2). 227-253.
- Lemnitzer, L. y Zinsmeister, H. 2006. *Korpuslinguistik: Eine Einführung*. Tubinga: Gunter Narr Verlag.
- Lenat, D.B. 1995. «Cyc: A Large-Scale Investment in Knowledge Infrastructure». *Communications of the ACM*, 38. 33-38.
- León Guzmán, K. y Parra Boyero, M. 2001. «Nuevas tendencias deportivas: Deportes de Sliz» [en línea]. *Educación física y deportes*, 6 (30). <<http://www.efdeportes.com/efd30/liz.htm>> [Consulta: 13/04/2011].

- L'Homme, M. C. 2004. «A lexico-Semantic Approach to the Structuring of Terminology». *Proceedings of 20th International Conference on Computational Linguistics (COLING'04)*. Ginebra. 7-14.
- . 2007. «Using Explanatory and Combinatorial Lexicology to Describe Terms», en L. Wanner. *Selected Lexical and Grammatical Topics in the Meaning-Text Theory. In Honour of Igor Mel'cuk*. Ámsterdam y Filadelfia: John Benjamins. 11-50.
- y Dancette, J. 2001. «Modélisation des relations sémantiques dans un dictionnaire spécialisé bilingüe», en Clas, A. et al. (ed.). *L'éloge de la différence: la voix de l'autre. Actes des 6^e Journées du Réseau LTT*. Actualité scientifique, Paris: AUF, 385-400.
- León Araúz, P., Reimerink, A. y Faber Benítez, P. 2009. «Puertoterm & Marcocosta: A Frame-based Knowledge Base for the Environmental Domain». *Journal of Multicultural Research*, 1. 47-70.
- Levenshtein, V. 1966. «Binary Codes Capable of Correcting Deletions, Insertions, and Reversals». *Sovietic Physics Doklady*, 10 (8). 707-710.
- Lew, R. 2011. «Studies in dictionary use: Recent developments». *International Journal of Lexicography*, 24 (1). 1-4.
- Localization Industry Standards Association (LISA). 2005. *LISA Terminology Management Survey: Terminology Management Practices and Trends* [en línea]. <<http://www.lisa.org/Terminology-Management.461.0.html>> [Consulta: 13/04/2011].
- . 2008. *Systems to manage terminology, knowledge, and content - TermBase eXchange (TBX)* [en línea]. <http://www.lisa.org/fileadmin/standards/TBX_2008_10_29.pdf> [Consulta: 13/04/2011].
- . 2008. *TBX-Basic* [en línea]. <https://www.socialtext.net/data/workspaces/terminology-sig/attachments/tbx_basic:20081024215407-0-19440/original/TBX_Basic_datacategoriesV2.pdf> [Consulta: 13/04/2011].
- Londero, R. 2006. «Traducir la publicidad para el turismo: referencias culturales en anuncios españoles e italianos», en M. V. Calvi y L. Chiexichetti (ed.). *Nuevas tendencias en el discurso de especialidad*. Berna: Peter Lang. 249-270.
- López Rodríguez, C.I., Tercedor Sánchez, M. y Faber Benítez, P. 2006. «Gestión Terminológica Basada en el Conocimiento y Generación de Recursos de Información sobre el Cáncer: el Proyecto ONCOTERM» [en línea]. *Revista e-Salud*, vol. 2 (8). 228-240. <<http://www.revistaesalud.com/index.php/revistaesalud/article/view/127/325>> [Consulta: 13/04/2011].
- Lorán, R. 2005. *Los buscadores son cada vez más útiles e imprescindibles* [en línea]. <<http://bibliotecavirtualut.suagm.edu/Publicaciones%20profesores/Roberto%20Lor%C3%A1n/Los%20buscadores%20son%20cada%20vez%20m%C3%A1s%20%C3%BAtiles%20e%20imprescindibles.pdf>>. 1-3. [Consulta: 13/04/2011].

- Lorente Casafont, M. 2001. «Teoría e innovación en terminografía: la definición terminográfica», en IULATERM. *La terminología científico-técnica*. Barcelona: Instituto Universitario de Lingüística Aplicada. 81-112.
- . 2005. «Ontología sobre economía y recuperación de información» [en línea]. *Hipertext.net*, 3. <<http://www.hipertext.net>> [Consulta: 13/04/2011].
- Lozano Tello, A. 2001. «Ontologías en la Web Semántica» [en línea]. *I Jornadas de Ingeniería Web '01*. <<http://www.informandote.com/jornadasIngWEB/articulos/jiw02.pdf>> [Consulta: 13/04/2011].
- . 2002. *Métrica de idoneidad de las ontologías*. Tesis doctoral. Universidad de Extremadura.
- Lüdeling, A. y Kytö, M. (eds). 2008. *Handbook on Corpus Linguistics*. Berlín: Mouton de Gruyter.
- Luque Gil, A. M. 2003. *Las actividades recreativo-deportivas y el uso turístico del medio natural*. Tesis doctoral. Servicio de Publicaciones de la Universidad de Málaga: Málaga.
- Llamas Arjona, C. 2009. *Marketing y gestión de calidad turística*. Madrid: Liber Factory.
- Macías, M. 1999. «Nuevas tecnologías aplicadas a la distribución turística. Oficina telemática de servicios turísticos de Andalucía». Congreso *TuriTec'99: Turismo y Tecnologías de la Información y las Comunicaciones*. Málaga.
- Mackenzie, R. y Vienne, J. 2000. «Resource research strategy : A key factor in teaching translation into non-mother tongue», en M. Grosman, M. Kadric, I. Kovačič y M. Snell-Hornby (eds.). 125-132.
- Mackintosh, K. 1998. «An empirical study of dictionary use in L2-L1 Translation», en S. Atkins (ed.). 1998. 121-141.
- Magris, M. 2004. «Verso una terminografia per il traduttore giuridico». *Linguistica antverpiensia*, 3. 53-65.
- Maia, B. 2002. «Do-It-Yourself, Disposable, Specialised Mini Corpora – Where Next? Reflections on Teaching Translation and Terminology through Corpora» [en línea]. *Cadernos de Tradução*, vol. 9 (1). 219-234. <www.cadernos.ufsc.br/online/cadernos9/belindamaia.pdf> [Consulta: 13/04/2011].
- Maia, B. 2003. «Using Corpora for Terminology Extraction: Pedagogical and computational approaches», en B. Lewandowska-Tomasczczyk (ed.). *PALC 2001. – Practical Applications of Language Corpora*. Lodz Studies in Language. Fráncfort: Peter Lang. 147-164.
- Malaisé, V.; Zweigenbaum, P. y Bachimont, B. 2005. «Mining defining contexts to help structuring differential ontologies». *Terminology*, 11 (1). 21-53.
- Maldonado Martínez, Á. y Abejón Peña, T. 2004. «Herramientas de control del vocabulario científico-técnico: Glosarios y tesauros del CINDOC». *II Congreso Internacional El español, lengua de traducción*. Toledo.

- Malkiel, Y. 1975. «A Typological Classification of Dictionaries on the Basis of Distinctive Features», en F.W. Householder y S. Saporta (eds.), *Problems in Lexicography*. Bloomington. 3-24.
- Marcos Marín, F. y Sánchez Lobato, J. 1991. *Lingüística aplicada*. Madrid: Síntesis.
- Marinai, E.; Peters, C. y Picchi, E. 1991. «Bilingual Reference Corpora: A System for Parallel Text Retrieval in Using Corpora». *Proceedings for the 7th Annual Conference of the UW Centre for the New OED and Text Research*. Oxford: Oxford University Press. 63-70.
- Maroto García, M. N. 2007. *Las relaciones conceptuales en la terminología de productos cerámicos y su formalización mediante un editor de ontologías*. Tesis doctoral. Universitat Jaume I. Castellón.
- Marshman, E. 2008. «Expressions of uncertainty in candidate knowledge-rich contexts. A comparison in English and French specialized texts». *Terminology*, 14 (1). 124-151.
- y L’Homme, M.C. 2006. «Disambiguation of lexical markers of cause and effect», en H. Picht (ed.). *Modern Approaches to Terminological Theories and Applications. Proceedings of the 15th European Symposium on Language for Special Purposes, LSP 2005*. Berna: Peter Lang. 261-285.
- , Morgan, T. y Meyer, I. 2002. «French Patterns for Expressing Concept Relations». *Terminology*, 8 (1). 1-29.
- Martí Marco, M. R. 2006. «Ökosprache in Tourismus», en M. Borrucco Rosa (coord.), *La especialización lingüística en el ámbito turístico*. Junta de Andalucía. Consejería de Turismo, Comercio y Deporte.
- . 2007. «Turismo y Publicidad: Análisis del discurso publicitario de Turespaña», en M. Borrucco Rosa (coord.), *El lenguaje publicitario en el turismo*. Junta de Andalucía. Consejería de Turismo, Comercio y Deporte.
- Martín, J.; Ruiz, R.; Santaella, J. y Escáñez, J. 1996. *Los lenguajes especiales*. Granada: Editorial Comares.
- Martin, W.J.R. 1998. «Frames as definition models for terms», en A. Munteanu (ed.). *Proceedings of the International Conference on Professional Communication and Knowledge Transfer*, 2. Vienna: Termnet. 189-221
- Martínez de Sousa, J. 1999. *Manual de edición y autoedición*. Madrid: Ediciones Pirámide.
- . 2004. «La lexicografía especializada del español actual», en C. Gonzalo García y V. García Yebra (eds.), *Manual de documentación y terminología para la traducción especializada*. Madrid: Arco/Libros, S.L.173-190.

- Martínez López, F. J., P. Luna Huertas, R. Fernández Carrión y J. L. Salmerón Silvera. 2001 [1997] *Internet para investigadores: relación y localización de recursos para investigadores y universitarios*. 3.^a ed. rev. y ampl. Huelva: Universidad de Huelva.
- Martínez Melis, N. 1997. *Évaluation et traduction: Cadre de recherche sur l'évaluation dans la didactique de la traduction*. Tesis doctoral. Barcelona: Universidad Autónoma de Barcelona.
- y Hurtado Albir, A. 2001. «Assessment in Translation Studies: Research needs». *Méta*. 46 (2). 272-287.
- Martínez Motos, R. 2005. «Análisis comparativo de las convenciones textuales de textos turísticos en inglés y en español: los folletos editados por organismos oficiales», en A. Fuentes Luque (ed.), *La traducción en el sector turístico*. Granada: Atrio.
- Martínez Unanue, R. 1999. *Alineación automática de corpus paralelos: una propuesta metodológica y su aplicación a un dominio de especialidad*. Tesis doctoral. Universidad de Deusto.
- Maurel D. 2008. «Prolexbase: A multilingual relational lexical database of proper names». *Proceedings of LREC 2008*. Marrakech. 334-338.
- Mcalester, G. 2000. «The Evaluation of Translation into a Foreign Language», en C. Schäffner y B. Adab (eds.). *Developing Translation Competence*. Ámsterdam: John Benjamins. 229-242.
- McArthur, T. 1986. «Thematic Lexicography», en R. R. K Hartmann (ed.). *The History of Lexicography*. Ámsterdam y Filadelfia: John Benjamins. 157-166.
- McEnery, T. 2004. *Corpus Linguistics: An Introduction*. Edimburgo: Edinburgh University Press.
- y Wilson, A. 1996. *Corpus Linguistics*. Edimburgo: Edinburgh University Press.
- McTait, K. 2001. «Linguistic Knowledge and Complexity in an EBMT System Based on Translation Patterns» [en línea]. *Proceedings of the Workshop on Example-Based Machine Translation (EBMT)-MT Summit VIII*. Santiago de Compostela. <<http://www.mt-archive.info/MTS-2001-McTait.pdf>> [Consulta: 13/04/2011].
- Měchura, M. B. 2008. «Giving them what they want: Search strategies for electronic dictionaries» [CD-ROM]. *Proceedings of the 13th EURALEX International Congress*. Euralex 2008, Barcelona.
- Mejía García, L.; Zapata Giraldo, G.; Plested Alvarez, M. C.; María Campo, Á.; Muñoz Torres, A.; Castrillón Cardona, E. R. y Restrepo Marin, R. 1998. *Recomendaciones para el diseño y la configuración de bancos de datos terminológicos*. Medellín: Sociedad para la terminología y la Transferencia de Conocimiento.
- Melby, A. K. 1998. «Data Exchange Standards from the Oscar and MARTIF Projects», en *Proceedings of the First International Conference on Language Resources & Evaluation (LREC)*. Granada (España). 28-30 mayo. 3-7.
- . 2008. «TBX-Basic Translation-oriented Terminology Made Simple» [en línea]. *Revista Tradumàtica*, 6.

- <<http://www.fti.uab.es/tradumatica/revista/num6/articles/02/02.pdf>> [Consulta: 13/04/2011].
- Mel'cuk, I. 1997. *Vers une linguistique sens-texte* [en línea]. Ponencia inaugural. Collège de France, Chaire Internationale. <<http://olst.ling.umontreal.ca/pdf/melcukColldeFr.pdf>> [Consulta: 13/04/2011].
- Mel'cuk, I. y Zolkovsky, A.K. 1988. «The Explanatory Combinatorial Dictionary», en M. Evens (ed). *Relational Models in the Lexicon*. Cambridge: Cambridge University Press. 41-74.
- Melgosa Arcos, F.J. 2004. «El régimen jurídico-administrativo de los alojamientos rurales en España», en F. J. Melgosa Arcos (ed.). *Derecho y Turismo*. Salamanca: Ediciones Universidad Salamanca.
- Merlo Vega, J. A. 2003. «La evaluación de la calidad de la información web: aportaciones teóricas y experiencias prácticas», en F. Zapico Alonso (coord.). *Recursos informativos: creación, descripción y evaluación*. Mérida: Junta de Extremadura, Consejería de Educación, Ciencia y Tecnología. 101-110. <<http://exlibris.usal.es/merlo/escritos/pdf/calidad.pdf>> [Consulta: 13/04/2011]
- . 2004. «Uso de la documentación en el proceso de traducción especializada», en C. Gonzalo García y V. García Yebra (eds.). 309-336.
- Mesa Lao, B. 2009. «Estació TERMINUS». *Terminalia*, 0. 45.
- Meyer, I. 1992. «Knowledge management for terminology-intensive applications: needs and tools», en J. Pustejovsky y S. Bergler (eds.). *Lexical Semantics and Knowledge Representation*. Berlín: Springer Verlag. 21-37.
- . 2001. «Extracting knowledge-rich contexts for terminography», en D. Bourigault, C. Jacquemin y M.C. L'Homme (eds.). 279-302.
- , Eck, K. y Skuce, D. 1997. «Systematic Concept Analysis within a Knowledge-Based Approach to Terminology», en S. E. Wright y G. Budin (eds.). Vol. 1. 98-118.
- y Mackintosh, K. 1996. «The Corpus from a Terminographer's Viewpoint». *International Journal of Corpus Linguistics*, 1 (2). 257-285.
- ; Skuce, D.; Bowker, L. y Eck, K. 1992. «Towards a new generation of terminological resources: An experiment in building a terminological knowledge base». *Proceedings of the 15th International Conference on Computational Linguistics (COLING'92)*. 956-960.
- Mezyed Rivas, R. 2005. *Turismo rural y turismo activo en España* [en línea]. TalkIt: Madrid. <<http://www.talk-it.se/Turismo%20rural&activo%20talkit.pdf>> [Consulta: 13/04/2011].
- Mindt, D. 2000. *An Empirical Grammar of English Verb Systems*. Berlín: Cornelsen.
- Ministerio de Industria, Turismo y Comercio. 2000. *Plan Integral de Calidad del Turismo Español. PICTE 2000-2006*. Madrid.

- . 2006. *Movimientos turísticos en Fronteras (Frontur)* [en línea]. Madrid: Instituto de Estudios Turísticos.
<<http://www.iet.tourspain.es/Informes/Documentacion/frontur/Movimientosturistico senfronteras2005.pdf>> [Consulta: 13/04/2011].
- . 2007a. *El turismo Español en cifras 2006* [en línea]. Madrid: Instituto de Estudios Turísticos.
<<http://www.iet.tourspain.es/informes/documentacion/publicaciones/TurisCifras2006Esp.pdf>> [Consulta: 13/04/2011].
- . 2007b. *Plan del Turismo Español. Horizonte 2020* [en línea]. Madrid.
<<http://www.turismo2020.es/index.php?/esp/documentacion>> [Consulta: 13/04/2011].
- . 2009. *Nota de coyuntura. Diciembre de 2009* [en línea]. Madrid: Instituto de Estudios Turísticos.
<<http://www.iet.tourspain.es/informes/documentacion/frontur/NotasMensuales/2009/Coyuntura1209.pdf>> [Consulta: 13/04/2011].
- . 2010a. *Balance del turismo. Resultados de la actividad turística en España. Año 2009* [en línea]. Madrid: Instituto de Estudios Turísticos.
<<http://www.iet.tourspain.es/informes/documentacion/FronturFamiliar/Balance%20del%20turismo%20en%20Espana%20en%202009.pdf>> [Consulta: 13/04/2011].
- . 2010b. *Informe anual 2009* [en línea]. Madrid: Instituto de Estudios Turísticos.
<<http://www.iet.tourspain.es/informes/documentacion/egatur/Frontur-Egatur2009.pdf>> [Consulta: 13/04/2011].
- . 2010c. *Nota de coyuntura. Diciembre 2010* [en línea]. Madrid: Instituto de Estudios Turísticos.
<<http://www.iet.tourspain.es/informes/documentacion/frontur/NotasMensuales/2010/Coyuntura1210.pdf>> [Consulta: 13/04/2011].
- . 2011. *Informe turístico sobre el año 2010* [en línea]. Madrid: Ministerio de Industria, Turismo y Comercio. <http://www.mityc.es/es-ES/IndicadoresyEstadisticas/DatosEstadisticos/VIII.%20Turismo/VIII_2.pdf> [Consulta: 13/04/2011].
- Minsky, M. 1975. «A Framework for Representing Knowledge», en P.H. Winston (ed.). *The Psychology of Computer Vision*. Nueva York. 211-277.
- Mira Rueda, C. 2008. *El discurso turístico en inglés y en español: su tratamiento lexicográfico*. Tesis doctoral. Universidad de Málaga.
- Mitkov, R. (ed.) 2003. *The Oxford Handbook of Computational Linguistics*. Oxford: Oxford University Press.
- ; Evans, R.; Orasan, C.; Ha, L. A. y Pekar, V. 2007. «Anaphora Resolution: To What Extent Does It Help NLP Applications?», en A. Branco (ed.). *Anaphora: Analysis, Algorithms and Applications*. Berlín: Springer-Verlag. 179-190.

- ; Pekar, V.; Blagoev, D. y Mulloni, A. 2008. «Methods for extracting and classifying pairs of cognates and false friends». *Machine Translation*, 21 (1). 29-53.
- Mizoguchi, R., Vanwelkenhuysen, J. e Ikeda, M. 1995. «Toward Very Large Knowledge Bases: Knowledge Building and Knowledge Sharing». *IOS Press*. 46-59.
- Montaner Montejano, J. y Antich Corgos, J. 1998. *Diccionario de Turismo*. Barcelona: Editorial Síntesis.
- Monterde Rey, A. M. 2003. «Creación de un programa informático generador de sistemas de conceptos como apoyo a la enseñanza de la terminología», en N. Gallardo San Salvador (dir.). *Terminología y traducción: un bosquejo de su evolución*. Granada: Atrio. 91-106.
- . 2006. «Genesis: Programa informático dirigido a estudiantes de terminología para la generación de sistemas de conceptos», en H. Pitch (ed.). *Modern Approaches to terminological theories and applications*. Berna: Peter Lang, vol 36 [*Linguistic Insights: Studies in Language and Communication*]. 301-327.
- Montiel Ponsoda, E. 2011. *Multilingual lexico-syntactic patterns for ontology modeling and linguistic information repository for ontology localization*. Tesis doctoral. Universidad Politécnica de Madrid.
- Moon, R. 1998. *Fixed Expressions and Idioms in English: A corpus-based approach*. Oxford: Clarendon Press.
- Moreno Ortiz, A. 2000. «Diseño e Implementación de un Lexicón Computacional para Lexicografía y Traducción Automática» [en línea]. *Estudios de Lingüística Española (ELiEs)*, 9. <<http://elies.rediris.es/elies9/index.htm>> [Consulta: 13/04/2011].
- . 2004. «Representación de la información terminológica en Ontoterm®: Un sistema gestor de bases de datos terminológicas basado en el conocimiento», en P. Faber Benítez y C. Jiménez (eds.). 25-70.
- . 2008. «Ontologías para la terminología: por qué, cuándo, cómo» [en línea]. *Revista Tradumática*. <<http://www.fti.uab.es/tradumatica/revista/num6/articles/03/03central.htm>> [Consulta: 13/04/2011].
- Mott, B. L. 1996. *A Course in Semantics and Translation for Spanish learners of English*. Barcelona: EUB.
- Mucchielli, R. 1993. *Le questionnaire dans l'enquête psycho-sociale. Connaissance du problème*. París: Les Éditions ESF.
- Murphy, J. y Medin, D. L. 1985. «The role of theories in conceptual coherence». *Psychological Review*, 92 (3). 289-300.
- Murphy, J. y Scharl, A. 2007. «An investigation of global versus local online branding». *International Marketing Review*, 24 (3). 297-312.

- Nakamoto, K. 1994. *Establishing Criteria for Dictionary Criticism: A Checklist for Reviewers of Monolingual English Learners' Dictionaries*. Tesis doctoral. Universidad de Exeter.
- Navarro Galindo, J. L. y Samos Jiménez, J. 2007. «Una panorámica actual de software para trabajar con ontologías» [en línea]. *Actas del I Simposio en Desarrollo Software*. Departamento de Lenguajes y Sistemas Informáticos. Universidad de Granada. 101-116. <<http://www.salmer.info/docu/sds07.pdf>> [Consulta: 13/04/2011].
- Neches, R.; Fikes, R.E.; Finin, T.; Gruber, T.R.; Senator, T. y Swartout, W. R. 1991. «Enabling technology for knowledge sharing». *AI Magazine*, 12 (3). 36-56.
- Neubert, A. 2000. «Competence in Language, in Languages, and in Translation», en C. Schäffner y B. Adab (eds.). *Developing Translation Competence*. Ámsterdam y Filadelfia: John Benjamins. 3-18.
- Newmark, P. 1981. *Approaches to Translation*. Oxford/Nueva York/Toronto/Sydney/Fránkfort: Pergamon Press.
- . 1988. *A Textbook of Translation*. Ámsterdam: Rodopi.
- Ndi-Kimbi, A. 1994. «Guidelines for terminological definitions: The adherence to and deviation from existing rules in BS/ISO 2382: Data Processing and Information Technology Vocabulary». *Terminology*, 1 (2). 327- 350.
- Nielsen, J. 1995. *Multimedia and Hypertext: the Internet and Beyond*. Boston: Academia Press.
- Niles, I. y Pease, A. 2001. «Origins of the IEEE Standard Upper Ontology» [en línea]. *Working Notes of the IJCAI-2001 Workshop on the IEEE Standard Upper Ontology*. <<http://projects.teknowledge.com/HPKB/Publications/SUO.ps>> [Consulta: 13/04/2011].
- Nobs, M. L. 2003. *Expectativas y evaluación en traducción de folletos turísticos: estudio empírico con usuarios reales*. Tesis doctoral. Universidad de Granada.
- . 2005. *La traducción de folletos turísticos, ¿qué calidad demandan los turistas?* Granada: Comares.
- Nord, B. 2002. *Hilfsmittel beim Übersetzen. Eine empirische Studie zum Rechercheverhalten professioneller Übersetzer*. Fránkfort: Peter Lang.
- Nord, C. 1991. *Text analysis in translation. Theory, methodology, and didactic application of a Model for Translation-Oriented Text Analysis*. Ámsterdam: Rodopi.
- . 1992. «Text analysis in translator training», en C. Dollerup y A. Loddegaard (ed.), *Teaching Translation and Interpreting*. Ámsterdam: John Benjamins. 39-48.
- . 1997. *Translating as a purposeful activity*. Manchester: St. Jerome Publishing.
- . 2003. *Kommunikativ handeln auf Spanisch und Deutsch: ein übersetzungsorientierter funktionaler Sprach- und Stilvergleich*. Wilhelmsfeld: Gottfried Egert.

- Noy, N. 2003. «Ontologies», en A. Farghaly (ed.). *Handbook for Language Engineers*. Stanford: CSLI Publications.
- y McGuinness, D. 2001. *Ontology Development 101: A Guide to Creating Your First Ontology*. Documento Técnico KSL-01-05 de Stanford Knowledge Systems Laboratory y SMI-2001-0880 de Stanford Medical Informatics.
- OCDE (Organización para la Cooperación y Desarrollo Económico). 2010. *OECD Tourism Trends and Policies 2010* [en línea]. OECD Publishing. <<http://browse.oecdbookshop.org/oecd/pdfs/browseit/8510011E.PDF>> [Consulta: 13/04/2011].
- Olivera Betrán, J. 1995. «Las actividades físicas de aventura en la naturaleza: análisis sociocultural». *Apunts: educación física i sports*, 41. 5-8.
- y Olivera Betrán, A. 1995. «Propuesta de una clasificación taxonómica de las actividades físicas de aventura en la naturaleza. Marco conceptual y análisis de los criterios elegidos». *Apuntes: Educación Física y Deportes*, 41. 108-123.
- OMT. 2002. *Tourism 2020 Vision: Europe*. Organización Mundial del Turismo, Madrid, España.
- . 2010. *Panorama del turismo internacional. Edición 2009* [en línea]. Organización Mundial del Turismo, Madrid, España. <http://unwto.org/facts/eng/pdf/highlights/UNWTO_Highlights09_sp_HR.pdf> [Consulta: 13/04/2011].
- Oostdijk, N. 1991. *Corpus Linguistics and the Automatic Analysis of English*. Ámsterdam: Rodopi.
- Oster, U. 2000. «La biunivocidad de los términos y la traducción técnica». *Proceedings of the 1st International Conference on Specialized Translation*. Barcelona: Universitat Pompeu Fabra.
- . 2005. *Las relaciones semánticas de términos polilexemáticos: estudio contrastivo alemán-español*. Fráncfort: Peter Lang.
- . 2006. «Classifying Domain-Specific Intraterm Relations: A schema-Based Approach». *Terminology*, 12 (1). 1-17.
- Otman, G. 1991. *Vocabulaire d'intelligence artificielle*. Nanterre: EC2.
- . 1996. *Les représentations sémantiques en Terminologie*. París: Masson.
- Orăsan, C., Ha, L.A., Evans, R., Hasler, L. y Mitkov, R. 2007. «Corpora for computational linguistics». *Ilha Do Desterro: A Journal of English Language, Literatures in English and Cultural Studies*. Special Issue in Corpus Linguistics, 52 (1). 65-102.
- PACTE. 2003. «Building a Translation Competence Model», en Alves, F. (ed.) *Triangulating Translation: Perspectives in process oriented research*. Ámsterdam: John Benjamins. 43-66.
- . 2005. «Investigating Translation Competence: Conceptual and Methodological Issues». *Meta*, 50 (2). 609-619.

- Palomares Perraut, R. 1999. «Evaluación de recursos de comunicación para el traductor en Internet», en M. Pinto Molina y J. A. Cordón García (eds.). *Técnicas documentales aplicadas a la traducción*. Madrid: Síntesis. 179-193.
- . 2000. *Recursos documentales para el estudio de la traducción*. Málaga: Servicio de publicaciones e intercambio científico de la Universidad de Málaga.
- y Gómez Camarero, C. 2004. *Plantilla para evaluar recursos electrónicos en línea*. Documento Técnico. [BFF2003-04616 MCYT/TI-DT-2004-4]. 1-3. <<http://turicor.com/privada/documentos/TI-DT-2004-4.pdf>> [Consulta: 13/04/2011].
- Pantel, P. y Pennacchiotti, M. 2006. «Espresso: Leveraging generic patterns for automatically harvesting semantic relations». *Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the ACL*. Sydney, Australia. 113-120.
- Parra Boyero, M. 2002. «Las actividades físicas y deportivas en la naturaleza. Todo un fenómeno sociológico». *Congreso sobre educación física, ocio y recreación. III Congreso Internacional de Educación Física*. FETE, UGT. Jerez de la Frontera (Cádiz). 175-189.
- Parra Galiano, S. 2005. *La revisión de traducciones en la Traductología: aproximación a la práctica de la revisión en el ámbito profesional mediante el estudio de casos y propuestas de investigación*. Tesis doctoral. Granada: Universidad de Granada.
- Pascau Canales, M. 1998. «El territorio aragonés como oferta de recursos naturales para las actividades deportivas en el medio natural». *Actas del XVII Seminario aragonés sobre municipio y deporte*. Gobierno de Aragón, Zaragoza.
- Patronato de Turismo de la Costa del Sol. *Observatorio Turístico de la Costa del Sol. Segundo Trimestre de 2006*. Estudio realizado por la SOPDE (Sociedad de Planificación y Desarrollo). <<http://profesional.visitacostadelsol.com/bd/ficheros.php?modulo=72>> [Consulta: 13/04/2011].
- Pavel, S. y Nolet, D. 2001. *Handbook of Terminology* [en línea]. Ottawa-Hull: Translation Bureau. <http://www.termiuplus.gc.ca/didacticiel_tutorial/english/lesson1/index_e.html> [Consulta: 13/04/2011].
- Pearson, J. 1998. *Terms in context*. Ámsterdam y Filadelfia: John Benjamins.
- Pedersen, V. H. 2000. «Translation into L2 – In practice, and in the classroom», en M. Grosman, M. Kadric, I. Kovačić, I. y M. Snell-Hornby (eds.). 109-116.
- Peñalver Torres, M. T. 2004. «El turismo activo como alternativa y complemento al modelo turístico en la región de Murcia». *Cuadernos de Turismo*, 14. 179-215.
- Pérez Hernández, C. 2002. «Explotación de los corpórea textuales informatizados para la creación de bases de datos terminológicas basadas en el conocimiento» [en línea]. *Estudios de Lingüística Española (ELiEs)*, vol. 18. <<http://elies.rediris.es/elies18/index.html>> [Consulta: 13/04/2011].

- Petruck, M. 1996. «Frame Semantics», en J. Verschueren, J. O. Östman, J. Blommaert y C. Bulcaen (ed.). *Handbook of Pragmatics*. Ámsterdam y Filadelfia: John Benjamins.
- Picht, H. y Draskau, J. 1985. *Terminology: An Introduction*. Guildford: Universidad de Surrey.
- Pierini, P. 2007. «Quality in Web Translation: An Investigation into UK and Italian Tourism Web Sites». *The Journal of Specialized Translators*, 8. 85-103.
- Pigeassou, C. 1997. «Sport et tourisme: emergence du sport dans l'offre touristique». *Les Cahiers Espaces*, 52. París: Editions Touristiques Européennes. 15-24.
- Pinto Molina, M. 1998. «Gestión de calidad en Documentación» [en línea]. *Anales de Documentación*, 1. 171-183. <<http://www.um.es/fccd/anales/ad01/ad0110.html>> [Consulta: 13/04/2011].
- Porta Zamorano, J. 2006. *Técnicas cuantitativas para la extracción de términos en un corpus* [en línea]. Curso de verano de Terminología e Ingeniería lingüística. Universitat Jaume I de Castellón de la Plana. <<http://arantxa.ii.uam.es/~jporta/uji06/uji.pdf>> [Consulta: 13/04/2011].
- Pozzi, M. 2001. «The Terminological Definition: Conflict between Theory and Practice», en F. Mayer (ed.). *LSP: Perspectives for the New Millenium*, 1. Tubinga: Gunter Narr Verlag. 272-281.
- Procter, P. (ed.) 1995. *Cambridge International Dictionary of English*. 2ª ed. Cambridge: Cambridge University Press.
- Prószéky, G. y Kis, B. 2002. «Development of a Context-Sensitive Electronic Dictionary». *Proceedings of EURALEX*. Copenhagen, Vol. I. 281–290.
- y Földes, A. 2005. «Between Understanding and Translating: A Context-Sensitive Comprehension Tool». *Archive of Control Sciences*, 15 (4). 637-644.
- Proyecto POINTER. 1996. *Proposals for an Operational Infrastructure for Terminology in Europe* [en línea]. <<http://www.mcs.surrey.ac.uk/AI/pointer/report>> [Consulta: 13/04/2011].
- Pujol, D.; M. Corrius y J. Masnou. 2006. «Print Deferred Bilingualised Dictionaries and their Implications for Effective Language Learning: A New Approach to Pedagogical Lexicography». *International Journal of Lexicography*, 19 (2).
- Pym, A. 1992. «Translation Error Analysis and the Interface with Language Teaching», en C. Dollerup y A. Loddegaard (eds.). *Teaching Translation and Interpreting. Training, Talent, and Experience*. Ámsterdam: John Benjamins. 279-290.
- Radev, D.; Fan, W.; Qi, H.; Wu, H. y Grewal, A. 2005. «Probabilistic question answering on the web» [en línea]. *Journal of the American Society for Information Science and Technology (JASIST)*, 56 (6). 571-583. <<http://filebox.vt.edu/users/wfan/paper/www/www.pdf>> [Consulta: 13/04/2011].
- Rapp, R. 1995. «Identifying Word Translations in Non-Parallel Texts». *Proceedings of the 33th Meeting of the Association for Computational Linguistics (ACL'95)*. Cambridge (EE.UU.). 320-322.

- Real Academia Española. 2008. Diccionario de la Real Academia de la lengua española [en línea]. Madrid: Espasa. <<http://www.rae.es>> [Consulta: 13/04/2011].
- Realiter (Red Panlatina de Terminología Realiter). 1996. *Principios metodológicos del trabajo terminológico* [en línea]. <<http://realiter.net/spip.php?article11>> [Consulta: 13/04/2011].
- Recorder, M. J. y Cid, P. 2004. «La documentación en la traducción especializada», en C. Gonzalo García y V. García Yebra (eds.). 73-87.
- Reiss, K. 1986. *Möglichkeiten und Grenzen der Übersetzungskritik. Kategorien und Kriterien für eine sachgerechte Beurteilung von Übersetzungen*. Múnich: Hueber.
- y Vermeer, H. J. 1991. *Grundlegung einer allgemeine Translationstheorie*. Tübinga: Max Niemeyer Verlag.
- Rey, A. 1970. «Typologie génétique des dictionnaires», *Langages*, 19. 48-68.
- Rey, A. 1976. «Presentation», en Rey, A. (ed.). *Néology en marche, série b: langues de spécialité 2*. Quebec: Éditeur officiel du Québec.
- . 1979. *La terminologie: noms et notions. Que sais-je?* París: Presses Universitaires de France (PUF).
- . 1995. *Essays on Terminology*. Ámsterdam y Filadelfia: John Benjamins.
- Rivera Mateos, M. 2007. «La emergencia del turismo y el ocio deportivo en la naturaleza en Andalucía y España: aproximación al estudio de su demanda reciente». *Actas de las I Jornadas sobre Turismo y Sociedad en Andalucía*. Instituto de Estudios Sociales Avanzados (IESA-CSIC). Córdoba. 1-26.
- Roberts, R. P. 1984. «Compétence du nouveau diplômé en traduction». *Traduction et Qualité de Langue. Actes du Colloque Société des traducteurs du Québec/ Conseil de la langue française*. Quebec: Éditeur officiel du Québec. 172-184.
- . 1997. *Using dictionaries efficiently* [en línea]. ATA. <<http://www.dico.uottawa.ca/articles-en.htm>> [Consulta: 13/04/2011].
- Roche, C. 2003. «Ontology: A Survey» [en línea]. *8th Symposium on Automated Systems Based on Human Skill and Knowledge*, IFAC. 22-24 de septiembre. Göteborg (Suecia). 1-6. <<http://ontology.univ-savoie.fr/condillac/files/docs/articles/2003-ifac-ontology-a-survey.pdf>> [Consulta: 13/04/2011].
- . 2007. «Le terme et le concept: fondements d'une ontoterminologie» [en línea]. *TOTh 2007 «Terminologie & Ontologie: Théories et Applications»*. Annecy (Francia). <<http://ontology.univ-savoie.fr/condillac/files/docs/articles/2007-toth-terme-et-concept-fondements-pour-une-ontoterminologie.pdf>> [Consulta: 13/04/2011].
- ; Calberg-Challot, M.; Damas, L. y Rouard, P. 2009. «Ontoterminology: A new paradigm for terminology» [en línea]. *Actas de International Conference on Knowledge Engineering and Ontology Development, KEOD 2009*. Madeira. <<http://ontology.univ-savoie.fr/condillac/files/docs/articles/Ontoterminology-a-new-paradigm-for-terminology.pdf>> [Consulta: 13/04/2011].

- Rogers, M. y Ahmad, D. 1994. «Computerized Terminology for Translators: The Role of Text», en M. Brekke, O. Andersen, T. Dhal y J. Mykings (eds.). *Applications and Implications of Current LSP Research*, 2. Noruega: Fagbokforlaget. 840-851.
- Rondeau, G. 1983. *Introduction à la terminologie*. Chicoutimi (Quebec): Gaëtan Morin.
- Rosch, E. 1973. «On the Internal Structure of Perceptual and Semantic Categories», en T. E. Moore (ed.). *Cognitive Development and the Acquisition of Language*. Nueva York: Academic Press. 111-144.
- . 1975. «Cognitive Representations of Semantic Categories». *Journal of Experimental Psychology*, 104. 192-233.
- . 1978. «Principies of Categorization», en E. Rosch y B. Lloyd (eds.). *Cognition and Categorization*. Hillsdale: Laurence Erlbaum Ass. 27-48.
- Rotta, M. 2002. «Gli scenari della ricerca sul Web», en P. Capitán (coord.). *Semantica e terminologia nei portali. Tavola rotonda*. 19-26. <http://eprints.unifi.it/archive/00000155/00/Semantica_e_terminologia_nei_portali.pdf> [Consulta: 13/04/2011].
- Rufin, R. 2002. *Las empresas turísticas en la sociedad de la información*. Madrid: Centro de Estudios Ramón Areces, D. L.
- Ruppenhofer, J.; Ellsworth, M.; Petruck, M.; Johnson, C. y Scheffczyk, J. 2006. *FrameNet II: Extended Theory and Practice* [en línea]. <<http://framenet.icsi.berkeley.edu/book/book.html>> [Consulta: 13/04/2011].
- Sáez Padilla, J. 2005. «Las actividades en el medio natural, aproximación terminológica» [en línea]. *Wanceulen: Educación Física Digital*, 1. <<http://www.wanceulen.com/revista/nos.anteriores/numero1.diciembre05/articulos/articulo%201-5.htm?phpMyAdmin=6e797b98330a177d30bc4cb1260dadfe>> [Consulta: 13/04/2011].
- Sager, J. C. 1989. «Quality and standards – the evaluation of translations», en C. Picker (ed.). *The translator's handbook*. Londres: Aslib. 91-102.
- . 1990. *A practical course in terminology processing*. Ámsterdam y Filadelfia: John Benjamins.
- . 1992. «Future Developments and Research in Phraseology and Terminology related to Translation». *Terminologie et Traduction*, 2 (3). 583-588.
- . 1993. *Language Engineering and Translation: Consequences of automation*. Ámsterdam y Filadelfia: John Benjamins. 21-49.
- . 1994. «What's wrong with "terminology work" and "terminology science?». *Terminology*, 1 (2). 375-381.
- ; Dungworth, D. y McDonald, P. F. 1980. *English special languages: principles and practice in science and technology*. Wiesbaden: Brandstetter.
- y Ndi-Kimbi, A. 1995. «The conceptual structure of terminological definitions and their linguistic realisations: A report on research in progress». *Terminology*, 2 (1). 61- 85.

- Sánchez, A.; Sarmiento, R.; Cantos, P. y Simón, J. 1995. *Cumbre. Corpus lingüístico del español contemporáneo. Fundamentos, metodología y aplicaciones*. Madrid: SGEL.
- Sánchez, A. 2001. «Plan General del deporte en Andalucía: el deporte para todos». *VI Encuentro andaluz de deporte municipal*. Granada: Instituto Andaluz del Deporte.
- Sánchez Ramos, M. M. 2005a. «Nuevos retos para la competencia instrumental: la lexicografía electrónica», en D. Sales Salvador (ed.). *La Biblioteca de Babel: documentarse para traducir*. Albolote: Comares. 201-219.
- . 2005b. «Research on Dictionary Use by Trainee Translators. Translator Journal». *Translation Journal*, 9 (2). <<http://accurapid.com/journal/32dictuse.htm>> [Consulta: 13/04/2011].
- Sarmiento, L.; Maia, B.; Santos, D.; Pinto, A. y Cabral, L. 2006. «Corpógrafo V3 - From terminological aid to semi-automatic knowledge engineering». *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC 2006)*. Genoa, Italia. 1502-1505.
- Sayadi, S. y Calatrava, J. 2001. «Agroturismo y desarrollo rural: Situación actual, potencial y estrategias en zonas de montaña del sureste español». *Cuadernos de Turismo*, 7. 131-157.
- Schroeder, R. 2006. «How can I tell if it is a scholarly source? Evaluating Web Sites». *The Virtual Earned Scholarly Average* [en línea]. Waverly: Universidad de Wartburg. <<http://www.wartburg.edu/library/vesa.html>> [Consulta: 13/04/2011]
- Schudson, M. 1979. «Review essay: on tourism and modern culture». *American Journal of Sociology*, 84 (5). 1249-1258.
- Shank, R. y Abelson, R. P. 1977. *Scripts, plans, goals and understanding*. Hillsdale: Lawrence Erlbaum Associates.
- Secretaría de Estado de Turismo y Comercio. 2005a. *El turismo de naturaleza en España y su plan de impulso* [en línea]. Madrid: Ministerio de Industria, Turismo y Comercio. Subdirección General de Calidad e Innovación Turística. <<http://www.proyectos.com/public/resources/re/tu/090308-Plan%20impulso%20turismo%20naturaleza.pdf>> [Consulta: 13/04/2011].
- . 2005b. *La Guía de Turismo Activo - Fitur Active 2006*. Madrid: Ministerio de Industria, Turismo y Comercio.
- Seghiri Domínguez, M. 2006. *Compilación de un corpus trilingüe de seguros turísticos (español-inglés-italiano): aspectos de evaluación, catalogación, diseño y representatividad*. Tesis Doctoral. Málaga: Servicio de Publicaciones de la Universidad de Málaga.
- . 2008. «Creating virtual corpora step-by-step», en P. Sánchez Hernández, P. Pérez-Paredes, P. Aguado Jiménez y R. Criado Sánchez (eds.). *Researching and teaching specialized languages: New contexts, new challenges. Proceedings of the AELFE 2008 Conference*. Murcia: Editum. 435-449.

- Senso, J. A.; Magaña Redondo, P. J.; Faber Benítez, P. y Vila Miranda, A. 2007. «Metodología para la estructuración del conocimiento de una disciplina: el caso de PuertoTerm». *El profesional de la información*, 16 (6). 591-604.
- Sierra, G., Alarcón, R., Aguilar, C. y Bach, C. 2008. «Definitional verbal patterns for semantic relation extraction». *Terminology*, 14 (1). 78-94.
- Simard, M.; Foster, G. y Isabelle, P. 1992. «Using Cognates to Align Sentences in Bilingual Corpora» [en línea]. Proceedings of the 4th Conference on Theoretical and Methodological Issues in Machine Translation (TMI). Montreal. 67-81. <<http://www.mt-archive.info/TMI-1992-Simard.pdf>> [Consulta: 13/04/2011].
- Sinclair, J. 1987. *Collins COBUILD English Language Dictionary*. Londres y Glasgow: Collins.
- . 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- . 1993. «Lexicographers' Needs». *Zeitschrift für Anglistik and Amerikanistik*, 41 (1). 5-14.
- . 2004. «Corpus and Text: Basic Principles» [en línea], en Wynne, M. (ed.) *Developing Linguistic Corpora: a Guide to Good Practice* [en línea]. Oxford: Oxbow Books. <<http://ahds.ac.uk/creating/guides/linguistic-corpora/chapter1.htm>> [Consulta: 13/04/2011].
- Skuce, D. y Lethbridge, T.C. 1995. «CODE4: A Unified System for Managing Conceptual Knowledge». *International Journal of Human-Computer Studies*, 42. 413-451.
- Slano, F. y Velardi, P. 2007. «TermExtractor: a Web Application to Learn the Shared Terminology of Emergent Web Communities», en R. J. Gonçalves, J. P. Müller, K. Mertins y M. Zelm. *Enterprise Interoperability II: New Challenges and Approaches*. Londres: Springer.
- Smith, E. y Medin, D. L. 1981. *Categories and Concepts*. Cambridge: Harvard University Press.
- Soler Puertes, V. y Alcina Caudet, A. 2008. «Patrones léxicos para la extracción de conceptos vinculados por la relación parte-todo en español». *Terminology*, 14 (1). 99-123.
- Sosinski, M. 2006. *Fraseología comparada del polaco y del español: su tratamiento en los diccionarios bilingües*. Tesis doctoral. Universidad de Granada.
- Spohr, D. 2010. «Towards a Multifunctional Electronic Dictionary Using a Metamodel of User Needs», en S. Granger y M. Paquot (eds.). *eLexicography in the 21st century: New Challenges, new applications. Proceedings of ELEX2009. Cahiers du Cental*, 7. Lovaina-La-Nueva. Presses Universitaires de Louvain. 195-205.
- Steiner, R. J. 1984. «Guidelines for reviewers of bilingual dictionaries». *Dictionaries. Journal of the Dictionary Society of North America*, 6. 166-181.
- Strehlow, R. 1997. «Frames and the Display of Definitions», en S. E. Wright y G. Budin (eds.). 203-208.

- Suau Jiménez, F. 2006. *El metadiscurso en el género 'Promoción de Servicios y Productos Turísticos' en inglés y español: importancia de su traducción como recurso para la persuasión del cliente* [en línea]. <<http://www.uv.es/suau/pdf/BAIRES.pdf>> [Consulta: 13/04/2011]
- Summers, D. (ed.). 1995. *Longman Dictionary of Contemporary English*. 3ª ed. Harlow: Longman.
- Sure, Y. y Studer, R. 2003. «A methodology for Ontology-based Knowledge Management», en J. Davies, D. Fensel y F. Van Hamelen (eds.) *Towards the Semantic Web. Ontology-Driven Knowledge Management*. Nueva York: John Wiley & Sons. 33-46.
- Tarp, S. 2008a. *Lexicography in the Borderland between Knowledge and Non-knowledge. General Lexicographical Theory with Particular Focus on Learner's Lexicography*. Tubinga: Max Niemeyer Verlag.
- . 2008b. «The Third Leg of Two-legged Lexicography». *Hermes – Journal of Language and Communication Studies*, 40. 117-131.
- . 2009. «Reflexions on lexicographical user research». *Lexikos*, 19. 275–296.
- Tebé, C. 2003. «Terminología: historia y organización». *Curso de postgrado online de Introducción a la terminología*. Barcelona: Institut Universitari de Lingüística Aplicada. Universitat Pompeu Fabra.
- Temmerman, R. 1995. «The process of revitalisation of old words: Splicing, a case study in the extension of reference». *Terminology*, 2 (1). 107-128.
- . 1997. «Questioning the univocity ideal. The difference between socio-cognitive Terminology and traditional Terminology». *Hermes. Journal of Linguistics*, 18. 51-91.
- . 2000. *Towards New Ways of Terminology Description. The sociocognitive approach*. Ámsterdam: John Benjamins.
- . 2003. «Innovative methods in specialised terminology». *Terminology*, 9 (1). 117-135.
- y Geentjens, S. 2006. «Termonography and domain-specific translation dictionaries. Tackling cross-linguistic and cross-cultural variation». *International Conference on Terminology: Terminology and Society*. 15-17 noviembre. Amberes (Bélgica).
- y Kerremans, K. 2003. «Termonography: Ontology Building and the Sociocognitive Approach to Terminology Description» [en línea], en Hajicová, E., Kotešovcová, A., Mírovský, J. (eds.), *Proceedings of CIL17*, Matfyzpress, MFF UK (CD-ROM). Praga. <http://www.hf.uib.no/forskingskole/temmerman_art_prague03.pdf> [Consulta: 13/04/2011]
- y Knops, U. 2004. *The Translation of Domain Specific Languages and Multilingual Terminology Management. Linguistica Antverpiensia New Series 3/2004*. Amberes: Hogeschool Antwerpen.
- Termcat. 1990. *Metodologia del treball terminològic*. Barcelona: Generalitat de Catalunya. Departamento de Cultura.

- Terziev I., Kiryakov A. y Manov D. 2003. *PROTON Ontology (PROTo ONtology)* [en línea]. Proyecto SEKT. <http://proton.semanticweb.org/D1_8_1.pdf> [Consulta: 13/04/2011].
- Teubert, W. 2000. «A province of a federal superstate, ruled by an unelected bureaucracy: Keywords of the Eurosceptic discourse in Britain», en A. Mussolff, C. Good, P. Points y R. Wittlinger (eds.). *Attitudes Towards Europe: Language in the Unification Process*. Aldershot: Ashgate. 45-86.
- . 2005. «My version of Corpus Linguistics». *International Journal of Corpus Linguistics*, 11 (1). 1-13.
- y Krishnamurthy, R. (ed.) 2007. *Corpus Linguistics. Critical Concepts in Linguistics*, vol. 1. Londres y Nueva York: Routledge.
- Tiedemann, J. 2003. *Recycling Translations -Extraction of Lexical Data from Parallel Corpora and their Application in Natural Language Processing*. Tesis doctoral. Universidad de Upsala.
- Toledo Báez, C. 2009. *El resumen automático como recurso documental para la traducción de artículos de investigación del ámbito jurídico-tecnológico (español-inglés-francés)*. Tesis doctoral. Málaga: Servicios de Publicaciones de la Universidad de Málaga.
- Torres Pombert, A. 2003. «El uso de los buscadores en Internet» [en línea]. *ACIMED*, 11 (3). <<http://eprints.rclis.org/1772/1/uso.pdf>> [Consulta: 13/04/2011].
- Torruella, J. y Llisterri, J. 1999. «Diseño de corpus textuales y orales» [en línea], en J. M. Bleca, G. Clavería, C. Sánchez y J. Torruella (eds.). *Filología e informática: Nuevas tecnologías en los estudios filológicos*. Barcelona: Seminario de Filología e Informática de la Universidad Autónoma de Barcelona y Ed. Milenio. 45-77. <http://liceu.uab.es/~joaquim/publicacions/Torruella_Llisterri_99.pdf> [Consulta: 13/04/2011].
- UNE 50-106-90. 1990. *Directrices para el establecimiento y desarrollo de tesauros monolingües: equivalente a ISO 2788-1986*. Madrid: AENOR.
- Vallez, M. y Pedraza-Jiménez, R. 2007. *El Procesamiento del Lenguaje Natural en la recuperación de Información Textual y áreas afines* [en línea]. <http://www.hipertext.net/web/pag277_print.htm> [Consulta: 13/04/2011].
- Valls, J. F. 1996. *Las claves del mercado turístico: cómo competir en el nuevo entorno*. Bilbao: Deusto.
- Van Heist, G.; Schreiber, A. y Wielinga, B. 1997. «Using Explicit Ontologies in KBS Development». *International Journal of Human-Computer Studies*, 45. 183-292.
- Varantola, K. 1984. *On Noun Phrases in Engineering English*. Turku: Turku Yliopisto.
- . 1994. «The Dictionary User as Decision-Maker». *Euralex 1994. Proceedings*. Ámsterdam: Vrije Universiteit. 606-611.
- . 2002. «Disposable corpora as intelligent tools in translation» [en línea]. *Cadernos de Tradução: Corpora e Tradução*, 1 (9). Florianópolis: NUT. 171-189. <<http://www.cadernos.ufsc.br/online/9/krista.htm>> [Consulta: 13/04/2011].

- Vargas Sierra, C. 2006a. «Diseño de un corpus especializado con fines terminográficos: el Corpus de la Piedra Natural» [en línea]. *Debate Terminológico*, 2 (7). París: RITERM (Red Iberoamericana de Terminología). <http://rua.ua.es/dspace/bitstream/10045/9426/1/chelo_vargas_sierra.pdf> [Consulta: 13/04/2011].
- . 2006b. «El proceso terminográfico multilingüe con WordSmith Tools», *CONFLUÊNCIAS –Revista de Tradução Científica e Técnica*. 84–107.
- . 2007. «El léxico especializado y las ontologías», en E. Alcaraz, J. Mateo y F. Yus (coords.). *Las lenguas profesionales y académicas*. Barcelona: Ariel. 41-52.
- . 2008. «La sistematización terminográfica: una propuesta metodológica para la elaboración de diccionarios traductológicos» [en línea]. *Actas del X Simposio Iberoamericano de Terminología*. Montevideo (Uruguay). <<http://rua.ua.es/dspace/bitstream/10045/13212/1/1453%20Vargas.pdf>> [Consulta: 13/04/2011].
- . 2009. «Do-it-yourself IT for Terminology o experiencias de bricolaje informático en la elaboración de diccionarios terminológicos». *Punto y Coma*, 115 (Suplemento especial de las VIII Jornadas de AETER). 42-49.
- Vázquez, J. C. 2002. «Datos para un análisis del mercado laboral en turismo deportivo», en M. Latiesa, S. Rebollo y J.L. Paniza (eds.). *Salidas profesionales en el campo del turismo deportivo*. Málaga: Instituto Andaluz del deporte, Junta de Andalucía. 37-49.
- Vázquez de Francisco, A. 2006. «Los compuestos nominales en el texto publicitario turístico en Internet: una aplicación didáctica», en C. P. Neumann, R. P. Alastrué y C. Pérez-Llantada Auría (eds.). *Actas de V Congreso Internacional AELFE*. 347-353.
- Véronis, J. (ed.). 2000. *Parallel Text Processing: Alignment and Use of Translation Corpora*. Dordrecht: Kluwer.
- Vermeer, H. 1989. «Wörterbücher als Hilfsmittel für unterschiedliche Typen der Translation», en F. J. Hausman, O. Reichmann, H. E. Wiegand y L. Zgusta. (eds.) *Wörterbücher, Dictionaries, Dictionnaires. An International Encyclopaedia of Lexicography*, 1^{er} vol. Berlín y Nueva York: de Gruyter. 171-174.
- Videon, C. 2006. *WWW Checklist* [en línea]. Delaware County (PA): Community College. <<http://faculty.dccc.edu/~cvideon/wwweval.htm>> [Consulta: 13/04/2011].
- Villayandre Llamazares, M. 2006. «La lingüística del corpus como metodología lingüística» [en línea], en M. Villayandre Llamazares (ed.). *Lingüística computacional II: Lingüística del corpus*. León: Universidad de León. 1-17. <http://www3.unileon.es/dp/dfh/Milka/LCII/Ling_corpus.pdf> [Consulta: 13/04/2011].
- Vintar, Š. 2001. «Using Parallel Corpora for Translation-oriented Term Extraction». *Babel*, 47(2): 121-132.
- Vivaldi Palatresi, J. 2003. *Sistema de extracción de candidatos a término YATE: manual de utilización* [en línea]. Barcelona: Institut Universitari de Lingüística

- Aplicada. <<http://www.iula.upf.edu/repositori/06inf043.pdf>> [Consulta: 13/04/2011].
- Voutilainen, A. 2003. «Part-of-Speech Tagging», en R. Mitkov (ed.). 219-232.
- Waddington, C. 2000. *Estudio comparativo de diferentes métodos de evaluación de traducción general*. Madrid: Universidad Pontificia Comillas.
- Weaver, D. (ed.). 2001a. *The Encyclopedia of Ecotourism*. Nueva York: CABI Publishing.
- . 2001b. «Ecotourism in the Context of Other Tourism Types», en D. Weaver (ed.). 73-83.
- y Lawton, L. 2006. *Tourism Management*. Brisbane: John Wiley and Sons Australia, Ltd.
- Welker, H. A. 2008. *Panorama general da lexicografia pedagógica*. Brasília: Thesaurus.
- . 2010. *Dictionary Use. A General Survey of Empirical Studies*. Brasília: Thesaurus.
- White, J. B. 1991. «Our meanings can never be the same: reflections on language and law». *Rhetoric Society Quarterly*, 21(3). 68-77.
- Wiegand, H. 1997. «A Multilingual Ontology-based Lexicon for News Filtering –The TREVI Project». *IJCAI-97 Workshop WP24. Ontologies and Multilingual NLP*. Japón.
- . 1998. *Wörterbuchforschung. Untersuchungen zur Wörterbuchbenutzung, zur Theorie, Geschichte, Kritik und Automatisierung der Lexikographie*. Vol. 1. Berlin: W. de Gruyter.
- Wight, P.A. 2001. «Ecotourists: Not a Homogeneous Market Segment», en D. Weaver (ed.). 37-62.
- Wilss, W. 1976. «Perspectives and limitations of a didactic framework for the teaching of translation», en R.W. Brislin (ed.). *Translation Applications and Research*, Nueva York: Gardner. 117-137.
- . 1996. *Knowledge and Skills in Translation Behavior*. Ámsterdam: John Benjamins.
- Wotjak, B. 2005. «Formas rutinarias en los diccionarios didácticos», en J. D. Luque Durán y A. Palies Bertrán (ed.). *La creatividad del lenguaje: colocaciones idiomáticas y fraseología*. Granada: Granada Lingüística (Serie Collectae).
- Wright, S. E. 1993. «Term entry formatting conventions». *Keystones of Communications: Proceedings of the 34th Annual Conference of the American Translators Association*. Medford: Learned Information. 179-196.
- . 1997. «Representation of Concept Systems», en S. E. Wright y G. Budin (eds.). Vol. 1. 89-97.
- . 2001. «Data Categories for Terminology Management», en S. E. Wright y G. Budin (eds.). Vol. 2. 552-571.

-
- y Budin, G. 1997. «The Handbook of Terminology Management: Introduction», en S. E. Wright y G. Budin (eds.). Vol. 1. 1-10.
- y Budin, G. (eds.). 1997. *Handbook of Terminology Management: Basic Aspects of Terminology Management*. Vol. 1. Ámsterdam y Filadelfia: John Benjamins.
- y Budin, G. (eds.). 2001. *Handbook of Terminology Management: Application-Oriented Terminology Management*. Vol. 2. Ámsterdam y Filadelfia: John Benjamins.
- y Wright, L. 1997. «Terminology Management for Technical Translations», en S. E. Wright y G. Budin. 147-159.
- Wüster, E. 1998. *Introducción a la teoría general de la terminología y a la lexicografía terminológica*. Barcelona: Institut Universitari de Linguística Aplicada [traducción].
- Zanettin, F. 1998. «Bilingual Comparable Corpora and the training of Translators». *Meta*, 43 (4). 1-4.
- . 2002. «DIY Corpora: The WWW and the Translator» [en línea], en Maia, B., Haller, J. y Urlrych, M. (eds.) *Training the Language Services Provider for the New Millennium*. Porto: Faculdade de Letras, Universidade do Porto. 239-248. <<http://www.federicozanettin.net/DIYcorpora.htm>> [Consulta: 13/04/2011].
- Željko, M. 2009. «Improvements of Dictionaries – Suggestions by Evroterm» [en línea]. *Proceedings of the INFUTURE2009: Digital Resources and Knowledge Sharing*. 269-278. <<http://infuture.ffzg.hr/papers/4-08%20Zeljko,%20Improvements%20of%20Dictionaries%20-%20Suggestions%20by%20Evroterm.pdf>> [Consulta: 13/04/2011].
- Zgusta, L. 1971. *Manual of Lexicography*. Praga: Academia.
- Zielinski, D. y Ramírez Safar, Y. 2005. «Research meets practice: t-survey 2005 - An online survey on terminology extraction and terminology management» [en línea]. *Proceedings ASLIB 27 - Translating and the Computer*. 24-25 Noviembre. Londres. 1-30. <http://ecolotrain.uni-saarland.de/download/sdv/malaga2007/articles/t-survey_aslib2005_zielinski.pdf> [Consulta: 13/04/2011].

ANEXOS

ANEXO I

A continuación, se muestra la encuesta diseñada en esta investigación para conocer las necesidades y requisitos de los traductores profesionales. Esta encuesta estuvo disponible durante los meses de junio a septiembre de 2008 en línea en la siguiente dirección: <<http://clg.wlv.ac.uk/surveys/survey.php?sid=29>>, en su versión inglesa, pero también se distribuyó a través de mensajes electrónicos en versiones en español, italiano y alemán. La versión que se incluye en este apartado equivale a la encuesta en español.

FUENTES TERMINOLÓGICAS PARA TRADUCTORES: ENCUESTA

Esta encuesta está orientada a conocer su opinión acerca de los recursos terminológicos disponibles en la actualidad con relación a sus necesidades y requisitos personales. Asimismo, algunas preguntas pretenden conocer su metodología de trabajo por cuanto respecta al uso de recursos durante la traducción.

DATOS SOBRE PROFESIÓN

1. Profesión

| | |
|---------------------|--------------|
| Gestor de proyectos | Subtitulador |
| Traductor | Terminólogo |
| Intérprete | |

2. Situación laboral

| | |
|----------------------------|-----------------------------|
| Autónomo | Empleado en una empresa |
| Propietario de una empresa | Empleado de una institución |

3. Años en la profesión

| | |
|----------|----------------|
| 0-2 años | 6-8 años |
| 2-4 años | 8-10 años |
| 4-6 años | Más de 10 años |

4. Formación relevante con su trabajo

| | |
|---|--|
| Título universitario | Otro tipo de formación |
| Miembro/certificado de un organismo profesional | Sin formación relacionada con su especialidad. |

5. Cursos específicos realizados

| | |
|--------------------------------------|-------------------------|
| Memorias de traducción | Recursos terminológicos |
| Extracción y/o gestión terminológica | Localización |
| Traducción especializada | Otros _____ |

MEDIO DE TRABAJO

6. Lenguas de trabajo

| | |
|-----------|------------|
| Inglés | Chino |
| Francés | Japonés |
| Italiano | Polaco |
| Español | Ruso |
| Alemán | Árabe |
| Griego | Otro _____ |
| Portugués | |

7. Áreas de especialización

| | |
|---|---|
| Agricultura y/o horticultura | Geografía y/o geología |
| Arquitectura y/o construcción | Humanidades |
| Arte, literatura y/o medios de comunicación | Matemáticas y/o estadística |
| Astronomía y/o cosmología | Propiedad intelectual (patentes, copyright) |
| Biología | Medicina y/o farmacia |
| Finanzas | Tecnología |
| Economía | Informática |
| Química | Física |
| Transporte | Textil y ropa |
| Cosmética y perfumería | Derecho |

Artesanía
Defensa
Ingeniería
Asuntos relacionados con la UE

Márketing
Turismo
Otro _____

9. ¿Para qué utiliza frecuentemente internet?

| | |
|---|-------------------------------|
| Para comunicarse (recepción y entrega de encargos de traducción, contacto con clientes/colegas, etc.) | Para descargar software |
| Como fuente de conocimiento (información de áreas especializadas, textos de referencia) | Para actualizar su página web |
| Como fuente de información lingüística (terminología, información gramatical, usos, ejemplos) | |

FUENTES TERMINOLÓGICAS

10. ¿Qué formato prefiere para las fuentes terminológicas?

Fuentes en línea
Fuentes sin conexión (por ej. CD-ROM)
Fuentes en papel

11. ¿Qué tipo de recursos lingüísticos prefiere con relación a la lengua?

| | |
|---------------------------------|-------------|
| Monolingüe (primera lengua, L1) | Bilingüe |
| Monolingüe (segunda lengua, L2) | Multilingüe |

12. ¿Qué utiliza más a menudo?

Fuentes terminológicas disponibles (por ej. diccionarios, bases de datos)
Recursos elaborados personalmente
Fuentes terminológicas de clientes/empresa

13. ¿Qué tipo de fuente terminológica utiliza cuando traduce?

| | |
|---|--|
| Diccionario/glosario monolingüe especializado (L1) | Búsquedas en motores de búsqueda (por ej. Google) |
| Diccionario/glosario monolingüe especializado (L2) | Tesaurus |
| Diccionario/glosario bilingüe especializado | Corpus paralelo (textos originales y sus traducciones) |
| Diccionario/glosario multilingüe especializado | Corpus comparable (textos originales en ambas lenguas) Base de datos terminológica |
| Diccionario visual monolingüe | Base de datos terminológica |
| Diccionario visual bilingüe | Enciclopedia |
| Sección Imágenes de un motor de búsqueda (por ej. Google) | Wikipedia |
| Listas de distribución | Otros |
| Foros de internet | |

14. De los recursos anteriores, ¿cuáles utiliza más a menudo?

| | |
|--|--|
| Diccionario/glosario monolingüe especializado (L1) | Búsquedas en motores de búsqueda (por ej. Google) |
| Diccionario/glosario monolingüe especializado (L2) | Tesaurus |
| Diccionario/glosario bilingüe especializado | Corpus paralelo (textos originales y sus traducciones) |
| Diccionario/glosario multilingüe especializado | Corpus comparable (textos originales en ambas lenguas) Base de datos terminológica |
| Diccionario visual monolingüe | Base de datos terminológica |

| | |
|---|--------------|
| Diccionario visual bilingüe | Enciclopedia |
| Sección Imágenes de un motor de búsqueda (por ej. Google) | Wikipedia |
| Listas de distribución | Otros |
| Foros de internet | |

15. Si utiliza, o ha utilizado Wikipedia alguna vez, ¿con qué fin la utiliza?

| | |
|--|--|
| Para encontrar información sobre un dominio. | Para encontrar equivalentes de traducción. |
| Para encontrar contexto. | Para encontrar formas desarrolladas de acrónimos/abreviaturas. |
| Para encontrar ejemplos de uso de palabras. | Otros |

16. Si no utiliza Wikipedia, ¿cuál es la razón?

Wikipedia no es una fuente fiable.
 Nunca he pensado en utilizar Wikipedia, pero la probaré.
 Nunca he pensado en utilizar Wikipedia, y no lo voy a hacer.
 Wikipedia no contiene información relevante para mí/mi trabajo.

17. ¿Compila su propio corpus (conjunto de textos electrónicos) cuando traduce?

Sí
 No

18.1. ¿Por qué compila su propio corpus?

Los recursos terminológicos de mi dominio no son fiables.
 Puedo comprobar los términos en su contexto.
 Puedo comprobar el uso real del lenguaje de un dominio.
 Puedo buscar ejemplos.
 Puedo entender mejor los significados.
 Puedo extraer los términos para hacer mi propio banco terminológico.
 Otro

18.2. ¿Por qué no utiliza corpus cuando traduce?

No entiendo lo que significa un corpus.
 Creo que es suficiente con las fuentes terminológicas disponibles.
 No tengo tiempo para compilar un corpus cuando traduzco.
 No creo que sea útil.

19. ¿Utiliza algún tipo de método de extracción terminológica cuando traduce?

| | |
|-------------------|-------------|
| Sí, siempre. | Casi nunca. |
| Sí, muy a menudo. | No |
| Sí, a veces. | |

19.1. Si su respuesta anterior ha sido “Sí”, ¿de dónde extrae los términos?

De los textos origen
 De los textos de referencia
 De memorias de traducción

19.2. ¿Cómo extrae terminología?

Manualmente
 Herramienta de extracción terminológica (por ej. Wordsmiths Tools)
 Sistema de memoria de traducción (Multiterm Extract, Multitrans, Chamblon Terminology)
 Otro

20. Si su respuesta ha sido “Casi nunca” o “No”, ¿por qué?

- El rendimiento de las herramientas de extracción terminológica no es demasiado bueno.
- El resultado de las listas de extracción terminológica contiene muchos errores en relación a los candidatos a términos.
- Las herramientas de extracción terminológica detectan muy pocos términos.
- La comprobación de la lista de términos requiere demasiado tiempo.
- El precio de las herramientas de extracción terminológica es muy elevado.
- No creo que sean útiles.
- No creo que sea necesario.
- Nunca he oído hablar de las herramientas de extracción terminológica.

EVALUACIÓN DE RECURSOS UTILIZADOS POR USTED Y SU OPINIÓN SOBRE LOS RECURSOS “IDEALES”

21. ¿Considera que existen fuentes terminológicas de calidad disponibles en su dominio de traducción especializada?

- | | |
|---|---|
| La mayoría de las fuentes son muy buenas. | No hay fuentes buenas. |
| Algunas de las fuentes son muy buenas. | Prefiero crear mis propias fuentes. |
| Es difícil encontrar buenas fuentes. | Hay fuentes de buena calidad pero el precio es demasiado elevado. |

22. ¿Evalúa las fuentes terminológicas antes de utilizarlas? (por ej. autor, actualización, comentarios realizados)

- | | |
|--|---|
| Sí, si la fuente no es fiable, no la utilizo. | No, no me importa. |
| Sí, pero no siempre es posible utilizar fuentes fiables. | No, nunca he pensado en evaluar esas fuentes. |
| Sí, alguna vez pero no siempre. | |

23. ¿De qué forma evalúa la calidad de las fuentes terminológicas? Indique del 1 al 7, siendo 1 la primera opción y 7 la última.

- | | |
|----------------------------|--------------------------------------|
| Autor de la obra/sitio web | Instrucciones de uso |
| Variedad de la información | Comentarios externos sobre la fuente |
| Actualización | Accesibilidad (tipos de búsqueda) |

24. Cuando considera que una fuente terminológica es de baja calidad, ¿cuáles son los problemas que le encuentra? (en su caso)

- | | |
|---|--|
| Falta de términos | Falta de explicaciones sobre varios equivalentes de traducción |
| Falta de otras unidades aparte de nombres (verbos, adjetivos, adverbios) | Falta de coherencia en los equivalentes de traducción. |
| Falta de definiciones | Falta de información gramatical (categoría, número, género) |
| Falta de definiciones claras | Falta de instrucciones de uso |
| Falta de información etimológica | Falta de información semántica (sinónimos, antónimos, hipernónimos, hipónimos) |
| Falta de indicaciones de pronunciación | Falta de imágenes |
| Falta de silabación | Presencia de errores |
| Falta de ejemplos | Falta de derivados y nombres compuestos |
| Falta de información fraseológica (colocaciones, expresiones idiomáticas) | Falta de especificaciones del dominio |

25. ¿Qué tipo de orden prefiere en una fuente terminológica?

- | | |
|------------------|------------|
| Orden alfabético | Ambos |
| Orden conceptual | Cualquiera |

26. ¿Qué cree que debería incluir una fuente terminológica adecuada para traductores?

| | Información necesaria | Información deseable | Información innecesaria |
|---|-----------------------|----------------------|-------------------------|
| Un mayor número de entradas | | | |
| Una mayor variedad de unidades (nombres, verbos, adverbios, adjetivos, nombres propios) | | | |
| Abreviaturas y acrónimos | | | |
| Definiciones | | | |
| Definiciones más clara | | | |
| Definiciones en ambas lenguas (si es bilingüe) | | | |
| Información etimológica | | | |
| Pronunciación | | | |
| Silabación | | | |
| Ejemplos | | | |
| Una mayor variedad de ejemplos | | | |
| Ejemplos tomados de contextos reales | | | |
| Información fraseológica (colocaciones, expresiones idiomáticas) | | | |
| Derivados y nombres compuestos | | | |
| Especificaciones de dominio | | | |
| Más equivalentes de traducción | | | |
| Una explicación de cada equivalente de traducción | | | |
| Información gramatical (parte de la oración, inflexión, número, género) | | | |
| Instrucciones de uso | | | |
| Información semántica (sinónimos, antónimos, hiperonimia, homonimia) | | | |
| Imágenes | | | |

27. ¿Tendría alguna otra sugerencia sobre el contenido que debería incluir una fuente terminológica para traductores?

ANEXO II

A continuación, se muestra la plantilla de evaluación *ad hoc* que se ha diseñado para evaluar recursos terminológicos tanto en formato electrónico como en papel.

| DESCRIPCIÓN | | Puntuación |
|--|--|------------|
| Título | | |
| Autoría (1, 2, 3) | | |
| URL (si procede) | | |
| Fecha de publicación/actualización (1, 2, 3) | | |
| Lenguas | <input type="checkbox"/> Monolingüe <input type="checkbox"/> Bilingüe <input type="checkbox"/> Multilingüe | |
| SopORTE | <input type="checkbox"/> Papel <input type="checkbox"/> Electrónico | |
| Especialización en TURISMO (1, 2, 3) | <input type="checkbox"/> Sí <input type="checkbox"/> No | |
| Puntuación total = | | |
| MACROESTRUCTURA | | Puntuación |
| Indicaciones (1, 2, 3) | <input type="checkbox"/> Sí <input type="checkbox"/> No | |
| Número de entradas (1, 2, 3) | | |
| Ausencia de errores (1, 2, 3) | <input type="checkbox"/> Sí <input type="checkbox"/> No | |
| Ergonomía (1, 2, 3) | <input type="checkbox"/> Sí <input type="checkbox"/> No | |
| Puntuación total = | | |
| MICROESTRUCTURA | | Puntuación |
| Definición (1, 2, 3) | <input type="checkbox"/> Sí <input type="checkbox"/> No | |
| Información gramatical (1, 2, 3) | <input type="checkbox"/> Sí <input type="checkbox"/> No | |
| Uso de las unidades (1, 2, 3) | <input type="checkbox"/> Sí <input type="checkbox"/> No | |
| Información sintagmática (locuciones, colocaciones y enunciados fraseológicos) (1, 2, 3) | <input type="checkbox"/> Sí <input type="checkbox"/> No | |
| Información paradigmática (sinónimos, antónimos) (1, 2, 3) | <input type="checkbox"/> Sí <input type="checkbox"/> No | |
| Ejemplos (1, 2, 3) | <input type="checkbox"/> Sí <input type="checkbox"/> No | |
| Remisiones/hiperenlaces (1, 2, 3) | <input type="checkbox"/> Sí <input type="checkbox"/> No | |
| Tipos diferentes de entradas (sustantivo, adjetivo, verbo, acrónimos, abreviaturas) (1, 2, 3) | <input type="checkbox"/> Sí <input type="checkbox"/> No | |
| Puntuación total = | | |
| Otra característica que destacar (3) | | |
| Puntuación final = | | |

ANEXO III

En las páginas que siguen encontrará un listado de Asociaciones de traductores e intérpretes profesionales que se han contactado para llevar a cabo la distribución de la encuesta realizada en esta investigación, cuya finalidad consiste en conocer las necesidades y requisitos propios de los traductores profesionales en la consulta de recursos terminológicos durante el proceso de traducción.

- AATI (Asociación Argentina de Traductores e Intérpretes). URL: <http://www.aati.org.ar>
- AATIA (Austin Area Translators and Interpreters Association). URL: <http://www.aatia.org>
- ADÜ Nord (Assoziierte Dolmetscher und Übersetzer in Norddeutschland e.V.). URL: <http://www.adue-nord.de>
- ***AIETI (Asociación Ibérica de Estudios de Traducción e Interpretación)***. URL: <http://www.aieti.es/>
- AIIC (International Association of Conference Interpreters). URL: <http://www.aiic.net>
- AITC (Association internationale des traducteurs de conférence). URL: <http://www.aitc.ch>
- AITI (Italian Translators and Interpreters Association). URL: <http://www.mix.it/AITI/AITI.htm>
- ALTA (American Literary Translators Association). URL: <http://wwwpub.utdallas.edu/research/cts/alta.htm>
- ATA (American Translators Association). URL: <http://www.atanet.org>
- ATIA (Association of Translators and Interpreters of Alberta). URL: <http://www.atia.ab.ca>
- ATICOM (Association of Professional Freelance Translators and Interpreters, Germany) . URL: <http://www.aticom.de>
- ATIDA (Arabic Translation and Intercultural Dialog Association). URL: <http://www.atida.org>
- ATINS (Association of Translators and Interpreters of Nova Scotia). URL: <http://www.chebucto.ns.ca/Commerce/ATINS/>
- ATIO (Association of Translators and Interpreters of Ontario). URL: <http://www.atio.on.ca/>
- ATPP (Asociación de Traductores Profesionales del Perú). URL: <http://www.atpp.org.pe>
- ATR Romanian Translators Association. URL: <http://www.atr.org.ro>
- AUSIT Australian Institute of Interpreters and Translators. URL: <http://www.ausit.org>
- BDÜ (German Translators and Interpreters Association). URL: <http://www.bdue.de>
- CATI (Carolina Association of Translators and Interpreters). URL: <http://www.catiweb.org>
- CBTIP-BKVTF (Belgian Chamber of Translators, Interpreters and Philologists). URL: <http://www.cbtip-bkvtf.org/>
- CTA (Colorado Translators Association). URL: <http://www.cta-web.org/base>
- FIT (Fédération internationale des traducteurs). URL: <http://fit.ml.org>
- FLATA (Florida Chapter of the ATA). URL: <http://www.atafl.com/>
- ITA (Irish Translators' Association). URL: <http://homepage.tinet.ie/~translation>

- Interpreters' Division of ATA. URL: <http://www.ata-divisions.org/ID/>
- IoL (Institute of Linguists). URL: <http://www.iol.org.uk>
- ITI (Institute of Translation and Interpretation, United Kingdom). URL: <http://www.iti.org.uk>
- LTAC/ATTLC (Literary Translators' Association of Canada). URL: <http://www.geocities.com/Athens/Oracle/9070/>
- NCATA (National Capital Area Chapter of the American Translators Association). URL: <http://www.ncata.org>
- NCTA (Northern California Translators Association). URL: <http://www.ncta.org>
- NETA (New England Translators Association). URL: <http://netaweb.org/>
- NMTIA (New Mexico Translators and Interpreters Association). URL: <http://www.cybermesa.com/~nmtia>
- NOTA (Northeast Ohio Translators Association). URL: <http://www.ohiotranslators.org/>
- NOTIS (Northwest Translators and Interpreters Society). URL: <http://www.NOTISnet.org>
- NYCT (New York Circle of Translators). URL: <http://www.nyctranslators.org/index.html>
- Österreichischer Übersetzer- und Dolmetscherverband. URL: <http://www.universitas.org/>
- OTIAC (Ordre des traducteurs et interprètes agréés du Québec). URL: <http://www.otiaq.org/>
- SCATIA (Southern California Area Translators Association). URL: <http://www.scatia.org>
- SFT (Société française des traducteurs). URL: <http://www.sft.fr>

ANEXO IV

El siguiente anexo presenta el manual de usuario de la herramienta de búsqueda y consulta diseñada en el seno de esta tesis doctoral, *Trandix*, donde se indican los pasos básicos que se deben efectuar para poder utilizarla de forma adecuada.

USAR TRANDIX PASO A PASO

Paso 1. Cargar los archivos de trabajo en *Trandix*. En primer lugar, es necesario cargar el archivo de texto en formato .txt en el campo Archivo TEXTO (1) y, en segundo lugar, seleccionar el archivo TBX (2). Una vez seleccionados los textos, elegir el idioma del texto (3) que se va a procesar y pulsar el botón *Aceptar* (4). Si no queremos usar este asistente, pulsamos *Cancelar* (5).

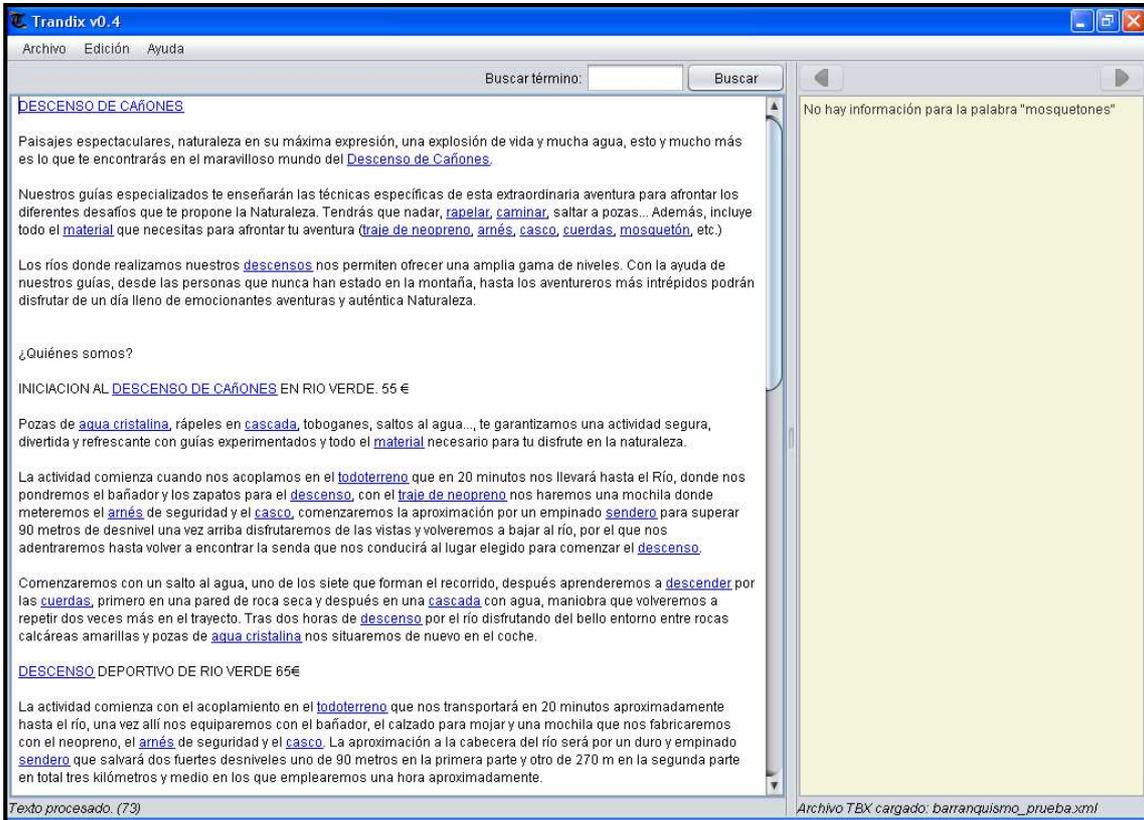


NOTA: Si se desea, estos pasos se pueden realizar directamente en la pantalla principal de *Trandix* a través del menú Archivo, una vez que se ha seleccionado *Cancelar*. Para cargar el archivo de texto, tenemos que pulsar en *Archivo* → *Abrir archivo de texto...* y elegirlo, o bien, seleccionar un archivo de la lista *Archivo* → *Abrir archivo de texto reciente*, si lo hubiera. Para cargar el archivo TBX debemos pulsar en *Archivo* → *Abrir archivo TBX...* y escogerlo, o bien, elegir un archivo de la lista *Archivo* → *Abrir archivo TBX reciente*, si lo hubiera. (*Archivo* → *Idioma del texto*), que por defecto es español, y deberemos cargar los ficheros también desde el menú Archivo.



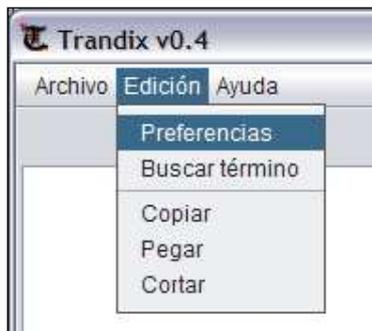
Paso 2. Una vez cargados los archivos, se inicia el procesamiento del texto automáticamente y se pasará a la ventana principal de *Trandix*. En esta ventana, se visualiza el texto de trabajo en el panel de la izquierda con los términos que se encuentran en la base de datos resaltados con hipervínculos. La

información terminológica referente a los términos se muestra en el panel de la derecha al hacer doble clic sobre el término en cuestión.

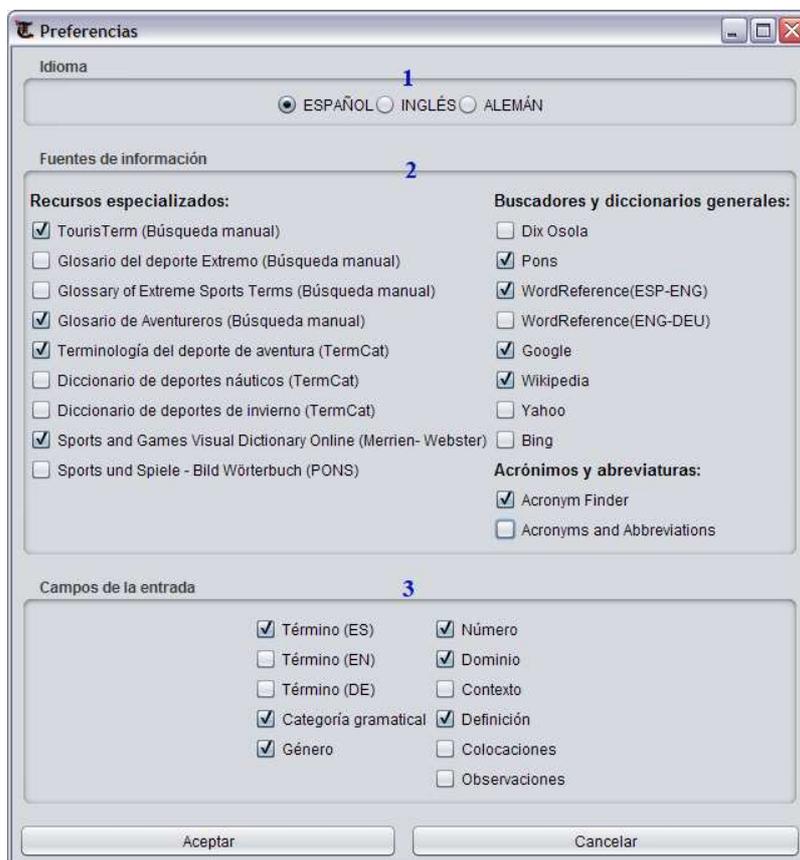


NOTA: La carga del fichero de texto no es obligatoria, ya que el panel de la izquierda es editable y se puede pegar el texto deseado directamente ahí. En este caso, tenemos que pedirle a Trandix explícitamente que procese el texto. Para ello, solo hay que pulsar *Archivo*→*Procesar*.

Paso 3. Antes de comenzar a utilizar *Trandix*, es recomendable acceder a la opción *Preferencias*, en el menú *Edición* (*Edición*→*Preferencias*) para proceder a la personalización de la herramienta por parte del usuario.

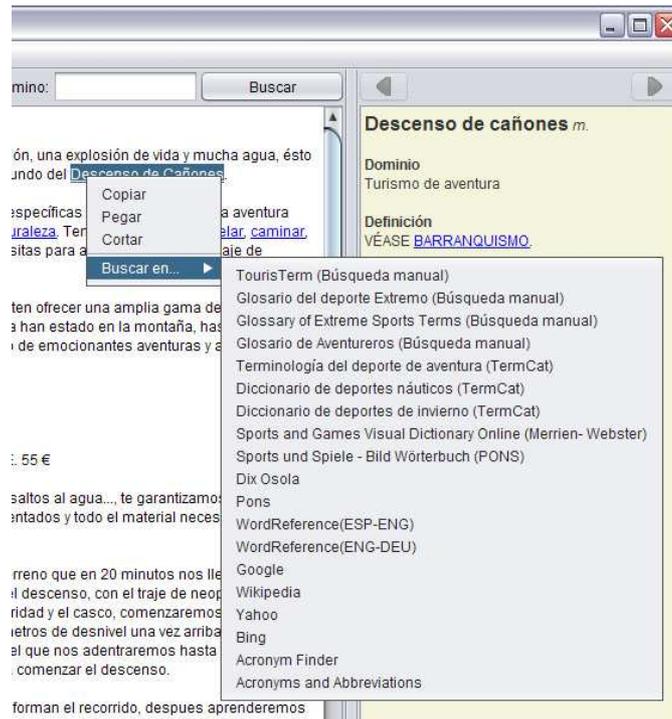


En esta opción, se podrá modificar el idioma de la interfaz (1), las fuentes de información que se desean utilizar (2) y los campos terminológicos que se quieren visualizar en la entrada terminológica. Una vez seleccionadas las opciones deseadas, se pincha en el botón *Aceptar*.



Paso 4. A continuación, nos encontramos de nuevo en la página principal de *Trandix*, donde se podrá proceder a realizar las búsquedas terminológicas en contexto, es decir, directamente en el texto de trabajo. Para realizarlas, simplemente se debe hacer doble clic en los términos que aparecen con hipervínculo. La información terminológica aparecerá en el panel derecho de la pantalla principal.

Paso 5. Para realizar búsquedas externas en las fuentes de información seleccionadas anteriormente, se debe proceder de la siguiente manera: 1. seleccionamos el término deseado, 2. Pinchamos el botón derecho del ratón sobre dicha selección y aparecerá un menú contextual, 3. elegimos *Buscar en...* y 4. Seleccionamos el recurso deseado de la lista de recursos disponibles seleccionados. Una vez que se ha elegido el recurso, se abrirá el navegador por defecto (Internet Explorer, Firefox, Safari...), que mostrará los resultados obtenidos por el recurso elegido para la búsqueda del término seleccionado.



Paso 6. Para realizar búsquedas de términos directamente a la base de datos cargada sin necesidad de usar un texto, es necesario escribir el término deseado en el campo de búsqueda que aparece en la parte superior derecha de la pantalla principal.



También es posible realizar este tipo de búsqueda a través del menú *Edición* → *Buscar término*. Con esta opción, se podrán acceder a los términos buscados más recientemente.



ANEXO V

Seguidamente se ofrece la clasificación que se ha realizado de los recursos terminológicos para el sector turístico en general y del segmento del turismo de aventura en particular según los diferentes criterios de clasificación aplicados para su catalogación, así como la evaluación de todas las fuentes.

| AVENTURA ⁶⁷⁴ | | LENGUA | FORMATO | TIPO | ORDEN | PUNTO | REFERENCIA |
|-------------------------|---|------------------------------------|----------|--------------------|-----------------------------|-------|---|
| 1 | Glosario sobre deportes de aventura | catalán, español, francés e inglés | en línea | Glosario | alfabético | 17 | http://www.termcat.cat/es/Diccionaris_En_Linia/28/Presentacio/# |
| 2 | Glosario del deporte extremo | inglés, español | en línea | Glosario | alfabético | 10 | http://www.elcastellano.org/gloxtrem.html |
| 3 | Enciclopedia de los deportes de aventura | español | en papel | Enciclopedia | alfabético | 12 | Biosca, C. 2004. <i>Enciclopedia de los deportes de aventura. La guía más completa de los deportes de aventura y riesgo</i> . Madrid: Edimat Libros. 159 páginas. ISBN: 84-9764-482-4 |
| 4 | Glosario de Aventureros | español | en línea | Glosario | alfabético | 10 | http://www.sappiens.com/sappiens%5Ccomunidades%5Cavenglosa.nsf/webtitulos!openform |
| 5 | Extreme Sports Terminology | inglés | en línea | Glosario | alfabético | 9 | http://extremesports.suite101.com/article.cfm/_what_is_index |
| 6 | Glossary of Extreme Sports Terms | inglés | en línea | Glosario | 1º temático y 2º alfabético | 13 | http://www.ixlactionsports.com/action-sports-terminology/ |
| 7 | Sports and Games - Visual Dictionary Online (Merriam-Webster) | inglés | en línea | Diccionario visual | temático | 16 | http://visual.merriam-webster.com/sports-games/cycling.php |
| 8 | Sports und Spiele - Bild Wörterbuch (PONS) | alemán, inglés | en línea | Diccionario visual | temático | 16 | http://www.bildwoerterbuch.com/en/sports-and-games/ |
| MONTAÑISMO | | LENGUA | FORMATO | TIPO | ORDEN | PUNTO | REFERENCIA |
| 9 | Glosario de Andarines.com | español | en línea | Vocabulario | alfabético | 11 | http://www.andarines.com/ |

⁶⁷⁴ En esta tabla se incluyen únicamente aquellos criterios que son relevantes para la clasificación y, por tanto, se han eliminado algunos de los criterios propuestos para la catalogación de recursos terminológicos en general. A continuación, presentamos los criterios eliminados y nuestra motivación para proceder a su eliminación: el criterio de la extensión de la especialidad, puesto que se realiza una clasificación propia en la misma tabla; el criterio de fuente de información utilizada, puesto que ninguno de los recursos incluidos indican la fuente de información, a excepción de alguno en papel, pero que no se considera relevante; el criterio de la función, ya que todos tienen una función descriptiva (frente a una función prescriptiva o normativa); el criterio de la aplicación, ya que su finalidad consiste en la comprensión del texto origen y ninguno presenta la función de codificación del texto meta; el criterio de tipo de unidades, puesto que la gran mayoría incluyen únicamente sustantivos, sin tener en cuenta otros tipos de unidades (sinónimos, adjetivos, verbos, etc.); y, por último, el criterio de usuario, puesto que ninguno especifica el grupo al que va dirigido de forma específica sino que se dirigen al público en general.

| 10 | Glosario montañero | español | en línea | Vocabulario | alfabético | 4 | http://www.8020internet.com/camycam/glosario.htm |
|----------|---|---|----------|-------------|------------|-------|---|
| 11 | Diccionario de terminología montañera | español | en línea | Glosario | alfabético | 11 | http://www.hyparion.com/web/diccionari/dics/muntanya.htm |
| 12 | Glosario de montaña y senderismo | español | en línea | Glosario | alfabético | 12 | http://www.geo.ya.com/ascensiones/glosario.htm |
| 13 | Glosario de montaña | español | en línea | Glosario | alfabético | 17 | http://www.monipedia.com/glosario.htm |
| 14 | Vocabulario referido a la montaña (o el monte) | español | en línea | Glosario | alfabético | 10 | http://www.profes.net/rep_documentos/Catal%C3%A1n/Actividad_montana.PDF |
| 15 | Glossary of Swansea University Mountaineering Club | inglés | en línea | Glosario | alfabético | 12 | http://www.sumc.org.uk/resources/13416/glossary |
| 16 | Glossary of Rock Climbing and Mountaineering Terms | inglés | en línea | Glosario | alfabético | 12 | http://www.interclimbing.co.uk/glossary.html |
| 17 | Mountaineering Dictionary | inglés | en línea | Glosario | alfabético | 7 | http://www.abc-of-mountaineering.com/mountaineering-dictionary.asp |
| 18 | German - English Mountaineering Dictionary | alemán, inglés | en línea | Léxico | alfabético | 7 | http://www.tourplaner-db.de/cgi-bin/dictionary.cgi?Lang=E |
| 19 | Alpines Wörterbuch | alemán, inglés, italiano, francés | en línea | Léxico | alfabético | 9 | http://www.bergsport-kolb.de/download/alpines_woerterbuch_alle_sprachen.html |
| 20 | Terminología en la Normativa de Espacios Naturales Protegidos | español | en línea | Léxico | alfabético | 9 | http://www.fedme.es/uploads/contenidos/Documento/880/pdf/Terminologia+en+la+normativa+de+espacios+naturales+protegidos.pdf |
| ESCALADA | | LENGUA | FORMATO | TIPO | ORDEN | PUNTO | REFERENCIA |
| 21 | Climbing Glossary | inglés, alemán, francés, italiano, español | en línea | Glosario | alfabético | 10 | http://www.dwmorrison.com/dictionary.htm |
| 22 | Dicionário de Escalada em Idiomas | alemán, inglés, portugués, español, francés | en línea | Léxico | alfabético | 9 | http://luis-ave-lar.planetaclix.pt/idiomas/idiomas_a.htm |
| 23 | Vocabulario Alpinístico Multi-Linguaggio | italiano, inglés, francés, alemán, español | en línea | Léxico | alfabético | 9 | http://www.feibat.com/Alpinismo/Alpinismo_vocabolario.html |

| | | | | | | | |
|---------------------|---|---|----------------|-------------|--------------|--------------|---|
| 24 | Glossar - Kletterm | alemán | en línea | Vocabulario | alfabético | 8 | http://www.kletterm-md.de/home/kletterm/wissen/glossar?pn=2 |
| 25 | Lexicon - Kletterm | alemán | en línea | Glosario | alfabético | 10 | http://www.bergleben.de/kletterm/2130--lexikon.html |
| 26 | Lexikon von Fachbegriffen rund um's Kletterm | alemán | en línea | Vocabulario | alfabético | 8 | http://www.on-sight.de/diesdas/index.html?glossar.html |
| 27 | Lexikon der Kletterbegriffe | alemán | en línea | Glosario | alfabético | 11 | http://www.kletterm.de/service/lexikon/lexikon-der-kletterbegriffe.155105.5.htm?skip=1 |
| 28 | Kletter - Lexikon | alemán | en línea | Glosario | alfabético | 10 | http://www.kletter-bergsport.de/html/lexikon-kletterm.html |
| 29 | Climbing Glossary | inglés | en línea | Glosario | alfabético | 10 | http://climb.mountainzone.com/glossary_a_1.html |
| 30 | Climbing Glossary of Terms | inglés | en línea | Vocabulario | alfabético | 11 | http://www.adventuresportsonline.com/terms.htm |
| ESPELEOLOGÍA | | LENGUA | FORMATO | TIPO | ORDEN | PUNTO | REFERENCIA |
| 31 | A Lexicon of Cave and Karst Terminology with Special Reference to Environmental Karst Hydrology | inglés | en línea | Glosario | alfabético | 17 | http://www.karstwaters.org/files/glossary.pdf |
| 32 | Espeleocedario | español | en línea | Glosario | alfabético | 14 | http://www.uv.es/~rzaragoz/espeleo/PDF/espeleo_cedario.pdf |
| 33 | Glosario de Turismo Subterráneo | español | en línea | Glosario | alfabético | 8 | http://rutas.entropia.es/noticia/18.htm |
| 34 | Caving Definitions - A Glossary of Terms | alemán | en línea | Vocabulario | alfabético | 13 | http://wasg.imnet.net.au/glossary.html |
| 35 | The Caver's Multi-Lingual Dictionary | inglés, búlgaro, alemán, español, francés, croata, húngaro, italiano, portugués, rumano | en línea | Léxico | alfabético | 13 | http://www.uisic.us-speleo.org/lexintro.html |

| BICICLETA | LENGUA | FORMATO | TIPO | ORDEN | PUNTO | REFERENCIA |
|----------------|---|----------|---------------|------------|-------|---|
| 36 | español | en línea | Glosario | alfabético | 16 | http://www.bikeweb.org/em_diccionario.htm |
| 37 | inglés | en línea | Glosario | alfabético | 14 | http://www.sheldonbrown.com/gloss_n-o.html#nonturn |
| 38 | inglés | en línea | Vocabulario | alfabético | 11 | http://www.adventuresportsonline.com/mtbterms.htm |
| 39 | alemán | en línea | Glosario | alfabético | 12 | http://www.mtb-biking.de/lexikon.htm |
| 40 | inglés, alemán, francés, italiano, español, holandés, etc. | en línea | Léxico | alfabético | 4 | http://www.cyclingaroundtheworld.nl/tipsadvice/bikevoca.htm |
| 41 | español, alemán, esperanto, francés, inglés, polaco, portugués y ruso. | en línea | Léxico | alfabético | 12 | http://ainome.net/bici/art_02.es.html |
| 42 | español, alemán, esperanto, francés, inglés, polaco, portugués, sueco, etc. | en línea | Léxico | alfabético | 9 | http://bemi.free.fr/biciklo/index.php&lingvo2=dehttp://bemi.free.fr/biciklo/index.php?lingvo1=es&lingvo2=de |
| MUNDO ECUESTRE | | FORMATO | TIPO | ORDEN | PUNTO | REFERENCIA |
| 43 | español | en línea | Glosario | alfabético | 8 | http://euskalhorse.net/hipica/archfiles/glosario/a.php |
| 44 | alemán, inglés | en línea | Glosario | alfabético | 13 | http://www.reiterin.ch/1/lexikon.htm |
| 45 | alemán, inglés | en línea | Base de datos | alfabético | 18 | http://www.iim.fh-koeln.de/webterm/deuterm/pferd/pferd_d.htm |
| 46 | inglés, francés, español, alemán | en línea | Léxico | alfabético | 8 | http://jcboulet.net/searchfilesvc2e.php |

| | | | | | | | |
|---|---|---|----------|---------------------------|-----------------------------------|----|---|
| 47 | Pferd & Reiter - Wörterbuch | alemán, inglés, francés, italiano y español | en línea | Léxico | alfabético | 7 | http://www.pferdreiter.de/pferdreiter/woerterbuch.html |
| 48 | Diccionario multilingüe del mundo ecuestre | español, inglés, francés y alemán | en papel | Diccionario especializado | alfabético | 16 | Estévez Paz, C. 2004. <i>Diccionario multilingüe del mundo ecuestre</i> . Madrid: Editorial Slovento. 289 págs ISBN: 84-609-1585-9 |
| VUELO SIN MOTOR | | | | | | | |
| LENGUA | | | | | | | |
| 49 | Diccionario de la Lengua Española (Parapente) | español | en línea | Glosario | alfabético | 5 | http://www.enelviento.com/diccionariodelvuelo.html |
| 50 | Skydiving Glossary | inglés | en línea | Glosario | alfabético | 15 | http://www.dropzone.com/safety/resources/handbook/gloss2.shtml |
| 51 | Skydiving Information - Glossary | inglés | en línea | Vocabulario | alfabético | 9 | http://www.bimson.com/skydiving/glossary.html |
| 52 | WebTerm - Gliders | alemán-inglés | en línea | Base de datos | alfabético | 18 | http://www.iim.fh-koeln.de/webterm/deuterm/Skydive/Skydive_e.htm |
| 53 | English-German Sailplane Dictionary | alemán-inglés | en línea | Léxico | alfabético | 10 | http://home.exetel.com.au/mrssa/index.php?option=com_content&view=article&id=23&Itemid=43 |
| PIRAGÜISMO, WINDSURF Y ACTIVIDADES ACUÁTICAS | | | | | | | |
| LENGUA | | | | | | | |
| 54 | Paddlers Wörterbuch | alemán | en línea | Glosario | alfabético | 9 | http://www.siegmund-schneider.de/Kanu/Dateien/Woerterbuch.htm |
| 55 | Kanu Begriffe | alemán | en línea | Vocabulario | alfabético | 9 | http://www.kanu-info-isar.de/begriffe.htm |
| 56 | Deutscher Kanu Verband | alemán | en línea | Glosario | alfabético y búsqueda por término | 12 | http://www.kanu.de/go/dkv/home/service/glossary.xhtml |
| 57 | Paddling Terms | inglés | en línea | Vocabulario | alfabético | 8 | http://www.adventuresportsonline.com/glossary.htm |
| 58 | Paddlers' Dictionary | inglés | en línea | Glosario | alfabético | 9 | http://paddlenow.com/html/dictionary%20A-C.html |
| 59 | Diccionario de deportes acuáticos | español | en línea | Glosario | alfabético | 17 | http://deportes.glosario.net/deportes-acuaticos |

| | | | | | | | |
|--------------|---|--|----------------|---------------|--------------|--------------|---|
| 60 | Vocabulario del Windsurfista | inglés, español | en línea | Vocabulario | alfabético | 6 | http://www.windsurfsp.com/s_vocabulario.asp |
| 61 | Windsurf-Dictionary | inglés, alemán, polaco, italiano, francés, español | en línea | Léxico | alfabético | 16 | http://www.home.umk.pl/~skoter/diction.htm |
| 62 | Glosario sobre deportes náuticos | catalán, español, francés, inglés | en línea | Léxico | alfabético | 17 | http://www.termcat.cat/ca/Diccionaris_En_Linia/10/ |
| BUCEO | | LENGUA | FORMATO | TIPO | ORDEN | PUNTO | REFERENCIA |
| 63 | Tauchsport Lexikon | alemán | en papel | Glosario | alfabético | 13 | König, U. y Fuckner, I. 2005. <i>Tauchsport-Lexikon. DiveLex</i> . Berlin: Delius Klasing Vlg GmbH. 712 págs. ISBN: 3768817016 |
| 64 | WebTerm - Tauchen | alemán, inglés | en línea | Base de datos | alfabético | 18 | http://www.imm.fh-koeln.de/webterm/deuterm/tauchen/tauchen_e.htm |
| 65 | Diccionario multilingüe de buceo y actividades subacuáticas | español, inglés, francés, italiano, holandés, alemán | en línea | Léxico | alfabético | 13 | http://www.foreignword.com/glossary/plongeon/spn/defb.htm |
| 66 | Bienvenue dans le dictionnaire multi-langues de la plongée | francés, alemán, inglés, español, noruego, sueco | en línea | Léxico | alfabético | 9 | http://iozzo.chez.com/dico.htm |

ANEXO VI

A continuación, mostramos el material elaborado durante la fase de preparación del experimento de evaluación de la herramienta *Trandix* para el taller práctico diseñado.

CUESTIONARIO I

El objetivo de este cuestionario consiste en conocer su metodología de trabajo durante las traducciones.

1. Indique su código: _____

2. Edad: _____

3. Lengua materna: _____

4. ¿Qué tipo de fuente terminológica ha utilizado durante la traducción del texto facilitado? Ordénelos del 1 al (siendo 1 el recurso más utilizado).

- Diccionario/glosario monolingüe especializado (L1) _____
- Diccionario/glosario monolingüe especializado (L2) _____
- Diccionario/glosario bilingüe especializado _____
- Diccionario/glosario multilingüe especializado _____
- Diccionario visual monolingüe _____
- Diccionario visual bilingüe _____
- Sección Imágenes de un motor de búsqueda (por ej. Google) _____
- Búsquedas en motores de búsqueda (por ej. Google) _____
- Tesauro _____
- Corpus paralelo (textos originales y sus traducciones) _____
- Corpus comparable (textos originales en ambas lenguas) _____
- Base de datos terminológica _____
- Enciclopedia _____
- Wikipedia _____
- Listas de distribución _____
- Foros de internet _____
- Otros. Especifique: _____

5. En general, ¿cómo considera la ayuda que obtiene de las herramientas disponibles actualmente durante la traducción?

- Muy bueno Bueno Regular Malo Muy malo

6. ¿Le ha resultado fácil la traducción que ha realizado? Indique del 1 al 5 el nivel de dificultad que ha encontrado (siendo 1 muy fácil y 5 muy difícil). Justifique su respuesta _____

7. ¿Cuáles han sido los aspectos más difíciles de la traducción según su opinión?

- Falta de fuentes de información
- Terminología desconocida
- Gramática y sintaxis
- Formato del TO y/o del TM
- Adaptación de referentes culturales
- Otro. Especifique: _____

8. Señale los 5 términos de mayor dificultad:

9. ¿Considera que ha tenido suficiente tiempo para realizar la traducción?

- Sí No

10. ¿Ha encontrado fácilmente la información que buscaba? Indique del 1 al 5 el nivel de dificultad que ha encontrado (siendo 1 muy fácil y 5 muy difícil). Justifique su respuesta _____

11. ¿Cuál es su nivel de satisfacción con la traducción realizada?

- Muy satisfactorio
- Satisfactorio
- Poco satisfactorio
- Nada satisfactorio

CUESTIONARIO II

El objetivo de este cuestionario consiste en conocer su opinión acerca de la herramienta Trandix y las ventajas/ inconvenientes que presenta con respecto a la metodología anterior.

1. Indique su código: _____
2. ¿Le ha resultado fácil la traducción que ha realizado? Indique del 1 al 5 el nivel de dificultad que ha encontrado (siendo 1 muy fácil y 5 muy difícil). Justifique su respuesta _____.
3. ¿Considera que ha tenido suficiente tiempo para realizar la traducción?
 Sí No
4. ¿Cuál es su nivel de satisfacción con la traducción realizada?
 Muy satisfactorio
 Satisfactorio
 Poco satisfactorio
 Nada satisfactorio
5. ¿Cuáles han sido los aspectos más difíciles de la traducción según su opinión?
 Falta de fuentes de información
Terminología desconocida
 Gramática y sintaxis
 Formato del TO y/o del TM
 Adaptación de referentes culturales
 Otro. Especifique: _____
6. Señale los 5 términos de mayor dificultad: _____
7. ¿Le ha resultado fácil el uso de la herramienta *Trandix*? Indique del 1 al 5 el nivel de dificultad que ha encontrado (siendo 1 muy fácil y 5 muy difícil). Justifique su respuesta. _____
8. ¿Le ha ayudado la herramienta *Trandix* a realizar la traducción?
 Totalmente Mucho Bastante Poco Nada
9. ¿Qué aspectos de la herramienta destacaría positivamente? Indíquelos.

10. ¿Qué aspectos de la herramienta destacaría negativamente? Indíquelos.

11. ¿Ha encontrado fácilmente la información que buscaba? Indique del 1 al 5 el nivel de dificultad que ha encontrado (siendo 1 muy fácil y 5 muy difícil) _____
12. Con respecto a los recursos que utiliza normalmente durante una traducción (especificados en el cuestionario de la primera parte), ¿cree que la herramienta *Trandix* aporta algún beneficio? Indíquelo.

13. ¿Ha utilizado algún otro recurso además de la herramienta *Trandix* y los recursos que incluye para las búsquedas externas?
 Sí. Especifique: _____
 No.

14. Si tuviera que elegir entre la metodología que utiliza normalmente para realizar sus traducciones y la que le ofrece la herramienta *Trandix*, ¿cuál preferiría?

Trandix La metodología anterior

15. ¿Propondría algún tipo de cambio en la herramienta *Trandix*?

Sí No

Especifique: _____

16. Si la herramienta *Trandix* estuviera disponible, ¿la utilizaría para sus traducciones?

Sí. No

17. ¿Conocía alguna herramienta similar?

Sí No

Especifique: _____

18. Valore del 1 al 5 los siguientes aspectos de *trandix*, siendo 1 muy bajo y 5 muy alto

Facilidad de manejo _____

Facilidad de acceso a la información _____

Flexibilidad _____

Contenido _____

Interfaz _____

19. ¿Desea añadir algún comentario sobre la herramienta *Trandix*?

ESPECIFICACIONES PARA EL ENCARGO DE TRADUCCIÓN

I. FINALIDAD O ESCOPO

La traducción del texto origen se publicará en una página web de una empresa de turismo de aventura como parte de su información promocional y publicitaria sobre las actividades de aventura en Reino Unido a fin de promocionar las actividades de aventura entre el mercado turístico español en este país.

II. SITUACIÓN COMUNICATIVA

- Participantes:

- **Autor:** desconocido.
- **Emisor:** Empresa de aventura en Reino Unido que necesita la traducción.
- **Iniciador:** el profesor de la asignatura.
- **Destinatario final:** la traducción está destinada a turistas españoles que se interesen por la práctica de esta actividad de aventura en Reino Unido.

- **Medio:** página web de la empresa de aventura.

- **Código:** lengua española de España.

- **Contexto:** cada vez más turistas españoles se interesan por la práctica de actividades de aventura en Reino Unido.

- **Temporalidad:** el texto origen está publicado entre 2008 y 2010, por lo que no se debe modificar ninguna indicación temporal del texto.

III. ENTREGA

La traducción se entregará a la finalización del seminario práctico a través de una tarea creada para tal efecto en la asignatura virtual correspondiente a la asignatura en la que se realiza el taller práctico, que se encuentra alojada en el Campus Virtual de la Universidad de Málaga. Para la entrega, es necesario que cada alumno titule su archivo según la codificación que se indicará el profesor antes de comenzar el taller. Por último, se seguirán las mismas normas de presentación aplicables a la elaboración de trabajos de la asignatura *Traducción especializada inglés/español-español/inglés (b): textos científico-técnicos*.